Functional Assessment for Adults with Disabilities

Committee on Functional Assessment for Adults with Disabilities

Paul A. Volberding, Carol Mason Spicer, and Jennifer Lalitha Flaubert,
Editors

Board on Health Care Services

Health and Medicine Division

A Consensus Study Report of
The National Academies of
SCIENCES · ENGINEERING · MEDICINE

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

This activity was supported by Contract/Task Order No. SS00-13-60048/0009 with the U.S. Social Security Administration. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

International Standard Book Number-13: 978-0-309-48938-6
International Standard Book Number-10: 0-309-48938-5
Digital Object Identifier: https://doi.org/10.17226/25376
Library of Congress Control Number: 2019945053

Additional copies of this publication are available from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Printed in the United States of America

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2019. *Functional assessment for adults with disabilities*. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/25376.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.nationalacademies.org.**

The National Academies of

SCIENCES • ENGINEERING • MEDICINE

Consensus Study Reports published by the National Academies of Sciences, Engineering, and Medicine document the evidence-based consensus on the study's statement of task by an authoring committee of experts. Reports typically include findings, conclusions, and recommendations based on information gathered by the committee and the committee's deliberations. Each report has been subjected to a rigorous and independent peer-review process and it represents the position of the National Academies on the statement of task.

Proceedings published by the National Academies of Sciences, Engineering, and Medicine chronicle the presentations and discussions at a workshop, symposium, or other event convened by the National Academies. The statements and opinions contained in proceedings are those of the participants and are not endorsed by other participants, the planning committee, or the National Academies.

For information about other products and activities of the National Academies, please visit www.nationalacademies.org/about/whatwedo.

COMMITTEE ON FUNCTIONAL ASSESSMENT FOR ADULTS WITH DISABILITIES

PAUL A. VOLBERDING (*Chair*), Professor, Department of Medicine; Director, AIDS Research Institute, Co-Director, University of California, San Francisco-Gladstone Institute of Virology and Immunology Center for AIDS Research, University of California, San Francisco

MARÍA P. ARANDA, Associate Professor, School of Social Work, University of Southern California

JACK T. DENNERLEIN, Professor, Department of Physical Therapy, Movement and Rehabilitation Sciences, Northeastern University

LISA DIXON, Professor, Department of Psychiatry, Columbia University; Director, Center for Practice Innovations, Division of Behavioral Health Services and Policy Research, New York State Psychiatric Institute

JUDITH GREEN-McKENZIE, Director of Clinical Practice, Associate Residency Director, Associate Professor, Division of Occupational Medicine, Department of Emergency Medicine, University of Pennsylvania Perelman School of Medicine

ALLEN W. HEINEMANN, Director, Center for Rehabilitation Outcomes Research; Professor, Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University

ANDREW J. HOUTENVILLE, Associate Professor, Paul College of Business and Economics; Research Director, Institute on Disability, University of New Hampshire

KURT L. JOHNSON, Director, Center for Technology and Disability Studies; Professor, Department of Rehabilitation Medicine; Head, Division of Rehabilitation Counseling, University of Washington

BARBARA L. KORNBLAU, Adjunct Professor of Occupational Therapy, School of Allied Health Science, Florida A&M University, Coalition for Disability Health Equity

PHILIP JORDAN MARION, Medical Director, Polytrauma Amputation Network Site, Washington, DC, Veterans Affairs Medical Center, Medical Director, Capital Medical Services

SUSAN McGURK, Professor, Department of Occupational Therapy; Senior Researcher, Center for Psychiatric Rehabilitation, Boston University

JUAN I. SANCHEZ, Professor, Knight-Ridder Byron Harless Eminent Chair in Management, College of Business, Florida International University

