Computer-Based Testing in Vocational Assessment and Evaluation: A Primer for Rehabilitation Professionals

Tina M. Anctil Peterman  
Portland State University, anctil@pdx.edu

Nancy J. Adams  
University of Wisconsin-Madison

Follow this and additional works at: https://pdxscholar.library.pdx.edu/coun_fac

Part of the Educational Assessment, Evaluation, and Research Commons

Let us know how access to this document benefits you.

Citation Details

This Article is brought to you for free and open access. It has been accepted for inclusion in Counselor Education Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.
Deb’s leadership, proceeded to shape a national plan that continued to evolve through new VEWA administrations and place us where we have come today.

Deb was equally as devoted to her work with the Commission on Certification of Work Adjustment and Vocational Evaluation Specialists (CCWAVES). Again, we shared common goals in our passionate feelings to ensure national certification for all. Deb served three stints as CCWAVES Commissioner, filling both terms for past Commissioners Dick Omang and Michael Rubin. Deb’s leadership roles included Secretary, Vice-Chair and Chair of CCWAVES.

During her tenure with CCWAVES, the national office changed physical locations three times. Supervising these transitions took a great deal of time and dedication and Deb was always completely enthralled in the process. Again, a number of initiatives evolved over these years including the establishment of a new certification: Certified Career Assessment Associate (CCAA). A national research study was conducted to determine minimal competencies for vocational evaluation and several publications were produced. CCWAVES embraced an opportunity to involve our vocational evaluation practitioners in Canada and establish a non-voting status on the board for such a representative. This has since evolved to a CCWAVES board representative who sits on the Commission, representing the interests and voices of our colleagues in Canada.

Deb gave much to the professions of vocational assessment and evaluation. She believed in the process. She valued the credential of Certified Vocational Evaluation Specialist (CVE). She moved forward, against the odds, and maintained success in her work with a public school system, as a private practitioner and as an expert witness for the Social Security Administration. While others lost their jobs in this business, Deb forged ahead. She will be missed in our professional community, in her contributions to the field and as a dear friend but her memory lives on in each of us. Our job now is to keep that flame alive in the profession and seek opportunities to ignite, engage and promote the value and purpose of vocational evaluation and assessment.

Computer-based Testing in Vocational Assessment and Evaluation: A Primer for Rehabilitation Professionals

NANCY J. ADAMS, M.S., CRC; DOCTORAL CANDIDATE & TINA M. ANCTIL, M.S., CRC; DOCTORAL CANDIDATE
University of Wisconsin-Madison
Dept. of Rehabilitation Psychology and Special Education

Abstract

Current technologies, including computerized assessments, assistive technology, and information/resource technology, are effective tools that offer the rehabilitation professional a variety of applications for vocational evaluation and work assessment. “The ability of vocational evaluators to effectively utilize computers to obtain useful information (e.g. availability of specific electronic devices, job accommodation techniques, job-matching) for vocational recommendations could ultimately affect the outcome goals achieved in the rehabilitation process” (Chan, Lam, Leahy, Parker, & Wong, 1989, p. 113). In order to appropriately use these technologies, rehabilitation professionals need to understand the issues surrounding the use of these tools (e.g., reliability, validity) and ethical concerns (e.g., equivalence, confidentiality) for their appropriate application to individuals with disabilities. The purpose of this paper is to provide an overview of the advantages and disadvantages of computerized assessment (including test administration, response recording/scoring, and data analysis/test interpretation) along with the ethical considerations. Additionally, information about how assistive technology and information technology can assist in vocational assessment and work evaluation is provided along with an appendix (A) of helpful Internet addresses.

Vocational evaluators who assess individuals with disabilities in order to facilitate career planning and employment must be aware of the concerns and issues surrounding computer use in vocational assessment and evaluation. Some argue that computer-assisted-technology, especially adaptive testing, has revolutionized the practice of assessment and is likely to be increasingly popular. Others, however, only see it as the transposition of the traditional paper-and-pencil test content onto the computer screen and the use of a keyboard as a replacement for the pencil (Ittenbach, Esters, & Wainer, 1997). It is crucial for vocational evaluators to have an understanding...
of the practical and ethical concerns and issues of computerized assessment in order to appropriately apply this technology to people with disabilities.

