Work Evaluations: Critique of the State of the Art of Functional Assessment of Work

Craig A. Velozo

Key Word: work capacity evaluation

The area of work evaluation is fertile ground for future research and development by occupational therapists. Current evaluations of work range from standardized work evaluations associated with vocational rehabilitation to highly technical physical capacity and work capacity instrumentation and equipment often associated with sports medicine. In addition, methods used to identify pain and abnormal illness behavior add the psychosocial component of work evaluation. The limitations of most of the traditional approaches to work evaluations are their lack of focus on the actual work environments and on the meaning of work to persons. Occupational therapy can play an important and unique role in linking work evaluations to psychosocial and environmental variables and in formulating comprehensive theoretical models of work that should improve and refine present work evaluations.

Work may be mere source of livelihood, or the most significant part of one's inner life; it may be experienced as expansion, or as exuberant expression of self, as bounded duty, or as the development of man's universal nature. Neither love nor hatred of work is inherent in man, or inherent in any given line of work. For work has no intrinsic meaning. (Mills, 1973, p. 6)

This eloquent description of work is germane to occupational therapy theory and practice. Truly representative of occupation, work, as defined by Kielhofner (1983), is "a behavior which is motivated by an intrinsic urge to be effective in the environment...influenced by cultural tradition and learned through the process of socialization" (p. 136). Consistent with this definition, work is the principal medium for expressing effectiveness in society and a major part of socialization into adulthood. Harvey-Krefting (1985), in her historical review of the therapeutic use of work in occupational therapy, suggested that work is derived from the human need for mastery and self-actualization and the occupational nature of the person. It is questionable whether work evaluations reflect these assumptions in present-day occupational therapy practice. In this paper I present a critique of the state of the art in the evaluation of work, reflect on how the construct of work has been captured by available evaluations, and provide recommendations on the future development of work evaluations.

Development of Work Evaluations

Although work has been a tenet throughout the history of occupational therapy practice, our involvement in developing standardized work evaluations has been limited. Occupational therapy played a relatively important role in developing work evaluations in the 1960s and early 1970s. In Van Allen and Loeber's (1972) review of in-hospital evaluations of work performance of psychiatric patients, four of the seven evaluations reviewed were derived from the occupational therapy literature (Clark, Koch, & Nichols, 1965; Distefano & Pryer, 1970; Ethridge, 1968; Wolff, 1961). Work, as assessed by these early instruments, took on a broader perspective than the evaluation of work skills or performance. These occupational therapy evaluations also attempted to assess the more global aspects of work, such as interpersonal relations, creativity, and energy.

Since the 1970s, despite the advent of occupational therapy's dominant role in work-hardening practice, occupational therapists have played almost no role in developing standardized work evaluations (OT BibSys, July 1991). Although it is assumed that occupational therapists have been creating or modifying work evaluations within their clinics, there is no published psychometric support for their instrumentation practices.

Occupational therapy's limited role in developing work evaluations can be partially attributed to the dominant position of vocational rehabilitation services in work...
assessment. Until recently, work evaluation was generally considered the domain of vocational rehabilitation. Since the enactment of the Carl Perkins Vocational Education Act of 1984 (Public Law 94–142), which mandated the assessment of all students with disabilities in public schools, vocational evaluators have aggressively developed commercially available vocational evaluations (Botterbusch, 1987).

Vocational Evaluations

Vocational evaluations (work evaluations developed by vocational evaluators) generally provide testing in one or more of the following areas: (a) vocational interests, (b) aptitude and achievement, (c) motor skill, (d) work samples (see Botterbusch, 1987; Hursh & Kerns, 1988; Jacobs, 1991). Vocational interest inventories, exemplified by tests such as the Strong-Campbell Vocational Interest Inventory (Strong & Campbell, 1981), Career Assessment Inventory (Johansson, 1986), and the Ohio Vocational Interest Survey (The Psychological Corporation, 1983), are used to determine the vocation that most closely fits the person’s interests. Aptitude and achievement tests such as the Wide Range Achievement Test (Jastak & Wilkinson, 1984) and the USES General Aptitude Test Battery (United States Employment Service, 1983) focus on assessing academic abilities (e.g., reading, math, abstract reasoning, visual perception). Motor skill tests in vocational rehabilitation focus on assessing dexterity, coordination, and use of tools (Jacobs, 1991). Some common motor skill tests include the Purdue Pegboard, Crawford Small Parts Dexterity Test, and the Bennett Hand-Tool Dexterity Test (cited in Jacobs, 1991). Finally, a number of vocational evaluations have standardized work samples to assess the potential of a person to work. Work samples involve tasks, materials, and tools that are identical or similar to those used in an actual job (Hursh & Kerns, 1988). These tasks can be fairly simple, such as using a foot stapler, or can be as complex as servicing a diesel engine. Some common vocational evaluations that extensively use work samples are the Vocational Evaluation System, Vocational Information and Evaluation Work Samples, and the Vocational Interest Temperament and Aptitude System (cited in Botterbusch, 1987).

