A Critical Review of Functional Capacity Evaluations

The role of functional capacity evaluations (FCEs) appears to be increasing as employers and insurers rely more heavily on them for decision making. To meet credibility requirements, and the American Physical Therapy Association's standards for measurement and documentation, all FCE suppliers need to validate and refine their systems. This article provides information that can be used to make informed decisions in the selection of an FCE and in functional assessment practices. Features of well-designed FCEs are discussed. Ten well-known FCE systems are analyzed according to these features and other common characteristics. Current issues such as the qualifications of the evaluators, reliability and validity, length of assessments, projection of endurance to 8-hour workdays, standards of practice, safety protocols, and behavioral assessment and management strategies are discussed. [King PM, Tuckwell N, Barrett TE. A critical review of functional capacity evaluations. *Phys Ther.* 1998;78:852–866.]

Key Words: *Disability, Evaluation, Functional capacity evaluation, Physical capacity.*

Phyllis M King

Nicola Tuckwell

Tanya E Barrett

Functional capacity evaluations (FCEs) have become part of practice in work injury prevention and rehabilitation. These tools are supposed to define an individual's functional abilities or limitations in the context of safe, productive work tasks.[1,2] A series of test activities is usually administered to measure whether an individual has the ability to meet the required job demands. Ideally, FCEs are used following an injury or illness to assist with cost-effective vocational rehabilitation.[3]

A multitude of factors are currently shaping the growth and evolution of FCEs. To contain health care costs, some efforts are under way to reduce the amount of time and money spent on administration of FCEs. Regulatory agencies such as the Occupational Safety and Health Administration are promoting the use of FCEs in an effort to ensure employee health and safety in job placement.[4] The Americans With Disabilities Act (ADA) places an emphasis on identifying an individual's physical abilities and limitations for employment and accommodation considerations.[4] In litigation cases, FCEs have become critical for the determination of whether a claimant has wage-earning potential based on physical abilities testing.[5] State welfare reforms and the Social Security Administration are opening new markets for FCE providers as they seek a streamlined process for disability determinations.[6]

The FCE has become a widely used tool for determining a person's readiness to return to work after injury, for performing pre-employment and post-offer screening, for making disability determinations, for setting goals and planning treatment for industrial rehabilitation, for monitoring progress throughout industrial rehabilitation, and for determining case closure.[7]

Insurance companies and the legal profession appear to rely heavily on data acquired from FCEs to make important reimbursement and return-to-work decisions.[1] The purpose of this article is to provide information that can be used to make an informed decision in the selection of an FCE system. Features of well-designed FCEs are discussed. Based on survey responses obtained from the designers of 10 well-known FCE systems, comparisons are made according to common characteristics. Current issues surrounding this area of practice also are discussed.

Differences Among FCEs

A great number of FCEs currently are available and in use.[8] Table 1 identifies 10 of these FCEs and compares their general characteristics. Differences among the various approaches to FCEs include variations in the number of measurements obtained, degree of standardization, clarity of the concepts and underlying theories, variety in choice of measuring instruments, adequacy of measurement for certain injury groups (eg, lifting assessments used with individuals with low back pain, use and availability of normative data, and ability to predict return to work or recurrence of injury).[9]

Tramposh[5] attempted to classify FCEs into 2 categories: the controlled FCE and the uncontrolled FCE. Table 2

PM King, PhD, OTR, FAOTA, is Associate Professor, Occupational Therapy Program, PO Box 413, University of Wisconsin–Milwaukee, Milwaukee, WI 53201 (USA) (pking@uwm.edu). Address all correspondence to Dr King.

N Tuckwell, BAppSc(OT), PGradDipHlthSc(Curtin), is Senior Occupational Therapist, Commonwealth Rehabilitation Services, Lake Joondalup Unit, Wangara, Western Australia, Australia.

TE Barrett, BAppSc(OT), PGradDipHlthSc(OHS), MAppSc, is Fellow, School of Occupational Therapy, Curtin University of Technology, Perth, Western Australia, Australia.

Table 1.
A Comparison of Functional Capacity Evaluations (FCEs)[a]

FCE System	Designer	No. of Years Available	Full or Module Format	Length of Assessment	Length of Report	Validity and Reliability	Peer-Reviewed Published Research	Standardized Instruction Manual
Blankenship	Keith Blankenship	13	. . .[b]	2.5–4 h	30 min with software	Based on published medical research	No	Yes
IWS	Susan J Isernhagen	10	Full format, modified available for specific requests	5 h over 2 d; modified—3 h over 2-part period	20–45 min to dictate; software available	No	Yes	Yes
ErgoScience (PWPE)	Deborah E Lechner	4	Both	3–4 h	30 min (less time with software)	Yes[10]	Yes	Yes
WEST-EPIC (Cal-FCP)	Leonard Matheson	4	Both	All 5 components=2 h	Cal-FCP=5 min; WEST-EPIC=15 min; software available	Validity and reliability tests on lifting capacity only[61]	Yes	Yes
WorkAbility Mark III	GC Heyde and J Shervinton	6	Both	2–4 h	2 h with software package	No	Yes	Yes
WorkHab	David Roberts and Sam Bradbury	1.5	Both	2–3.5 h	1.5–2 h; software available	No	No	Yes
AssessAbility	Michael Coupland	1	Both	2 h	. . .	Based on methods-time measurement standards and research	No	Yes
ARCON	Dana Rasch	10	Both	<2 h	. . .	No	Yes	Yes
ERGOS	Work Recovery Inc	10	Both	4 h	40 min	Reliability and validity studies performed[c,d]	Yes	Yes
Key	Glenda Key	17	Both	3.5–4 h	15–20 min with software	Yes[e]	No	Yes

Table 1.
Continued

FCE System	Norm or Criterion Referenced	Portable	Training Costs[f]	Train-the-Trainer Options	Train-the-Trainer Costs[c]	Equipment Costs,[c] Including Software	Licensing Costs[c]	Ongoing Costs[c]
Blankenship	Both	Yes	One professional included in purchase price; additional staff trained for $1,500	Yes	Cost depends on time required	$19,950–$75,000	. . .	Data access charge of $25 per evaluation
IWS	Criterion (some normative data available for specific test items)	. . .	$3,500 for training (included in cost of FCE)	Yes	$550 per person or a group of 5–9 for $2,750 (total)	$1,000–$2,000	$8,000 ($4,500 for manual and $3,500 for training)	None
PWPE	Criterion (some hand dexterity normative data used)	Yes	$1,800 for first person; $400 for additional (maximum of 6 per group)	Yes	Approximately $600 to train trainer; royalties of $300 per person trained externally go to ErgoScience	$3,500 without software, $5,500 with software	$3,000	Pack of 50 evaluations=$75 + shipping
WEST-EPIC (Cal-FCP)	Predominantly criterion, WEST-EPIC 1 (lift capacity) is normative	No	For WEST-EPIC 1 (lift capacity), included in equipment price for 2 staff, training in San Diego only; training for remaining modules via manuals and videos	. . .	. . .	$23,530	Included in equipment costs	None
WorkAbility Mark III	Criterion	Yes	$1,711	Yes	No extra cost if program is accredited; accreditation fee $1,120 (for 2 people)	$10,400	None	None
WorkHab	Criterion	Yes	$850 per person	Negotiable	. . .	Computer package=$1,500	$75 yearly licensing fee	None
AssessAbility	Criterion	Yes	None	Yes	None	$400 plus computer	None	$30 per client
ARCON	Both	Yes	None	Yes	None	$40,000	None	None
ERGOS	Criterion	Yes	Included in purchase price	Yes	None	$20,000–$100,000; depends on number of modules	Included in purchase price	None
Key	Both	Yes	$400 per day per individual	Yes	$15,000	$20,000, includes software	None	$60 per assessment; $10 per job placement analysis