Our goal in preparing this article was to broaden the rehabilitation professionals understanding of three important technological tools that can assist rehabilitation professionals in vocational evaluations and work assessments: (1) computerized assessments, (2) assistive technology, and (3) information technology/links to other resources. Section one provides an overview of the issues surrounding computerized assessment and evaluation, including (a) test administration, (b) adaptive tests, (c) response recording and scoring, (d) reliability and validity, (e) data analysis and test interpretation, and (f) important ethical considerations (equivalence and confidentiality) that affect client outcomes. We have endeavored to present both the strengths and limitation of computerized assessment, including test interpretation, in order to objectively review the issues and concerns with this technology. Section two reviews the application of assistive technology resources and services to people with disabilities, including testing accommodations. Lastly, section three discusses the role of information technology and offers a list of resources (see Appendix A), which are valuable tools for career exploration, placement resources, and the identification of potential employers. These technologies are useful tools to affect successful client outcomes when used appropriately and ethically by informed rehabilitation professionals.

**Computerized Vocational Assessment and Evaluation**

Computer applications in testing and assessment have now been in use for almost four decades and has resulted in the creation of a substantial body of knowledge. The goal of this section is to explore the potential benefits and problems associated with computer-assisted testing and assessment. For more detailed descriptions of the specific options available for computer-assisted test administration, scoring, and report generation, please refer to Butcher (1995), Moreland (1987), Madsen (1986), and Sampson, Kolodinsky & Greeno (1997).

**Test Administration**

Both item response theory and computerized adaptive testing represent new horizons in psychological testing that have developed during the late 1970s and early 1980s. Computerized adaptive tests, also known as computer-assisted technology or cognitive ability testing, are generally called CATs. Computer assessment utilizes Item Response Theory (IRT), which is a family of mathematical models that can be used to describe the characteristics of test items. Based on his/her responses to a set of test items with known characteristics, mathematical models allow for the estimation of an individual's trait level. At the individual level, IRT describes how an individual is likely to respond to a test item as a function of his/her trait level, thus creating inferences from the test performance of an individual. The IRT test item parameters permit the creation of item banks designed for specific testing purposes from which a computerized test can draw its items for a given examinee. IRT provides a means by which different sets of items, as administered in a computer-based test, can be scored on a common scale (Itenbach et al., 1997). The computerized format makes it possible to individualize the tasks to be measured by selecting items from the item banks and creating different versions of the same test. Additionally, IRT provides for an estimation of the error of measurement of an individual's trait level estimate. This allows for the creation of adaptive tests and non-adaptive tests. Both are useful for measurement and classification applications.

**Adaptive Tests**

Adaptive tests tailor the difficulty of the items to the ability of the examinee being tested. The item selection is sequential, meaning each new item is selected on the basis of how the test taker performed on the previous item. An estimate of the examinee's skill level is made after each item response and is the basis for selection of the next item. This method gives rough approximations and requires a number of responses before the precision improves. The adaptive test prevents high-ability individuals from becoming bored by too many easy items, and low-ability test takers from becoming frustrated by too many difficult questions. With the non-adaptive test, the items are randomly selected from the item bank so that each test taker gets a different test form thus simplifying test security and repetitive testing.

Essentially, computerized adaptive tests meet the specific needs and abilities of each examinee, while offering efficiency and control over measurement precision. CATs are efficient because they obtain the most information about the examinee per item administered, resulting in fewer test items. For example, the paper-and-pencil test may require the examinee to answer 400 items, while the individually administered CATs may only require the testee to answer 200 items. Green (1984) found CATs to be more precise, allowing estimates to be moved up and down the scale until a sufficiently accurate ability estimate is found. Often, the accurate ability estimate was determined after only half of the number of items usually found on a conventional test were completed. The measurement quality has been established to be equivalent or better to that of a conventional test with the same number of items. Adaptive testing can reduce test length by an average of 50% with individual test length reductions of up to 80% without compromising measurement quality (Weiss & Vale, 1987). The Educational Testing Service publication, *Computer-Based Testing: From Multiple Choice to Multiple Choices*, provides additional insight on the advantages of computer-adaptive testing (Drummond, 1996).

Computerized adaptive testing, however, does have its critics. Helms (1997) warns examiners and examinees that psychometricians do not have informed models for investigating the possible racial, cultural, and/or socioeconomic limitations of CATs. Butcher (1987) also discussed other disadvantages of CATs. He found that some traditional instruments may not be easily adapted to computer formats and actually have different psychometric properties once adapted for computer administration.