Vocational evaluators have clearly taken the lead in the development of work evaluations. In creating these evaluations, concerted efforts have been made to assess factors that appear critical to the construct of work. Developers have made headway in assessing the vocational interests of the person, the underlying skills involved in work, and work performance as simulated by working samples. In addition, several of these evaluations have at least some psychometric support in terms of reliability and validity.

In spite of this relative level of sophistication, vocational evaluations generally are applicable only to a circumscribed population. Since vocational evaluations have emerged from the mandate of the Carl Perkins Vocational Education Act, most have been designed to assess the potential worker, particularly the student who is disabled and who is making a first attempt to enter the work force. The content of vocational evaluations often is not applicable to the large proportion of clients who are attempting to return to their original job after an injury. For example, interest inventories may be unnecessary for the injured worker because he or she often intends to return to the original job. Performance tests, aptitude and achievement tests, or work samples are appropriate only if they are closely linked to the client’s original job or if the injured worker intends to make a vocational change. More importantly, none of these evaluations focuses on work from the unique perspective of the occupational therapist: the ability to actually engage in the process of doing meaningful work.

Evaluations for the Injured Worker

Unlike traditional vocational assessments, no single discipline can lay claim to developing evaluations for the injured worker. Although some of these evaluations can be traced to vocational rehabilitation services, others have their origins in psychology, behavioral medicine, and sports medicine. In addition, with the rapid advancement of work or industrial rehabilitation programs, new evaluations reflect the specific concerns related to rehabilitating the injured worker. Evaluations for the injured worker can be classified into one of two types, disability evaluations and functional capacity evaluations.

Disability Evaluations

A major emphasis in the area of work rehabilitation has been disability evaluation. Because work injuries, such as those that result in low back pain, often have no definitive pathology, there is frequently a question of whether the injury is the real cause of disability from work. A number of instruments have been used primarily to determine whether the injured worker is malingering or displaying symptom magnification. Although there is debate over the use and misuse of these terms, generally both imply that the person is consciously or unconsciously magnifying the disability for some secondary gain (e.g., monetary benefits, attention). Malingering and symptom magnification most clearly fit under the concept of “abnormal illness behavior,” defined by Pilowsky (1978) as “the persistence of an inappropriate or maladaptive mode of perceiving, evaluating, and acting in relation to one’s state of health” (p. 133). Although a detailed description of the tools used for identifying abnormal illness behavior is beyond the scope of this paper, a brief overview is presented (see Niemeyer, 1989 for a comprehensive review). Methods that have been commonly used to identify
Malingering or symptom magnification include pain drawings or pain rating scales, personality tests, self-reports, and measures of maximum voluntary effort.

Pain drawings and pain rating scales. Pain drawings and pain rating scales are widely used in chronic pain, physical therapy, and work hardening clinics as a method for clients to express the location, severity, and subjective characteristics of their pain (Niemeyer, 1989). For pain drawings, the client locates and draws areas of pain on the outlines of the human body. For pain rating scales, the client rates his or her pain on a numeric scale from 0 (no pain) to 10 (severe pain) or marks the level of pain on a 10-cm vertical line on which the bottom represents no pain and the top represents severe pain. Generally, these instruments suggest malingering or symptom magnification when the expressions of pain have little physiological basis or are out of proportion to the client's functional level (e.g., an indication of 9 or 10 on a pain scale when the client is capable of ambulation).

Personality tests. The personality test most often used to identify malingering is the Minnesota Multiphasic Personality Inventory (MMPI), a 550-item self-report questionnaire used to produce a personality profile of the client. Evaluation of the Hysteria, Hypochondriasis, and Depression scales, with the Depression scale slightly lower than the other two, produces the “Conversion V” profile, which is believed characteristic of symptom magnification, when the client shows little insight into his or her psychological problems (Mayer & Gatchel, 1988). Clients who have all three scales elevated, a situation called the neurotic triad, show more awareness of their symptoms and are better able to express their anxiety and stress (Mayer & Gatchel, 1988).