[a] Responses from a survey completed by FCE designers in March 1997.
[b] Ellipsis indicates no comment.
[c] Cooke C, Dusik L, Menard M, et al. Relationship of performance on the ERGOS Work Simulator to illness behavior in a Workers' Compensation population with low back versus limb injury. *J Occup Med.* 1994;36:757–762.
[d] Dusik L, Menard M, Cooke C, et al. Concurrent validity of the ERGOS Work Simulator versus conventional functional capacity techniques in a Workers' Compensation population. *J Occup Med.* 1993;35:759–767.
[e] Waite H. *Use of a New Physical Capacities Assessment Method to Assist in Vocational Rehabilitation of Injured Workers.* Thesis submitted to the Faculty of the Graduate School of the University of Colorado for the degree of Master of Science in Public Health.
[f] Monetary values are in US dollars. Costs are determined from May 1997 data.

Table 2.
Types of Functional Capacity Evaluations[a]

	Actual Simulation	Predicts Ability
Controlled	Authors' instruction in process Tests are actual simulations of physical demands	Authors' instruction in process Tests simulate components of physical demands and predict physical demands
Uncontrolled	Instruction may be by test authors, self-taught, or by others who have been instructed Tests are actual simulation of physical demands	Instruction may be by test authors, self-taught, or by others who have been instructed Tests simulate components of physical demands and predict physical demands

[a] Reprinted from Tramposh[5] with permission from Hanley & Belfus Inc, Philadelphia, Pa.

Table 3.
Advantages and Disadvantages of Various Types of Functional Capacity Evaluations[a]

Type	Advantages	Disadvantages	Examples
"Controlled" (Actual Simulation)	Reasonable chance of inter/intra-rater reliability Content validity easy to show Easy to use in court due to standard protocol Patients see job-relatedness	Evaluator lacks flexibility to simulate a specific job	Isernhagen FCE
"Controlled" (Predicts Ability)	Best change of inter/intra-rater reliability Easier standardization of test Easy to use in court due to standard protocol Easiest to research due to control and standardization	Evaluator lacks flexibility to simulate a specific job Relies on construct validity More difficult for patients to see job-relatedness Tends to be the most expensive (for therapists to purchase)	ERGOS system
"Uncontrolled" (Actual Simulation)	Content validity easy to show Flexibility to simulate a specific job Clients see job-relatedness Most accessible for therapists	Inter/intra-rater reliability difficult to control More difficult to use in court—lack of protocol control Hardest to research due to lack of standardization and protocol control	Blankenship FCE
"Uncontrolled" (Predicts Ability)	Easier standardization of test	Inter/intra-rater reliability difficult to control Relies on construct validity More difficult to use in court—lack of protocol control More difficult for patients to see job-relatedness	Isometric/isokinetic equipment

[a] Reprinted from Tramposh[5] with permission from Hanley & Belfus Inc, Philadelphia, Pa.

details the 2 categories, and Table 3 describes some of the advantages and disadvantages of both categories. The comparison of uncontrolled versus controlled FCEs deals with issues such as type of training for FCE administration, degree of work simulation, ability to alter the test design, and generic versus job-specific testing. The classification into controlled and uncontrolled categories focuses on different aspects of these issues, but the classification is arbitrary and overlapping. A new classification system is needed, one that further defines the methodological differences among the various types of FCEs.

The commercially available FCEs and those developed within individual clinics all share the common goal of attempting to measure work-related functional performance objectively. Whether they accomplish this objective can be answered only with research. The differences in FCEs revolve around the way they assess cooperation and sincerity of effort and safety, determination of end points for stopping clients during performance of manual material-handling tests, use of isometric testing, training processes, degree of work simulation, ability to alter the test design, generic versus job-specific testing, expense of equipment, use of algorithms for scoring, methods of projecting endurance to an 8-hour workday,

degree of standardization, evidence of reliability and validity, and so on.

Choosing an FCE

Although the practice of administering FCEs for the determination of an individual's physical capacities has been around for over a decade, research to justify the use of FCEs is lacking. Little is known regarding the reliability and validity of data obtained with FCEs. In addition to not being standardized, many of the testing systems lack comprehensiveness and objectivity in data collection.[1,5,10]

Given the lack of scientific evidence to support the use of the various FCEs, the considerable financial investment when purchasing an FCE, and the magnitude of medical, legal, and financial implications arising from the administration of FCEs, the decision of which FCE to choose becomes crucial to the provider's practice.

A Well-Designed FCE

In our view, functional capacity testing requires the evaluator to use tests that are most appropriate for a given client. We believe there is no single most appropriate test for any one client or for any one assessment situation, because no one assessment can provide 100% of all the answers concerning work injury and return to work.[3,11,12]

A clear understanding from the referral source regarding the purpose of the FCE is essential in choosing an FCE. In evaluations where return to work is the major focus, a job analysis should be performed to determine the tasks required for the job. The results of the FCE can then be compared with the job's physical requirements. In cases where an individual is being placed in a new job, we believe that a more comprehensive and generic assessment is needed. A range of physical demands must be tested to yield as much information as possible in order to consider a variety of job possibilities. We do not have data, however, to suggest how much information is necessary to infer from an FCE whether a person can safely function on the job. In our opinion, if disability assessment is the purpose of the FCE, then the evaluation often can be more limited in scope and correspond to the information requested by whoever is determining the level of disability.

Although the application of FCE testing is dependent on the purpose and setting, we contend that every attempt should be made to follow standardized procedures when available. Jobs may differ from one FCE to another, but consistency should be adhered to in the methods for designing and conducting the assessment. We lack data, however, to demonstrate that such standardization leads to reliability.