PAUL SHATTUCK, Associate Professor, A.J. Drexel Autism Institute, Drexel University
LYNNE WARNER STEVENSON, Professor of Medicine and Lisa M. Jacobson Chair in Cardiovascular Medicine, Vanderbilt University Medical School; Director of Cardiomyopathy Program, Vanderbilt Heart and Vascular Institute, Vanderbilt University Medical Center
ROBERT B. WALLACE, Irene Ensminger Stecher Professor of Epidemiology and Internal Medicine, Department of Epidemiology, College of Public Health, The University of Iowa

Study Staff

CAROL MASON SPICER, Study Director
JENNIFER LALITHA FLAUBERT, Associate Program Officer
TOM CARTAXO, Research Assistant
SHARYL NASS, Director, Board on Health Care Services

Reviewers

This Consensus Study Report was reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise. The purpose of this independent review is to provide candid and critical comments that will assist the National Academies of Sciences, Engineering, and Medicine in making each published report as sound as possible and to ensure that it meets the institutional standards for quality, objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We thank the following individuals for their review of this report:

David Cella, Northwestern University Feinberg School of Medicine
Naomi L. Gerber, George Mason University
Howard H. Goldman, University of Maryland at Baltimore School of Medicine
Clarion E. Johnson, ExxonMobil (Retired)
Kevin Liebkemann, Legal Services of New Jersey
David B. McNaughton, The Pennsylvania State University
Mark R. Munetz, Northeast Ohio Medical University
Marc A. Norman, University of California, San Diego
Gbenga Ogedegbe, New York University
Jackson Roush, University of North Carolina School of Medicine
Gwendolyn Sowa, University of Pittsburgh School of Medicine
Craig A. Velozo, Medical University of South Carolina
Sheila K. West, Johns Hopkins Hospital

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations of this report nor did they see the final draft before its release. The review of this report was overseen by **Sara Rosenbaum,** The George Washington University, and **Elaine Larson,** Columbia University School of Nursing. They were responsible for making certain that an independent examination of this report was carried out in accordance with the standards of the National Academies and that all review comments were carefully considered. Responsibility for the final content rests entirely with the authoring committee and the National Academies.

Preface

The U.S. Social Security Administration (SSA) annually reviews more than 3 million claims from adults requesting a determination of disability for physical or mental health conditions that prevent them from engaging in substantial gainful employment. Assessment of individuals' functional abilities relevant to work requirements should be an important part of determining whether they are able to meet workplace demands and sustain work performance on a regular and continuing basis. Over the years, many tests, surveys, and instruments have been developed to collect information about individuals' physical and mental functional abilities. With support from SSA, the Health and Medicine Division (HMD) of the National Academies of Sciences, Engineering, and Medicine convened a committee to identify and describe ways to collect information about individuals' physical and mental abilities relevant to function in the workplace, to discuss the types of information that support findings of limitations in work-related function, and to provide findings and conclusions about the collection of information and assessment of functional abilities relevant to work requirements.

On behalf of the committee, I sincerely thank the SSA staff who worked closely with the committee to define details of the committee's task, including Megan Butson, Gina Clemons, Joanna Firmin, Deborah Harkin, Mary Beth Rochowiak, and Melissa Spencer. We also thank the expert and dedicated HMD staff who worked tirelessly to bring together information used in our deliberations. Led by study director Carol Mason Spicer, assisted by Jennifer Lalitha Flaubert (associate program officer) and Tom Cartaxo (research assistant), the committee accomplished a great deal, none of which would have been possible without their hard work, focus, and expertise.

We extend thanks as well to other HMD staff members who played a key role in the production of this report, including Sharyl Nass (senior board director, Board on Health Care Services); Karen Helsing (senior program officer); Patrick Burke, Julie Wiltshire, and Micah Winograd (financial associates). Research assistance was provided by Rebecca Morgan (senior research librarian, the National Academies). We are indebted as well to Rona Brière and Alisa Decatur, who provided exceptional editorial assistance in preparing the final report.