Self-reports of disability. A number of self-reports of disability are available to determine the client's perception of his or her pain-related disability. For example, in the Disability Questionnaire (Roland & Morris, 1983), the client checks any of 24 statements that describe the level of back pain and how it affects everyday functioning. The more statements checked, the greater the expression of disability. The Oswestry Low Back Pain Disability Questionnaire (Fairbank, Davies, Cooper, & O'Brien, 1980), using a slightly different methodology, also attempts to capture the person's perception of his or her level of pain and how it affects daily functioning. It consists of 10 sections, each of which represents a description of the pain condition or an area of functioning (e.g., ability to tolerate pain, sitting tolerance, standing tolerance). From a set of five statements under each of these categories, the client chooses the statement that best describes his or her condition. As with the pain scales, self-reports of disability are indicative of abnormal illness when the self-perceptions are inconsistent with a physiological explanation or when these perceptions suggest disability that is out of proportion to the client's functional level.

Evaluation of effort. Recently, considerable energy has been directed at identifying the injured worker who is malingering through the tests of consistency of effort or maximum voluntary effort. Underlying this evaluation procedure is the assumption that a person who is malingering will consciously or unconsciously fail to give maximum effort during strength testing. With instruments that accurately control and assess strength, lack of effort can be detected by inconsistency in performance (Niemeyer, 1989).

The Cybex Sagittal Strength Device has been extensively used to test maximum voluntary effort. This device, which isolates lumbar movement by restricting pelvic and lower extremity motion, measures strength in terms of the torque produced at varying velocities of lumbar flexion and extension. The device then graphically plots these as curves (torque over angular distance). Kishino and his colleagues (1985) have shown that maximum effort is displayed as consistent, replicable flexion and extension curves for repeated trials, even when pain and inhibitory mechanisms result in reduced forced output. Failure to show consistent flexion and extension curves suggests that the person is consciously restraining efforts.

Limitations of Disability Evaluations

Although the information gained from disability assessments is of interest to the occupational therapist, such tests provide data that are only indirectly related to a person's functional ability with regard to work. Tests of pain, psychological distress, self-perceptions, and effort, while providing us with descriptions of the client's illness experience, do not specifically assess the person's ability to work. Moreover, the intentional use of instruments to identify a client as malingering has serious ethical implications. In spite of efforts to lessen the negative connotations of labeling (e.g., by using symptom magnification instead of malingering) these terms continue to connote negative stereotypes that can have serious detrimental effect on the person (e.g., layoff, loss of injury compensation). Exaggerated pain symptoms, overstated perceptions of disability, or lack of maximum effort may represent complex personal reactions to the injury and disability and are more likely expressions of interactions of physical, psychological, and social factors (Matheson, 1991; Niemeyer, 1989). Although it is important to understand the clients' expression of their condition, negative labeling has little positive benefit for the person or the rehabilitation process in general, or occupational therapy functional capacity evaluations can be subdivided into

Functional Capacity Evaluations

Functional capacity evaluations can be subdivided into
physical capacity evaluations and work capacity evaluations. Physical capacity evaluations examine isolated parts of the body or functional units, which represent a component of the body that serves relevant biomechanical purposes. For example, the most studied functional unit is the lumbar region because it is critical for the transmission of loads, locomotion, and bending (Mayer & Gatchel, 1988). Work capacity evaluations, on the other hand, assess performance involving several functional units. For example, lifting involves transferring forces from the hands, upper extremities, shoulder girdle, lumbar region, and lower extremity to the foot (Mayer & Gatchel, 1988).

**Physical capacity evaluations.** Out of the sports medicine approach for low back pain treatment has emerged highly technical equipment to measure lumbar strength (e.g., Cybex Sagittal Strength Device, IsoStation B-200² [see Mayer & Gatchel, 1988, for an overview]). These devices constrain pelvic and lower extremity motion to isolate the lumbar area as a functional unit. Isometric and isokinetic tests of strength can be accomplished and performance can be measured in terms of peak torque, acceleration time, work, power consumption, and curve variability (Mayer & Gatchel, 1988). Because they control extraneous movement, these devices provide relatively reliable evaluation of lumbar strength.