The National Institute of Occupational Safety and Health (NIOSH) *Work Practices Guide to Lifting* cited by both Miller[3] and Matheson[13] and the American Physical Therapy Association's (APTA) *Standards for Tests and Measurements in Physical Therapy Practice*[14] indicate that therapists need to consider the following questions:

- Is the test safe to administer?
- Does the test give reliable results?
- Is the test valid specific to job requirements for predicting a safe level of work?
- Is the test valid specific to work-related abilities?
- Is the test practical to administer?
- Does the test predict the risk of future injury or illness?

For FCEs, the principles of scientific measurement should be considered, as they are for any other test. Functional capacity evaluations, therefore, should yield reliable and valid measurements. Standardization is one way FCE developers attempt to enhance reliability. Additionally, an FCE should be comprehensive, feasible, safe to administer, and have flexibility for job-specific testing.

Standardization

Standardization refers to the development of a clear set of procedures for administering and scoring tests. These procedures should be written in an easy-to-use instruction manual that describes the general approach and philosophy for the development and administration of the test and the specific methods for administering all items in the evaluation.[14] The instructions, task demonstrations, subject placement, and data collection and analysis should be documented and followed each time the evaluation is administered and should not change, regardless of the individual administering the assessment.[15,16]

We believe that each task should be defined and described according to the equipment needed and the procedures to be followed. Verbal instructions for administration of the test, in our opinion, are essential to minimize examiner bias. All terminology should be defined to avoid misinterpretation of meaning. Specifications for scoring should be clear and easy to interpret to promote accuracy in recording of results. If the examiner needs to extrapolate or project data, the manual should provide instructions for accomplishing this task.[16]

Some assessments, such as manual dexterity tests, have standardized procedures and equipment by which normative data have been established. For example, the Minnesota Rate of Manipulation Test[17] measures the ability to make skillful, controlled arm-hand manipulations of larger objects. Norms based on older, unem-

ployed adults and on young adults are available in percentile tables. Caution, however, should be used when comparing an individual's performance with normative data, because the ADA,[18] for example, prohibits using this method to make decisions regarding return to work. Decisions need to be made based on the functional abilities and limitations of each individual as they relate to his or her job situation. Denying a job to an individual with a disability based on data that compares his or her functioning with that of the general population is illegal under the ADA.

There are procedure manuals for all 10 FCEs reviewed for this article.[19–28] The degree, however, to which they incorporate the components described varies.

Objectivity

The term "objective" is used to indicate a measurement that has a degree of reliability and is relatively free from examiner bias.[29] Objectivity in testing can be promoted when the procedures, variables for observation, and scoring system are operationally defined. Visual observations can be objective if operational definitions and scoring criteria are applied.[7]

Reliability

The importance of reliability and validity of FCE measurements cannot be overstated. If an FCE measurement does not have established reliability, test results could be different with each administration. Without validity testing, there is no way of knowing whether the results are accurate.

Reliability refers to consistency in measurement. Interrater and test-retest reliability have been purported to be the 2 most important forms of reliability in FCE testing.[1,30]

Interrater reliability refers to the ability to achieve similar scores on an evaluation when administered by different evaluators. In industrial rehabilitation programs, one therapist may administer an FCE to a client prior to treatment, whereas another therapist may administer an FCE following the rehabilitation program. Interrater reliability helps to ensure that the differences between the 2 FCE scores were not due to the differences in the therapists' test administration and scoring.

Interrater reliability is tested by having 2 or more administrators give a test independently to the same group of subjects. The administrators' scores are then compared by calculating correlation coefficients to determine how closely the scores are related.

Test-retest reliability or *intrarater reliability* refers to the stability of a score derived from one administration of an FCE to another when administered by the same rater. Variables potentially affecting the results of this type of measurement include the time between the 2 evaluations, the stability of the client's physical condition, and the treatment received by the client between the FCEs. For the FCE, it is important to determine a time interval that is long enough to minimize a learning effect from the first test for both the client and the examiner but short enough so that the client's medical condition will not have changed substantially between tests.[10] Most of the reliability studies of FCEs were performed on subjects without disabilities.[1] The reliability of FCE measurements on persons seen in clinical practice, therefore, is unknown. Caution should be exercised when applying these results to various disabilities, because conclusions based on such extrapolations may be highly misleading.[1,14,31]

From the systems reviewed for this article, only the Physical Work Performance Evaluation (PWPE)[19,*] and the WEST-EPIC[20,†] (lifting-capacity section only) FCEs have been examined for intrarater and interrater reliability with results published in peer-reviewed journals.[10,32] Some components of the Blankenship FCE[28,‡] are based on methods developed and studied by other investigators.[33–47] Some of these tests, however, were developed for persons without injuries. No published peer-reviewed research documenting the reliability and validity of measurements obtained with the Blankenship FCE is currently available. This can also be said for the WorkAbility Mark III,[§] Isernhagen Work System,[‖] ARCON,[#] Key Method,[**] WorkHab,[††] AccessAbility,[‡‡] and ERGOS[§§] systems.

Validity

A score is considered valid if it measures the properties that it purports to measure and can be used to make inferences. In FCE testing, this means that the score predicts real-world function. The interpretation of the test results should predict or reflect the client's performance in a target work setting or predict an overall level of work if there is no target work setting.[11] Validity is an essential requirement for all measurements.[16,48] Establishing the validity of a physical work performance measure is more difficult than establishing its reliability

* ErgoScience, 3929 Glenwood Ave, Birmingham, AL 35222.
† Employment Potential Improvement Corporation, PO Box 3897, Ballwin, MO 63022.
‡ The Blankenship System, 3620 Eisenhower Pkwy, Suite 7, Macon, GA 31206.
§ WorkAbility Mark III, Unit 22/6-8, Price St, Ryde, New South Wales 2112, Australia.
‖ Isernhagen Work Systems, 2202 Water St, Duluth, MN 55812-2145.
Applied Rehabilitation Concepts, 309 McLaws Cir, Suite F, Williamsburg, VA 23185.
** Key Method, 1010 Park Ave, Minneapolis, MN 55404.
†† WorkHab Australia, PO Box 1761, Bundaberg, Queensland, 4670 Australia.
‡‡ MediSys Rehabilitation Inc, 1801 N Lamar St, Suite 202, Austin, TX 78737.
§§ Work Recovery Systems Inc, 2301 S Friebus, Unit 1, Tucson, AZ 85713.

due to the need for a gold standard for comparison with predictions.