Our thanks go to all of the presenters who shared their time and expertise during the committee's information-gathering sessions (see Appendix A); to the authors of three papers commissioned by the committee, Laura Ball, Ryan McCreery, Mark Wilkinson, Gislin Dagnelie, and Christopher Johnson; and to the reviewers for their invaluable feedback on an earlier draft of the report.

Finally, I thank the committee itself. As is so often the case, the National Academies are able to bring together experts across a wide range of expertise, each able to contribute to specific aspects of the task we were assigned. But the ability to bring that individual expertise together to a coherent whole demands an openness to learning from each other and a willingness to work as a team. Our committee did that and more. We each learned a great deal and trust that our report will be useful to SSA as it continues to improve in its essential mission.

Paul A. Volberding, *Chair*
Committee on Functional Assessment for Adults with Disabilities

Contents

Boxes, Figures, and Tables

BOXES

FIGURES

TABLES

Acronyms and Abbreviations

ADL	activity of daily living
ALS	amyotrophic lateral sclerosis
APS	Attending Physician Statement
ASD	autism spectrum disorder
ASHA	American Speech-Language-Hearing Association
BAC	Brief Assessment of Cognition
BLS	U.S. Bureau of Labor Statistics
CAPS	Clinician-Administered PTSD Scale
CAT	computer adaptive testing
CCSE	Cognitive Capacity Screening Examination
CI	confidence interval
CIDI	Composite International Diagnostic Interview
CNSLBP	chronic nonspecific low back pain
CPP	Canada Pension Program
CPTH	chronic posttraumatic headache
DASH	Disabilities of the Arm, Shoulder and Hand Questionnaire
DBQ	Disability Benefits Questionnaire
DPI	Depression Prognosis Index
DSM	*Diagnostic and Statistical Manual of Mental Disorders*
DSP	Disability Support Pension

EHR	electronic health record
EKG	electrocardiogram
ESII	extended speech intelligibility index
EWPS	Endicott Work Productivity Scale
FAAM	Foot and Ankle Ability Measure
FCE	functional capacity evaluation
FGA	Functional Gait Assessment
GAD	generalized anxiety disorder
GAF	Global Assessment of Functioning
GCS	Glasgow Coma Score
HAM-D	Hamilton-Depression Rating Scale
HAP	Human Activity Profile
HHIA	Hearing Handicap Inventory for Adults
HPQ	World Health Organization Heath and Work Performance Questionnaire
HRQOL	health-related quality of life
IADL	instrumental activity of daily living
ICC	intraclass correlation coefficient
ICF	*International Classification of Functioning, Disability and Health*
IME	independent medical examination
IRT	item response theory
K10	Kessler Psychological Distress Scale
KCCQ	Kansas City Cardiomyopathy Questionnaire
L&I	Washington State Department of Labor and Industries
LEFS	Lower-Extremity Functional Scale
LOC	loss of consciousness
MC	medical consultant
MDD	major depressive disorder
METS	metabolic equivalents
MHQ	Michigan Hand Outcomes Questionnaire
MIRECC	Mental Illness Research, Education and Clinical Center
MLHFQ	Minnesota Living with Heart Failure Questionnaire
MMI	maximal medical improvement
MMSE	Mini-Mental State Exam

MoCA	Montreal Cognitive Assessment
MRFC	Mental Residual Functional Capacity Assessment
NDI	Neck Disability Index
Neuro-QoL	Quality of Life in Neurological Disorders
NIH	National Institutes of Health
NRS	numeric rating scale
NYHA	New York Heart Association
OAE	otoacoustic emission
OCD	obsessive-compulsive disorder
ODI	Oswestry Disability Index
OFS	Occupational Functioning Scale
OIS	Occupational Information System
OMCT	Short Orientation-Memory-Concentration Test of Cognitive Impairment
OMPQ	Orebro Musculoskeletal Pain Questionnaire
O*NET	Occupational Information Network
ORS	Occupational Requirements Survey
PC	psychological consultant
PCL	PTSD Checklist
PHQ-9	Patient Health Questionnaire-9
PREE	Patient-Rated Elbow Evaluation
PROMIS	Patient-Reported Outcomes Measurement Information System
PRTF	Psychiatric Review Technique Form
PRWE	Patient-Rated Wrist Evaluation
PSP	Personal and Social Performance Scale
PTA	posttraumatic amnesia
PTSD	posttraumatic stress disorder
QCL	Quality of Communication Life Scale
RFC	residual functional capacity
RMDQ	Roland-Morris Disability Questionnaire
ROM	range of motion
SCL	Symptoms Checklist Depression Scale
SD	standard deviation
SF-36	36-Item Short Form Health Survey
SGA	substantial gainful activity