**Response Recording and Scoring**

**Scales, norms, and score comparability.** Is there a difference in norms between computer administered tests and paper-and-pencil tests? Green (1984) concludes that conventional and computer-presented tests may yield scores that are not directly comparable. He calls for different norms for the two versions, or the creation of a new scale system that will permit statistical transformation allowing the same norms to be used. According to TheStandards for Educational and Psychological Testing (1999):

A clear rationale and supporting evidence should be provided for any claim that scores earned on different forms of a test may be used interchangeably. In some cases, direct evidence of
score equivalence may be provided. In other cases, evidence may come from a demonstration that the theoretical assumptions underlying procedures for establishing score comparability have been sufficiently satisfied. The specific rationale and the evidence required would depend in part on the intended uses for which score equivalence is claimed (p. 57).

This standard applies to computerized adaptive testing as well as alternate forms of paper-and-pencil administered tests. Additionally, it applies to test forms administered in different formats (e.g., large print, Braille, etc.) to accommodate people with disabilities. Some testing accommodations may only affect the dependence of test scores on capabilities irrelevant to the construct the test is intended to measure. For example, use of a large print edition assures that performance does not depend on the individual’s ability to perceive standard-size print. In such cases, relatively modest studies or professional judgment may be sufficient to support claims of score equivalence.

Reliability and Validity

Does the computer automation of testing affect the instrument’s reliability and validity? Given the economic potential of testing software, this question has tended to be investigated after software has been developed and marketed and is thus already in use. Although a number of researchers have challenged the validity of this software (Adams & Heaton, 1985; Matarazzo, 1986; Moreland, 1985) guidelines for studying their validity have frequently been ignored. However, in reviewing the studies on computer-based testing and conventional testing, Bunderson, Inouye, and Olsen (1989) conclude that the reliabilities of the computer and conventional tests are very similar.

It is important for the informed vocational evaluator to also understand the advantages and disadvantages of CATs. First, mastery levels can be demonstrated faster allowing the test to be terminated quickly once a test taker has reached the criterion level of mastery. This is often helpful in licensure or certification competency assessments. Second, CATs provide test items or questions at a level that is consistently appropriate and challenging to the test taker. Lastly, CATs allow for increased psychometric precision over a broad range of proficiency levels.

One must also recognize, however, that there are many potential sources of error when using a computer-based testing instrument. Common sources of error are data entry errors, programming bugs, poor conceptualization of the system, and the lack of research efforts that go into the development of the program (Chan et al., 1989). The substitution of tests obtained from one particular test (i.e. the General Aptitude Test Battery) with another similar test (i.e. the Differential Aptitude Test) may confound the outcome of the computer results.

Data Analysis and Test Interpretations

Butcher, Perry and Atlis (2000) stated,

Computers have played an integral role in scoring psychological tests virtually since their introduction, almost a half-century ago. Initially, computer-based applications involved scoring and data processing; however as their general use became more widespread, their potential advantage to the process of interpretation came to be recognized (p. 6).

Automated-assessment computer programs are already widely used to interpret test results but there is great variation in the quality of the computerized-based test interpretations (CBTI). CBTIs, or automated assessment computer programs, largely perform a look-up-and-list-out function.

Basically, a broad range of interpretations are stored in the computer from various test indexes and the computer simply sorts the stored information and retrieves the appropriate report components based on the scale score levels of the given test. Therefore, there is little actual decision making conducted by the computer. This stored information consists of a series of “if-than” statements that are written by expert clinicians. For example, if the F scale on the MMPI-2 is 12 points or more higher than the K scale, the computer program may generate a notation that the client could be malingering. The appropriateness of computer output largely depends on the reliability of the data input and the integrity of the program itself (Chan et al., 1989).

Can the computer produce a test interpretation that virtually replaces the complex cognitive process of the evaluation professional? Research in computer generated narrative reports for personality assessments has generally revealed that the interpretive statements contained in the reports are comparable to clinician generated statements. However, Moreland (1985) reviewed validity studies of computer-based interpretations and pointed to the fact that the conclusions drawn from the research must be evaluated in light of several problems. These problems include: (a) small sample sizes; (b) inadequate external criterion measures with which to compare the computer-based test interpretation statements; (c) lack of information regarding the report’s base-rate accuracy; (d) failure to investigate the internal consistency of the report’s interpretations; and, (e) several issues pertaining to the report raters (e.g., lack of familiarity with the interpretive system used, lack of expertise in the area of interest, and possible bias secondary to the theoretical orientation held by the rater (Butcher, Perry, & Atlis, 2000). Vocational evaluators who use CBTIs must be aware of the possibility of excessive generality of results as well as the high potential for misuse of the results due to their increased availability (Butcher, 1987).