There are various ways to evaluate the validity of FCE measurements. The most common, *content validity*, asks whether a representative sample of the components of the task in question are incorporated into the assessment.[14,48] For example, content validity of FCE measurements can be implied based on the clinician's professional judgment as to the degree to which the test measures the job's demands, but this approach is considered relatively weak and provides no more than a logical argument for validity. We believe a thorough and systematic job analysis is necessary to accurately determine job demands. If an FCE is not job-specific, content validity can be established by covering all 20 physical demands of work described by the *Dictionary of Occupational Titles* (DOT).[39,40] Because content validity deals with theoretical arguments, it should not be used as the sole basis for suggesting that an FCE is valid. From an examination of the existing FCEs, it appears that several FCEs have good content validity when judged by the US Department of Labor's 20 demands.[19,21,22,24,26]

Criterion-related validity deals with whether there is evidence that a measure can be used to make an inference. In the case of FCEs, the most common inference is whether the measure predicts an individual's performance in specified activities. *Concurrent validity* deals with the correctness of an inference at the time of the measurement and may be used to determine whether a client needs therapy. If FCE scores distinguish between those clients who are currently unable to perform at a certain level of physical ability and those clients who can perform at this level, then the FCE is said to have good concurrent validity.

Predictive validity refers to a measure's ability to predict future performance. In the case of an FCE, if a client's scores predict a certain level of performance and the client successfully returns to work at that level, the FCE is said to have good predictive validity. Defining and measuring the criterion by which clients are compared becomes the greatest challenge to this test. External factors such as the work environment, the work pace, and the work schedule may affect the client's level of performance and ability to return to work.

In addition, for individuals with low back pain, we believe that ability or disability at any age should be evaluated relative to their current aspirations for a "normal" life. Thus, individual, cultural, and economic factors must be taken into account.[49]

Responses obtained from the 10 FCE designers chosen for this article indicate publication of a validity study in a refereed scientific journal exists only for the PWPE.[10] Lack of peer-reviewed publications for the FCEs reporting the completion of validity studies leaves open the question of whether the FCEs are acceptable. With the exception of the PWPE, the FCEs reviewed for this article do not provide the validity studies that are seen as the prerequisite for demonstrating that a measure is credible.

Components of FCEs

Data Gathering

Table 4 lists the major components of FCEs. Most methods begin with collecting a client's medical, social, and vocational history. This information assists with determining a client's perceptions of his or her own abilities.

Data may be obtained through record reviews, a client interview, or a self-administered client questionnaire. A medical history can identify conditions that may indicate that an FCE is contraindicated or that require precautions to be taken during testing. We believe it is important to determine how healing from an injury has progressed and how the client is affected by his or her condition. Gathering information regarding exercise programs, home and recreational activities, and level of functioning in activities of daily living is helpful as part of the FCE in order to establish a baseline on the client and reduce the risk of reinjury.[50]

Some FCEs incorporate psychological screening and pre-vocational testing into the initial data-gathering phase.[5] A vocational history is particularly helpful if job exploration is necessary.

Physical Examination

Some FCEs include a musculoskeletal examination, whereas other FCEs incorporate the musculoskeletal component into the evaluation.[24] A third approach is to perform a musculoskeletal screening only when "red flags" (eg, high blood pressure, elevated heart rate, recent surgery) are raised in the intake history or review of the medical record.[19] Regardless when it is performed, the musculoskeletal examination, along with the questionnaire and interview information, are taken into account for comparison during the performance of the physiological and functional assessment portions of the FCE. The primary reason for performing a physical assessment is to identify clinical signs related to conditions that are contraindications for testing or that should be monitored closely during testing.

Physiological Measurement

Measures include items such as muscular endurance and cardiovascular endurance. Muscular endurance is usu-

Table 4.
Common Components of Functional Capacity Evaluations

Record review
Self-administered questionnaire
Interview
Physical measures/musculoskeletal evaluation
Physiological measures
Functional measures
Comparison of testing with job requirements

ally measured by describing the demands of the activity performed and the duration of muscle performance prior to fatigue. Submaximal protocols, which have predetermined termination points based on a percentage of the client's estimated maximum heart rate, are a method of assessing cardiovascular endurance.[15]

A variety of approaches are used to predict the weights that can be lifted occasionally and frequently. The Isernhagen Work Systems FCE[22] notes changes in body mechanics, accessory muscle use, heart rate increase, and other physical signs to determine whether the effort expended to lift a load is light, moderate, or maximum. Blankenship[28] calculates the weight lifted frequently from the weight lifted occasionally, using 1 of 3 formulas. This system also provides for direct measurement of lifting performance by requiring the client to perform one lift every 5 seconds, with additional weight being added every 4 lifts. The Key Method[24] calculates the weight lifted frequently as 50% to 60% of the weight that can be lifted occasionally. None of these protocols have been published in peer-reviewed journals. Furthermore, a recent study attempted to predict weight lifted frequently by regression analysis and included factors such as subject's age, time since injury, and weight that could be lifted occasionally.[51] A large amount of error in the estimates led the researchers to conclude that the use of formulas in predicting the weight that can be lifted frequently is questionable.

Grading of exercise intensity and the accuracy achieved in predicting maximal oxygen uptake provided by treadmill and bicycle ergometer testing make these tests more popular for use in clinical and research environments.[52–55]

There is little reference in the literature to the inclusion of fitness evaluations in FCEs.[56] It appears that these assessments are usually conducted as "stand-alone" fitness evaluations rather than as part of an FCE.[57]

Functional Performance

Most comprehensive FCEs include the physical demands of work as specified by the US Department of Labor in the DOT.[39,40] The job factors listed in the DOT express both the physical requirements of jobs and the physical capacity a worker must have to meet those demands. Developing a method of measuring physical capacity in terms of job factors should allow a direct translation of which jobs a client can perform,[58] but only research can demonstrate whether this is true. These 20 physical demands are often used by rehabilitation specialists and vocational counselors to classify jobs. The DOT is a valuable resource in initially attempting to identify potential work opportunities. Just because a task is covered by a test, however, does not mean that it has been evaluated thoroughly or that the data produced are objective, specific, and quantified. An FCE may not have a standard protocol or an objective measurement tool.

Examiners need to determine not only whether a skill is evaluated in a particular FCE but also how it is evaluated. Testing should mimic the tasks and equipment as they occur in industry, and validity studies are the means of determining whether this has been done successfully. The FCEs that depend on very sophisticated technologies tend to cover fewer than the 20 physical demands of work. Whether the data generated by this equipment are more reliable or valid is yet to be determined.[1]

Critics of the DOT suggest it is not comprehensive enough in defining all types of jobs.[58,60] Despite these deficiencies, the DOT remains the most comprehensive set of occupational characteristics currently available.[58]

Physical capacity as it relates to lifting abilities is most often defined using the DOT classifications of physical work (Tab. 5).[39,40] In our opinion, broadly classifying the client as having sufficient strength to perform at the various levels of physical demands usually is not sufficient information for safe, productive job placement.[30] We believe, therefore, that it is important to test the client's lifting capacity under a variety of conditions and with the client in different postures.