**Work capacity evaluations.** Although procedures that assess isolated lumbar strength are effective for obtaining information about the injured body part, they are not structured to assess whole-body tasks common in the workplace. These work tasks include lifting, pushing, pulling, and carrying. The most studied of these is lifting because of its high correlation with industry's commonly reported injury, low back pain (Rowe, 1969). The most well-established technique for assessing lifting is isometric testing, which involves a static lift against a strain gauge (e.g., Cybex Liftask³). Isometric lift testing is used extensively as a means of employer selection for strenuous jobs and has been recognized by the National Institute of Occupational Safety and Health as a way to reduce occupational injuries (Mayer & Gatchel, 1988). This methodology has been extensively researched, but there is considerable controversy over its effectiveness in identifying workers who are prone to overexertion injuries in the workplace. Isoinertial lifting involves lifting loads over preset ranges (e.g., WEST 2 [WEST 2 User's Manual, 1985] and the Progressive Isoinertial Lifting Evaluation [Mayer et al., 1988]). Although isoinertial lifting most closely approximates actual lifting in the workplace, this methodology lacks sufficient studies on reliability and validity.

In addition to lifting, other common functional work activities such as standing, sitting, walking, crawling, climbing, pushing, and pulling are often evaluated. The Baltimore Therapeutic Equipment (BTE) Work Simulator⁴ is a popular instrument specifically designed to measure upper extremity performance. Different devices (e.g., knob, screwdriver, lever, steering wheel) are attached to this electrically controlled brake assembly (Niemeyer, Matheson, & Carlton, 1989). The resistance can then be adjusted to simulate specific work demand and thereby measure the amount of force produced during simulated tasks. Another popular device is the ERGOS Work Simulator⁵, a multi-station device used to measure dynamic and static strength, whole-body range of motion, work endurance, standing tolerance, and seated range of motion for a variety of work-simulated activities (Brandon & Snyder, 1989). Both of these devices, especially the ERGOS, have limited published studies to support their reliability and validity.

Finally, several proprietary work evaluation systems have been developed that require little equipment but instead emphasize therapist observations of work-simulated activities such as lifting, pushing, pulling, carrying, and so on (e.g., Functional Capacity Assessment of the Polinsky Medical Rehabilitation Center, KEY Functional Assessment) (cited in Jacobs, 1991). These evaluation systems include specialized training or equipment. Access to reliability and validity support for these evaluation systems is limited because of their proprietary nature.

**Limitations of Functional Capacity Evaluations**

Of the few studies that have included biomechanical factors in their predictive models of return to work, none has shown physical performance factors (e.g., trunk range of motion, trunk flexor-extensor strength, lifting ability) to be statistically related to return to work (Galagher et al., 1989; Milhous et al., 1989; Polatin et al., 1989; Velozo, Lustman, Cole, Montag, & Eubanks, 1991). Although further studies are necessary, these initial findings suggest that it is unwise to base assessments of return-to-work entirely on physical capacity and work capacity evaluations. Finally, Trombly (1993) and Mathiowetz (1993) stressed the need for occupational therapists to implement research that clarifies the relationship between components of functioning and occupational performance. Clearly, the available evidence suggests the need for this type of research specifically focused on the injured worker. This would enable the occupational

---

²Manufactured by Cybex, a Division of Lumex, Inc., 2100 Smithtown Avenue, Ronkonkoma, New York 11779.
³Manufactured by Isomedic Technologies, Inc., 328 Elizabeth Brady Road, Hillsboro, North Carolina 27279.
⁴Manufactured by Cybex, a Division of Lumex, Inc., 2100 Smithtown Avenue, Ronkonkoma, New York 11779.
⁵Manufactured by Baltimore Therapeutic Equipment Company, 74554 New Ridge Road, Hanover, Maryland 21076.
therapist to more carefully select appropriate functional capacity evaluations.

Recommendations for Improving Work Evaluations

Evaluations designed for potential workers and injured workers have generally originated from sound hypotheses or theories about the relationship of particular constructs to work. For example, interest inventories have been developed out of the assumption that success at work is related to a person's interests. Lifting ability has been hypothesized to be relevant to actual job performance. Although largely untested, these hypotheses carry considerable merit. A constructive strategy for the development of future work evaluations would be to build on these present hypotheses. The following approaches are proposed for future development of work evaluations: (a) improve measurement of constructs hypothesized to be relevant to work, (e.g., pain, personality, effort, physical and work capacity) (b) determine whether these constructs are actually related to work or return to work, and (c) develop new theoretical models and constructs related to work. Reliance on face validity is insufficient for determining an evaluation's adequacy in assessing a construct related to work. Analytic techniques such as Rasch analysis can be used to determine how well the items of an evaluation measure a single hypothetical construct. Through this technique one can determine whether particular test items misfit or do not measure the construct proposed, determine whether test items are redundant, and further obtain insight for creating test items that could be added to more accurately measure a particular construct (see Fisher, A. G., 1993, and Fisher, W. P., 1993, for further explanation and applications of Rasch measurement models). Other analytic techniques such as discriminant analysis, logistic regression, and confirmatory factor analysis are valuable statistical methods that can be used to determine the relationship of multiple variables to work or return to work. These procedures can provide insight into the relative importance of our constructs (e.g., physical capacity and work capacity) to work and assist us in developing occupational therapy models of work or return to work.