Matarazzo (1986) criticized the interpretations’ failure to recognize test takers’ uniqueness, the tendency for the interpretations to be unsigned (ostensibly leaving no one directly accountable for the interpretations’ contents), and the inclination to be viewed as an end rather than as means to an end. The practitioner must be aware that computer-human interactions are confined to written material, and consequently that potentially critical nonverbal cues such as speech patterns, vocal tone, and facial expressions cannot be accounted for in CBTIs, as they currently exist. Therefore, each computer-based application needs to be evaluated carefully. There is a need for more current research on the accuracy of the information contained in the computer-based interpretative reports (Butcher et al., 2000).

Instruction in the use and avoidance of misuse of CBTIs is essential for all professionals who use them. Even though computer-based reports have been validated in some settings, this does not guarantee their validity and appropriateness for all applications. If a test has not been developed or validated in a particular setting, then computer-based applications in that context are not warranted.

Additionally, there are four important factors for the vocational evaluation professional to be mindful of when assessing computer-assisted test interpretations. First, the credentials of the system author should be carefully evaluated. Second, the documentation of the computer-based interpretation system should be thorough. Third, one should determine if a scholarly review of the system has been completed. Finally, the practitioner must tryout the system before administering the test and utilizing the CBTI with clients (Moreland, 1992).
Ethical Issues

Equivalence. Given that the basic objective of computer-based tests has been the transfer of paper-and-pencil tests to computers, a logical ethical and research question to ask is about the equivalence of procedures, particularly with the computer administration of test matters. Some studies have found no differences between traditional and computer-administered versions of tests (Drumond, 1996). However, the majority of the literature indicates that the comparability of computerized and standard administration of various measures appears to vary.

Kline (2000) discussed several factors specific to computerized test administration that could be instrumental in yielding results that are not comparable. These include (a) individuals who experience discomfort with computers and consequent awkwardness when dealing with them; (b) the type of equipment used and the nature of the test material; and, (c) respondents who are willing to reveal their true feelings to a computer than to a human being, which may lead to atypical results. This last point is particularly relevant when test item content deals with sensitive and personal information. Research has also shown that computer-administered tests can reveal elevated negative affect scores (George, Lankford, & Wilson, 1990) indicating greater anxiety in the test taker with computer-based procedures (Hedl, O’Neil, & Hansen, 1973, as cited in Meier, 1994). This may alter the test taker’s rate of omitting items (Mazzeo & Harvey, 1988), as well as increase his or her “faking good” responses (Davis & Cowles, 1989, as cited in Meier, 1994).

Green (1984) states that when a conventional test is transferred to computer presentation, there is no assurance that the test performance will be equivalent. He concludes that time limit differences and response differences could affect the total test score. Research has indicated that when speed tests are translated from pencil-and-paper to a computer mode of presentation, responses come much faster on the computer. Green & Green (1986) found a large mean difference in favor of the computer. Additional pertinent issues, such as the impact of ergonomic factors on the computer administration process, also warrant further consideration.

Confidentiality. Confidentiality is the issue most often mentioned in discussions of ethical problems regarding computers. Vocational evaluators need to ensure confidential data are restricted to appropriate professionals as the widespread availability of microcomputers, computer networks, and communication links between microcomputers and large mainframe computers using telephone lines may increase the possibility for unauthorized access to confidential information (Sampson & Pyle, 1983). Additionally, unwarranted electronic requests for releases of confidential information among rehabilitation and human service agencies are increasing with the advent of mainstream information technology. Vocational evaluators must be particularly aware of these potential threats to confidentiality and develop appropriate procedures to safeguard the security of client files stored on their computer (Chan et al., 1989).

Assistive Technology

The advent of computers has brought the world to many individuals with a variety of physical disabilities. The ability of persons’ with physical disabilities being evaluated to physically access and use the computer is fundamental to any form of computer-based assessment. With the constant changing and technological advances in computer adaptive software and hardware, many individuals with physical and cognitive disabilities can now access computers and thus utilize computer-based assessments.

Assistive technology is an empowerment approach that, when appropriately applied, increases the life and work choices for a person with a disability (Vocational Evaluation and Work Adjustment Association, 1997). Assistive technology, as defined by the Technology-Related Assistance for Individuals with Disabilities Act of 1994 (PL 103-218), is an assistive technology device or an assistive technology service. The role of assistive technology in assessment and vocational evaluation is summarized as follows:

The use of assistive technology, also known as rehabilitation technology, in the assessment process is often required in order to reach effective outcomes. A fundamental goal of the field of assessment and vocational evaluation is to assist individuals with disabilities to reach their maximum potential. For many individuals this potential will be severely restricted without the benefit of assistive technology. The use of assistive technology within vocational evaluation, assessment, and work adjustment to enhance the performance of individuals is essential in determining their functional capacities (Vocational Evaluation and Work Adjustment Association, 1997, p.1).