Many lifting procedures related to functional testing have been identified.[61–63] Historically, 3 types of force assessments have been used: isometric, isokinetic, and isoinertial.[44,64,65] Isometric testing requires exertion of a force against an object that does not move. This type of testing is relatively simple, quick, inexpensive to administer, and easy to control.

With isokinetic testing, movement is restricted to a specific speed. This type of testing is usually limited to muscle testing rather than to assessment of task performances because tasks are almost never performed at constant speeds.

Isoinertial testing is described as lifting progressively heavier weights at a set frequency over a specific vertical range.[44,65] This type of testing has been adapted to test

adults with medical problems, and this version relies on established criteria such as cardiovascular response and the patient's requests for stopping the test. Isoinertial testing has been reported to be reliable for the amount of weight lifted.[13,46,50,66]

The client self-report method, sometimes called "psychophysical method" of testing, requires a subject to randomly adjust a load during a period of 20 to 25 minutes until an acceptable load is found that the subject estimates can be lifted repetitively during a regular 8-hour work shift.[67] This method relies on the client to determine capacity. Researchers have proven, however, that subjects tend to overestimate the maximum acceptable weight of lifting.[50] Critics of this method suggest that clients may have reasons to be fearful and, as a result, that measurements may not be reliable.[4]

Another approach to evaluating lifting capacity is to observe movement to determine maximal effort. This method of testing is described as lifting progressively heavier weights until the examiner makes a decision to stop the test.[4,68–70] The examiner relies on observational cues and cardiorespiratory signs to determine the degree of safety of the lifting tasks.

The results of isometric or isokinetic tests of individual muscle or whole-body torque correlate poorly with performance of functional activities,[44] and these tests do not appear to be valid for predicting function. These tests, therefore, may be more useful in measuring impairments. Direct tests of material-handling capabilities also have the practical advantages of simplicity, low cost, and portability. Although reliability has been demonstrated in lifting studies, these free-lifting protocols have been criticized for inadequate anatomic stabilization, subjectivity of the psychophysical end point, and lack of control for speed and acceleration variables.[13,62,67]

To ensure safe and consistent documentation of maximum functional levels, logical and ethical considerations should guide the development of lifting-capacity testing.[30] Most clinicians initially subject the client to low levels of weight at a given frequency and gradually increase the load while monitoring the client's response.[46,66] The clinician should encourage the client to lift slowly, smoothly, and continuously to avoid exacerbation of the injury.[71]

Report Writing

Several authors[1,3,50,72] have emphasized the importance of logical and clear reporting of observations, and that reports be in an easy-to-read format and free of jargon. The reader should get a clear picture of the client's physical capabilities and limitations as they relate to critical demands of the job. Optimally, we contend, this comparison should be made using the results of a job analysis, identifying both the essential physical demands and the necessary qualifications for the targeted job.[73] Obtaining a job analysis, however, is not yet a standard in many workplaces.

Table 5.
Dictionary of Occupational Titles[39,40] Physical Demands of Lifting and Carrying

	Maximum Lift Capacity (lb)	Maximum Carrying Capacity (lb)
Sedentary	10	5
Light	20	10
Medium	50	25
Heavy	100	50
Very Heavy	>100	>50

Sometimes, comparisons of the worker's abilities with the job demands need to be based on the clinician's estimates and on reports from the employer as well as from the employee. If the results of a job analysis are not available for comparison with the client's abilities and the employer will not cooperate with a job analysis, then job demands can be determined through client and employer self-report. The source of information on job demands should always be documented in the report. This documentation is particularly important if the client is involved in litigation. Client cooperation, consistency of the client's performance during the assessment, and safety of movements and body mechanics should be noted as well.[50]

As part of an FCE report, there should be clearly written descriptions of proposed job modifications and recommendations for any further modifications. Such descriptions are important because the results of FCEs are utilized by a wide range of professionals. A logical and clear reporting style is essential when assessing clients for Workers' Compensation purposes. All recommendations, therefore, should be substantiated by the data obtained during the FCE.[1,50] Where appropriate, the final report of an FCE should compare an injured worker's status with the requirements of the job to which he or she is returning.

Issues in FCE Development

Qualifications of the Evaluator

There is some debate surrounding the issue of who is best qualified to administer FCEs. Occupational therapists and physical therapists appear to have been the first health care professionals to perform FCEs.[50] These therapists have traditionally evaluated function, performed task analyses, and have an educational background that provides an understanding of pathology

and the musculoskeletal system, including muscle function and movement patterns.[3,7,50]

Other health care professionals currently involved in administering FCEs include athletic trainers, chiropractors, vocational evaluators, nurses, physical therapist assistants, certified occupational therapy assistants, exercise physiologists, psychologists, kinesiologists, and physicians. These individuals often claim to have acquired skills and knowledge in task and movement analysis from clinical practice and continuing education courses.

The issue of who is qualified to administer FCEs is being addressed by determining the competencies required to perform a safe, reliable, and valid FCE. Once determined, these competencies guide mandatory training sessions in which clinicians eventually demonstrate competence through written and practical skills as well as testing procedures.

Training

Given the difference in skill levels among the disciplines, evaluator training should, in our view, become an important means of ensuring quality and consistency in evaluation, scoring, and report writing related to FCEs. All FCEs on the market provide for training; however, there is variance in how this training is delivered. Some FCE designers use the "train-the-trainer" approach,[19,21,23,26–28] whereas other FCE designers insist on certifying each individual desiring to administer their FCE.[19,22,24] Regardless of the type of approach used, we believe that training is essential, but research has yet to provide data as to how much training is needed to obtain valid and reliable measurements. Training ensures that administrators of FCEs are knowledgeable and competent in carrying out the established procedures.

Length of Assessment

The length of time to administer an FCE varies among designers. Some FCEs require less than 2 hours to administer,[20,23,27] whereas other FCEs require 3 to 4 hours to administer.[19,25,27] One FCE is administered over a 2-day period.[22] Lechner et al[1] and Isernhagen[50] reported that the shorter FCEs (1–2 hours) seem to provide less reliable data and appear to be less comprehensive, but they did not provide data to support their contention. The shorter FCEs are less likely to include all 20 physical demand items identified by the DOT.

Miller[3] and Isernhagen[2,50] recommended that the standard FCE be 4 to 6 hours in length and assess general work demands such as lifting, carrying, reaching, sitting, standing, and walking, as well as hand strength and coordination. Tramposh[5] indicated that, to meet consumer demand, an FCE should be conducted in one session. Isernhagen,[2,50] however, stated that the most reliable format for conducting an FCE is over a 2-day period, with the most critical tasks being repeated on the second day. The 2-day format allows for retesting for accuracy and for evaluating the effect of the first day's assessment on the client, but Isernhagen did not provide data to support this assertion.