One of the most critical limitations of present work evaluations is their focus on the client with almost total disregard of the actual work environment. In addition to the person's aptitudes, performance, and skills, work also encompasses the person's environment. Environmental factors are likely to have a major effect on whether a person is successful at entering the work force or whether an injured worker returns to work. If there are major physical or psychosocial barriers in the work environment, it may be unlikely that the potential worker gets a job or an injured worker returns to work. Not only is the actual work environment important, but the worker's perception of this environment, or more broadly, what work means to a person, may be crucial to his or her success at entering the work force or returning to work. Recent studies have shown that job satisfaction is related to the report of back pain (Bigos et al., 1991) and that employer–worker relations may be related to disability claims (Wood, 1987). In addition, cultural norms may be critical in determining whether a person with back pain works or assumes a disabled role (Waddell, 1985).

Although the work environment and the meaning of work for the person have been generally overlooked by traditional work evaluations, several instruments may show promise in this area. The Position Analysis Questionnaire (PAQ) (McCormick, Jeanneret, & Mecham, 1975) offers a method to accomplish a structured job analysis by providing quantitative data on 194 job elements and 44 job dimensions. This instrument is well developed and is supported by extensive research. The Work Environment Questionnaire (Moos & Insel, 1974) is a self-report that measures the worker's perception of his or her work environment on three conceptual scales: relationship, personal growth, and system maintenance and change. This instrument is supported by a large normative sample and shows good psychometric properties. Finally, the Worker Role Interview (Velozo, Kielhofner, & Fisher, 1990), which is based on the Model of Human Occupation (Kielhofner, 1985), is a semi-structured, scoreable interview designed to be used as the psychosocial and environmental component of the initial rehabilitation assessment for the injured worker. Although this instrument is only in the early stages of development, it represents an initial attempt to identify psychosocial and environmental variables that may affect the injured worker's ability to return to work.

In light of recent research, traditional approaches to the evaluation of work are clearly limited. The above assessments represent attempts to expand this circumscribed perspective of work. Use of these instruments may serve as catalysts to expand our occupational therapy theories and consequently develop new evaluations that represent a more comprehensive view of work.

Conclusion

This overview of work evaluations suggests that, although there has been substantial development, the limitations of these evaluations need to be recognized. These limitations can appear in the form of restricted application of evaluations (i.e., vocational evaluations generally being designed for the student with disabilities), or in the intent of the assessment procedure (i.e., the identification of malingering). In addition, we must be cautious of the lure of high-tech performance evaluations or work capacity (component) evaluations, because there is little evidence that these evaluations accurately predict work outcome (occupational performance). Finally, it is important to
recognize that our present work evaluations provide only a limited perspective of work. Few of these evaluations are designed to assess the work environment or the meaning of work to the person.

The area of work evaluation provides fertile ground for future development within occupational therapy. Improved measurement of work-related constructs, such as work capacity and worker role identity, and the determination of the relevance of their measurements to actual work performance are needed. In addition, the formulation of comprehensive theoretical models of work should do much to encourage the development of evaluations that examine work from a broader perspective.

Recent studies that suggest the relevance of psychosocial variables to work are consistent with the education and theoretical perspectives of occupational therapists. These findings may serve as an incentive to remove the artificial boundaries that have existed between physical disability and psychosocial occupational therapy and lead to the development of work evaluations to assess functional capacity, psychosocial elements of work, and work performance. Occupational therapy practitioners have been among the leaders in the development of industrial rehabilitation programs and may serve as important resources for instrument development in this field. In addition, with the recent focus on measurement by the American Occupational Therapy Association and the American Occupational Therapy Foundation, occupational therapy researchers are gaining and sharing expertise in test development. The collaboration of these expert work clinicians with measurement experts could represent an effective formula for the development of work evaluations that will be both sound in measurement and clinically practical. ▲

Acknowledgments

This manuscript is based on a paper presented at the Symposium on Measurement and Assessment: Directions for Research in Occupational Therapy at the University of Illinois at Chicago, October 16–18, 1991. The symposium was jointly sponsored by the American Occupational Therapy Association, the American Occupational Therapy Foundation, and the Occupational Therapy Center for Research and Measurement at the University of Illinois at Chicago.

References


Mills, C. W. (1973). The meanings of work throughout his-