The application of assistive technology is described in four forms, including (a) site assessment for general accessibility, (b) modification of assessment tools and instruments, (c) use of technologies such as visual aids and computer adaptations, and (d) assessment recommendation which addresses further assistive technology issues. “Assistive technology resources and services should be integral components of all comprehensive vocational evaluation programs as well as vocational assessment and work adjustment services” (Vocational Evaluation and Work Adjustment Association, 1997, p.1).

Furthermore, adaptive devices, as well as other rehabilitation engineering devices such as augmentative communication aids, and environmental controls can all improve a person’s ability to complete tasks and function independently, as well as impacting vocational evaluation practices. Both the vocational evaluator and the individual served could benefit from these technological developments. Vocational evaluators may need to expand their knowledge of specific computer applications in vocational assessment as well as the availability of rehabilitation engineer aids in order to insure that the individual with a disability being evaluated has been given every chance for success in the evaluation process.

In addition to physical barriers to computer access, the vocational evaluator must also be aware of cognitive limitations to computerized testing. Many tests may not be suitable for individuals with cognitive limitations and it can be difficult for the examiner to make a determination of test suitability in many situations. An advantage of Item Response Theory is that test performance can be evaluated before a test is administered, using the test information function. A good instructional sequence will permit examinees to practice entering various kinds of responses and will incorporate remedial sequences of screens for examinees that are having problems. Some individuals may experience frustration with the forced response format, which will not permit skipping forward or backward through the questions. However, this computerized control also results in a “clean” data file with no missing data or invalid responses. It addition, it reduces the advantages of coaching on specific item content which is especially useful when individuals have to be retested, such as on minimum-level skills or essential-skills tests.

Information Technology

Information technology via the Internet, such as the O*Net, may also benefit the vocational evaluator as it complements the psychometric assessment process well. For example, once basic
aptitudes, abilities, and personality tests are completed, the job needs of the person with a disability must be accurately matched with his or her needs and abilities. (See appendix A for a list of occupational resources.) “Placement resources and job banks are numerous on the Internet which offers another valuable tool for career exploration, placement resources, and the identification of potential employers” (Patterson, 2000, p. 5). Thus, a decision to rule-out or pursue a particular career path may be made quickly. Patterson goes on to note, “As with any assessment, rehabilitation professionals need to use these tools selectively and as a basis for career exploration or counseling sessions. Their purpose is to promote self-knowledge, identify potential problems (Luciano, 1997), and enhance career counseling, not replace it” (Patterson, 2000, p. 5).

"A potential exists for incompetent use of computer resources by inadequately trained or overworked practitioners” (Sampson, 1986). For example, improper network access by vocational evaluators to computer-based test interpretations may be used to inappropriately compensate for a lack of vocational evaluator training or time. Vocational evaluators need to consistently monitor the knowledge and skill requirements of software on the network against their current competencies and only use software that is congruent with their capabilities. Vocational evaluators may use new software to expand their competencies, if adequate training and supervised experience opportunities are available (Sampson, Kolodinsky & Greeno, 1997).

Conclusion

Computer technology has had a profound impact on the testing, assessment and information gathering process. Most popular instruments used by practitioners today can be administered, scored and interpreted by computers. In another sense however, little empirical evidence is available to suggest that the computer is being used, on widespread basis, to do a better job of helping clients (Sampson, 2000). Whether or not the computer actually helps practitioners to substantially improve rehabilitation outcomes for clients depends more on practitioner attitudes, understanding, and skills in using this technology than on future advances in computer hardware and software. “Both the vocational evaluator and the individual served could benefit from these technological developments, however, the rehabilitation practitioner may need to expand their knowledge of specific computer applications in vocational assessment as well as the availability of rehabilitation engineer aids” (Chan, Lam, Leahy, Parker, & Wong, 1989, p. 110). It is essential for vocational evaluators to become knowledgeable consumers of computer-based testing products including understanding such key issues as test administration, response scoring and recording, data analysis and test interpretation, and interplay of assistive technology in computer access. In addition, the vocational evaluator must give thoughtful consideration and time to ethical issues such as the equivalence of computer-based tests as well as the confidentiality of the test results. Through understanding these issues and concerns of using computerized adaptive testing, the rehabilitation professional may utilize these computer resources in assisting individuals with disabilities to reach successful rehabilitation outcomes.

References


Patterson, J. B. (2000). Using the Internet to facilitate the rehabilitation process. The Journal of Rehabilitation, 66(1), 4-10.