There are no data to support the selection of a specific length of time for FCE testing. Functional capacity evaluations that do not take long to administer may compromise content validity because addressing all of the physical demands of the job may not be possible. Likewise, validity may be affected in the performance of a 2-day evaluation if soreness is acquired from testing on the first day. Functional abilities may be decreased the second day, and the results of the FCE may underestimate the client's abilities, particularly once he or she has acclimated to the work environment.

There is growing pressure from a number of sources to reduce the amount of time spent on FCEs. One strategy is to not include the musculoskeletal evaluation in the FCE, but provide it as a separate service. Another strategy is to perform an abbreviated FCE, which is customized to address factors in a specific injury or job scenario. The module or job-specific FCE could be composed of a subset of tasks selected from a larger group of tasks that have been shown to yield useful measurements. A shorter FCE, however, may not determine endurance for a full work shift. In any circumstance, we argue that consumer demand should not be the sole criterion for decisions regarding the time taken for FCE testing. Research addressing reliability and validity, when complete, could provide data to guide evaluation decision making.

Projection to an 8-Hour Workday

Because work typically relates to a full day and week (approximately 8 hours a day, 5 days a week), an FCE needs to relate to these time periods. We believe it is important that endurance and tasks be evaluated so that tolerance for an 8-hour workday can be determined. Documentation of heart rate, endurance factors, change in body mechanics, and fatigue can assist with this projection.[4] Specific "formulas" for proposing 8-hour workday functions are not available. Each individual has a unique mix of physical capacities that cannot, at present, be generalized, which is why generic formulas are not accurate.[69] Abdel-Moty et al[71] reported that measurements of a client's abilities through an FCE do not necessarily predict the client's ability to perform over an 8-hour workday; that is, they lack validity. These authors believe that the attempt to extrapolate data from a 1- to 2-hour assessment to an 8-hour workday is a major conceptual error in the design of some FCEs. They contend that the determination of an individual's ability

to handle loads repetitively requires the use of evaluation methods that combine the measurement of weights lifted with assessment of heart rate and oxygen consumption.[75–77]

The measurement of endurance is helpful in correlating repetitious activity at work with the functional testing completed. With an FCE, heart rate, blood pressure, and respiratory rate can be measured at rest and during activity, and these measurements then can be analyzed to note the changes that occur with activity. In many cases, maximal permissible limits must be set for a person and then monitored to ensure the limits are not exceeded during an FCE. Matheson[78] suggested that clients should not be required to exceed a cardiovascular effort of 65% of their predicted maximum heart rate. According to the NIOSH[63] and Astrand and Rodahl,[79] continuous 8-hour expenditures should not exceed 33% of a worker's aerobic capacity.

Behavior Management and Assessment

Physicians are often asked to determine when reasonable progress has occurred or when a plateau has been reached and there is residual impairment. Input from other health care professionals is often sought when making these decisions. Once a worker's condition has plateaued, we contend that the client's physical functional abilities should be determined and compared with job demands. Confusion may arise when the client's self-report of disability does not match the signs (measurements) of physical impairment obtained by the examiner. Potential causes of magnified illness behavior include (1) unrecognized physical severity of the medical impairment, (2) psychological distress related to the duration, amount, and failure of treatment or dislike of the job or employer, and (3) voluntary exaggeration to influence legal proceedings.[30]

Most FCEs include a mechanism to address the issue of effort. In some FCEs, the client is asked to perform isometric strength tests (eg, grip strength), and the evaluator judges the consistency of the client's performance. There is no evidence, however, that an inconsistent response on isometric testing indicates submaximal performance on the remainder of the battery of tests.[1] (The reader is referred to the article by Lechner et al in this issue.)

When a Workers' Compensation claim is involved, full client participation in the testing may be questioned due to the financial gains involved. Reporting a lack of full participation to the employer or insurer may have repercussions for the client. Therefore, a reliable and valid method of determining subject participation is vital, but none have been supported by current research.[1,10]

Standards of Practice Policies

Functional capacity evaluations have been conducted for many years by physical therapists, occupational therapists, vocational evaluators, and psychologists. Only recently have their professions established practice standards. These standards are necessary, and we believe they should be developed on an interdisciplinary basis.[11] Because the individual therapist is vulnerable to the policy and practices of his or her employer, we believe it is essential that the various professional organizations negotiate standards of practice and that health care employers agree to abide by them.[29]

Therapists should be aware that they are legally responsible for the consequences of performing FCEs. For this reason, policies and procedures of testing must be documented. Client safety is often an issue, and the amount of documentation regarding client safety in test performance is an important part of any defense.[14,50]

Safety

The issue of safety in FCEs focuses on the prevention of further injury during the testing process. Injury may occur because the client is asked to put forth maximum voluntary effort. In our view, clinicians should intervene at the point where task performance begins to appear unsafe. This intervention is possible only with careful visual assessment and when criteria are available for determining maximal effort. These criteria may be based on consensus from many therapists regarding which visual signs are indicators of unsafe performance. An example of a visual observation of an unsafe act is loss of postural trunk alignment when lifting heavy weights.[7] Sound clinical decision making and professional judgment are extremely important.

Ogden-Niemeyer[80] reported 2 schools of thought regarding therapist intervention during testing to ensure client safety: (1) intervention decreases validity and reliability and (2) no intervention places the client at an unacceptable risk of injury. No data are available to support either contention. Isernhagen[2,4] stated that therapists should allow only safe, controlled body mechanics in testing situations. Unsafe procedures are potentially harmful to the worker, and it is clearly inappropriate if an evaluator notices unsafe procedures and allows the behavior to continue.

If unsafe movements are used by the client during a task, we believe that correction with instructions from the clinician should be attempted. If the individual is unable to continue to maintain safe, correct body mechanics, then the activity should stop and the reason for the stoppage should be documented.[68]

Currently, there is no infallible method for determining a safe stopping point of FCEs.[81] Some FCE designers contend that the therapist is in the best position to make this decision.[50] Other FCE designers argue that the client should determine the stopping point.[81]

Safety should, in our opinion, be one of the main concerns of a therapist when he or she starts to formulate recommendations for the client's return to work. Advice on modifying the environment or provision of specialized equipment is a fundamental component of the FCE.

Medicolegal Implications

Functional capacity evaluations are increasingly being used in the legal arena. This use of FCEs often leads to scrutiny of test standardization, and particularly to scrutiny of the qualifications of the evaluator.[50] Using his or her background in showing how the client was evaluated, the clinician can serve as an expert witness. Medicolegal credibility must be compatible with the development of more efficient, clearer, and safer FCEs.

Conclusion

An FCE is not a stand-alone evaluation. Rather, an FCE consists of an interview and client history, a physical examination, test components, and a comparison of a client's abilities with the demands of the job. The use of evaluation via observation of visual movement (biomechanics) as well as the client self-report and physiological measures (eg, heart rate, blood pressure) is necessary for a safe, objective, and valid report. Functional capacity evaluations of the future should be driven by science, medicolegal implications, and reimbursement issues.

Acknowledgments

We thank Commonwealth Rehabilitation Services of Western Australia and Curtin University of Technology for their support of this work.

References

1 Lechner DE, Roth D, Straaton K. Functional capacity evaluation in work disability. *Work.* 1991;1:37–47.

2 Isernhagen SJ. *Work Injury Management and Prevention.* Gaithersburg, Md: Aspen Publishers Inc; 1988.

3 Miller M. Functional assessments: a vital component of work injury management. *Work.* 1991;1:6–10.

4 Isernhagen SJ. Functional capacity evaluation: rationale, procedure, and utility of the kinesiophysical approach. *Journal of Occupational Rehabilitation.* 1992;2:157–164.

5 Tramposh AK. The functional capacity evaluation: measuring maximal work abilities. *Occup Med.* 1992;7:113–124.

6 Jette AM. Concepts of health and methodological issues in functional status assessment. In: Granger CV, Gresham GE, eds. *Functional Assessment in Rehabilitation.* Baltimore, Md: Williams & Wilkins; 1987:46–64.

7 Lechner DE. Functional capacity evaluation: an evidence-based approach. In: King PM, ed. *Sourcebook of Occupational Rehabilitation.* New York, NY: Plenum Press. In press.

8 Smith RO. Computer-assisted functional assessment and documentation. *Journal of Occupational Therapy.* 1993;47:988–992.

9 Abdel-Moty E, Khalil T, Sadek S, et al. Functional capacity assessment of chronic low back pain patients. In: Karwowski W, Yates JW, eds. *Advances in Industrial Ergonomics and Safety.* Vol III. Bristol, Pa: Taylor & Francis Publishers Inc; 1991:475–479.

10 Lechner DE, Jackson JR, Roth DL, Straaton KV. Reliability and validity of a newly developed test of physical work performance. *J Occup Med.* 1994;36:997–1004.

11 Hart DL, Isernhagen SJ, Matheson LN. Guidelines for functional capacity evaluation of people with medical conditions. *J Orthop Sports Phys Ther.* 1993;18:682–686.

12 Sen S, Fraser K, Evans OM, Stuckey R. A comparison of the physical demands of a specific job and those measured by standard functional capacity assessment tools. In: *Proceedings of the 27th Annual Conference of the Ergonomics Society of Australia Inc.* Canberra, New South Wales, Australia: Ergonomics Society of Australia Inc; 1991:263–268.

13 Matheson LN. Evaluation of lifting and lowering capacity. *Vocational Evaluation and Work Adjustment Bulletin.* Fall 1986:107–111.

14 Task Force on Standards for Measurement in Physical Therapy. Standards for tests and measurements in physical therapy practice. *Phys Ther.* 1991;71:589–622.

15 Owens LA, Buchholz RL. Functional capacity assessment, worker evaluation strategies, and the disability management process. In: Shrey DE, Lacerte M, eds. *Principles and Practices of Disability Management in Industry.* Winter Park, Fla: GR Press; 1995.

16 Anastasi A. *Psychological Testing.* New York, NY: Macmillan; 1988.

17 *The Minnesota Rate of Manipulation Test, Examiner's Manual.* Circle Pines, Minn: American Guidance Service; 1969.

18 *Americans With Disabilities Act: Technical Assistance Manual.* Washington, DC: Equal Employment Opportunity Commission; 1992.

19 *Physical Work Performance Evaluation Users Guide.* Birmingham, Ala: ErgoScience.

20 *WEST-EPIC Users Guide.* Ballwin, Mo: Employment Potential Improvement Corporation.

21 *WorkAbility Mark III Users Guide.* Ryde, New South Wales, Australia.

22 *Isernhagen Work System Users Guide.* Duluth, Minn: Isernhagen Work Systems.

23 *ARCON Users Guide.* Williamsburg, Va: Applied Rehabilitation Concepts.

24 *Key Method Users Guide.* Minneapolis, Minn: Key Method.

25 *WorkHab Users Guide.* Bundaberg, Queensland, Australia: WorkHab Australia.

26 *ERGOS Users Guide.* Tucson, Ariz: Work Reovery Systems Inc.

27 *AssessAbility Users Guide.* Austin, Tex: MediSys Rehabilitation Inc.

28 *The Blankenship System Users Guide.* Macon, Ga: The Blankenship System.

29 Menard MR, Hoens AM. Objective evaluation of functional capacity: medical, occupational, and legal settings. *J Orthop Sports Phys Ther.* 1994;19:249–260.

30 Wickstrom R. Functional capacity testing. In: Scheer SJ, ed. *Multidisciplinary Perspectives in Vocational Assessment of Impaired Workers.* Gaithersburg, Md: Aspen Publishers Inc; 1990:73–88.

31 Mathiowetz V. Role of physical performance component evaluations in occupational therapy functional assessment. *Am J Occup Ther.* 1993;47:225–230.

32 Matheson LN, Mooney V, Grant JE, et al. Standardized evaluation of work capacity. *Journal of Back & Muscle Rehabilitation.* 1996;6: 249–264.

33 Aghazadeh F, Ayoub MM. A comparison of dynamic- and static-strength models for prediction of lifting capacity. *Ergonomics.* 1985;28: 1409–1417.

34 *Guide to the Evaluation of Permanent Impairment.* Chicago, Ill: American Medical Association; 1993.

35 *Testing and Training of Apparently Healthy Individuals: A Handbook for Physicians.* Dallas, Tex: American Heart Association; 1972.

36 Bohannon RW. Test-retest reliability of hand-held dynamometry during a single session of strength assessment. *Phys Ther.* 1986;66: 206–209.

37 Bohannon RW, Andrews AW. Interrater reliability of hand-held dynamometry. *Phys Ther.* 1987;67:931–933.

38 Borg G. Psychophysical basis of perceived exertion. *Med Sci Sports Exerc.* 1982;14:377–381.

39 US Department of Labor, Employment and Training Administration. *Revised Dictionary of Occupational Titles.* Vol I. 4th ed. Washington, DC: US Government Printing Office; 1991.

40 US Department of Labor, Employment and Training Administration. *Revised Dictionary of Occupational Titles.* Vol II. 4th ed. Washington, DC: US Government Printing Office; 1991.

41 Fairbank JC, Davies JB, Couper J, O'Brien JP. The Oswestry Low Back Pain Disability Questionnaire. *Physiotherapy.* 1980;66:271–273.

42 Hildreth DH, Breidenbach WC, Lister GD, Hodges AD. Detection of submaximal effort by use of the rapid exchange grip. *J Hand Surg Am.* 1989;14:742–745.

43 Keyserling WM, Herrin GD, Chaffin DB. An analysis of selected work muscle strength. In: *Proceedings of the 22nd Annual Meeting of the Human Factors Society.* Santa Monica, Calif: Human Factors Society; 1978.

44 Kroemer K. An isoinertial technique to assess individual lifting capacity. *Hum Factors.* 1983;25:493–506.

45 Mathiowetz V, Kashman N, Volland G, et al. Grip and pinch strength: normative data for adults. *Arch Phys Med Rehabil.* 1985;66: 69–74.

46 Mayer TG, Barnes D, Kishino ND, et al. Progressive isoinertial lifting evaluation, I: a standardized protocol and normative database. *Spine.* 1988;13:993–997.

47 Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain.* 1975;1:277–299.

48 MacDonald WA. Performance capacity evaluation: some comments on its validity. In: *Proceedings of the 28th Annual Conference of the Ergonomic Society of Australia Inc.* Canberra, New South Wales, Australia: Ergonomics Society of Australia Inc; 1992:189–196.

49 Beurskens AJ, de Vet RC, Koke AJ, et al. Measuring the functional status of patients with low back pain: assessment of the quality of four disease-specific questionnaires. *Spine.* 1995;20:1017–1028.

50 Isernhagen SJ. Advancements in functional capacity evaluation. In: D'Orazio BP, ed. *Back Pain Rehabilitation.* Boston, Mass: Butterworth; 1993:180–204.

51 Saunders RL, Beissner KL, McManis BG. Estimates of weight that subjects can lift frequently in functional capacity evaluations. *Phys Ther.* 1997;77:1717–1728.

52 Arstila M. Pulse-conducted triangular exercise-ECG test: a feedback system regulating work during exercise. *Acta Med Scand Suppl.* 1972; 529:3–109.

53 Jones NL, Campbell EJ, McHardy GJ, et al. The estimation of carbon dioxide pressure of mixed venous blood during exercise. *Clin Sci.* 1967;32:311–327.

54 Nagle FJ, Bauke B, Naughton JP. Gradational step tests for assessing work capacity. *J Appl Physiol.* 1965;20:745–748.

55 Spiro SG, Juniper E, Bowman P. An increasing work rate for assessing the physiological strain of submaximal exercise. *Ergonomics.* 1974;46:191–206.

56 Gibson L. *Functional Capacity Evaluation: Recommendations for Use and Provision in CRS QLD.* Brisbane, Queensland, Australia: Department of Human Services and Health. Unpublished document.

57 Fraser M. *Fitness for Work.* Bristol, Pa: Taylor & Francis Publishers Inc; 1992.

58 Fishbain DA, Abdel-Moty E, Cutler RK, et al. Measuring residual functional capacity in chronic low back pain patients based on the Dictionary of Occupational Titles. *Spine.* 1994;19:872–880.

59 Maze M. Using the DOT. *National Association of Rehabilitation Professionals Society J News.* 1989;5:141–145.

60 Miller AR, Treiman DJ, Cain PS, Roos PA. *Work, Jobs, and Occupations: A Critical Review of the Dictionary of Occupational Titles.* Washington, DC: National Academy Press; 1980.

61 Matheson LN, Mooney V, Grant JE, et al. A test to measure lift capacity of physically impaired adults, part 1: development and reliability testing. *Spine.* 1995;20:2119–2129.

62 Alpert J, Matheson L, Beam W, Mooney V. The reliability and validity of two new tests of maximum lifting capacity. *Journal of Occupational Rehabilitation.* 1991;1:13–29.

63 *Work Practices Guide for Manual Lifting.* Washington, DC: National Institute of Occupational Safety and Health, US Public Health Service; 1981.

64 Keyserling WM, Herrin GD, Chaffin DB. Isometric strength testing as a means of controlling medical incidents on strenuous jobs. *J Occup Med.* 1980;22:332–336.

65 Jacobs I, Bell DG, Pope J. Comparison of isokinetic and isoinertial lifting tests as predictors of maximal lifting capacity. *Eur J Appl Physiol.* 1988;57:146–153.

66 Mayer TG, Barnes D, Nichols G, et al. Progressive isoinertial lifting evaluation, II: a comparison with isokinetic lifting in a disabled chronic low-back pain industrial population. *Spine.* 1988;13:998–1002.

67 Mital A. A psychophysical approach in manual lifting: a verification study. *Hum Factors.* 1983;25:485–491.

68 Isernhagen SJ. Functional capacity evaluation and work hardening perspectives. In: Mayer T, Mooney V, Gatchel R, eds. *Contemporary Care for Painful Spinal Disorders.* Malvern, Pa: Lea & Febiger; 1991:328–345.

69 Isernhagen SJ. Return to work testing: functional capacity and work capacity evaluation. *Orthopaedic Physical Therapy Clinics.* 1992;1:83–98.

70 Johnson LJ. The kinesiophysical approach matches worker and employer needs. In: Isernhagen SJ, ed. *The Comprehensive Guide to Work Injury Management.* Gaithersburg, Md: Aspen Publishers Inc; 1995: 399–409.

71 Hart DL, Stobbe TJ, Jaraiedi M. Effect of lumbar posture on lifting. *Spine.* 1987;12:138–145.

72 Williams K. Functional capacity evaluation of the upper extremity. *Work.* 1991;1:48–64.

73 US Department of Labor, Employment and Training Administration. *The Revised Handbook for Analyzing Jobs.* Indianapolis, Ind: JIST Works; 1991.

74 Abdel-Moty E, Fishbain D, Khalil T, et al. Functional capacity and residual functional capacity and their utility in measuring work capacity. *Clin J Pain.* 1993;9:168–173.

75 Legg J, Pateman CM. A physiological study of the repetitive lifting capability in healthy young males. *Ergonomics.* 1984;27:259–272.

76 Legg J, Myles WS. Maximum acceptable repetitive lifting workloads using psychophysical and subjective rating methods. *Ergonomics.* 1981; 24:907–916.

77 Mital A. Comprehensive maximum acceptable weight of lift database for regular 8-hour work shift. *Ergonomics.* 1984;27:1127–1138.

78 Matheson LN. "How do you know that he tried his best?" The reliability crisis in industrial rehabilitation. *Industrial Rehabilitation Quarterly.* 1988;1:10–12.

79 Astrand I, Rodahl K. *Textbook of Work Physiology.* New York, NY: McGraw-Hill; 1970.

80 Ogden-Niemeyer L. *Procedure Guidelines for the WEST Standard Evaluation: Assessment of Range of Motion Under Load.* Long Beach, Calif: Work Evaluation Systems Technology; 1989.

81 Key GL. *Industrial Therapy.* St Louis, Mo: CV Mosby Co; 1997.