SII	speech intelligibility index
SLOF	Specific Level of Functioning Scale
SLP	speech-language pathologist
SOFAS	Social and Occupational Functioning Assessment Scale
SSA	U.S. Social Security Administration
SSDI	Social Security Disability Insurance
SSI	Supplemental Security Income
SSQ	Speech, Spatial, and Qualities of Hearing Questionnaire
TBI	traumatic brain injury
UPSA	University of California, San Diego, Performance-Based Skills Assessment
VA	U.S. Department of Veterans Affairs
VAS	visual analog scale
VBA	Veterans Benefits Administration
VDS	Vantaa Disability Study
VHA	U.S. Veterans Health Administration
WAI	Work Ability Index
WAIS-IV	Wechsler Adult Intelligence Scale, fourth edition
WD-FAB	Work Disability Functional Assessment Battery
WHO	World Health Organization
WHODAS 2.0	World Health Organization Disability Assessment Schedule 2.0
WLQ	Work Limitations Questionnaire
WPAI	Work Productivity and Activity Impairment Questionnaire
Y-BOCS	Yale-Brown Obsessive Compulsive Scale

Summary[1]

The U.S. Social Security Administration (SSA) provides disability benefits through the Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI) programs. The SSDI program, established in 1956, provides benefits to eligible adults with disabilities who have paid into the Disability Insurance Trust Fund, as well as to their spouses and adult children who are unable to work because of severe long-term disabilities. Enacted in 1972, SSI is a means-tested program based on income and financial assets that provides income assistance from U.S. Treasury general funds to adults aged 65 and older, individuals who are blind, and adults and children with disabilities. As of December 2017, SSDI had approximately 10.4 million beneficiaries and SSI about 7.1 million recipients who were classified as blind or disabled.

To receive SSDI or SSI disability benefits, an individual must meet the statutory definition of disability, which is "the inability to engage in any substantial gainful activity [SGA] by reason of any medically determinable physical or mental impairment which can be expected to result in death or which has lasted or can be expected to last for a continuous period of not less than 12 months." SSA uses a uses a five-step sequential process to determine whether an adult applicant meets this definition. The agency gathers information, including functional information, from the applicant, relevant health care providers, and third parties about the applicant's

[1] With the exception of Box S-2, this summary does not include references. Citations to support the text and conclusions herein are provided in the body of the report.

impairment-related limitations that may affect what he or she can do in a work setting. At step 1, SSA considers applicants' work activity in the past year. If an applicant is engaging in SGA (determined by earnings), SSA will not proceed with a disability determination. At step 2, SSA determines whether the applicant has a medically determinable physical or mental impairment or a combination of impairments that meets the severity and duration requirements for disability. If so, at step 3 the agency determines whether the applicant's impairment(s) meets or medically equals one of the listings in the *Listing of Impairments*, a number of which include functional criteria. For applicants whose impairment(s) do not meet or equal one of the listings, SSA then determines and considers their *residual functional capacity* (RFC) at steps 4 and 5. RFC is defined as "the most [an applicant] can still do despite [his or her impairment-related] limitations" or restrictions on "a regular and continuing basis," currently defined as 5 days per week, 8 hours per day, or an equivalent work schedule.

In 2017, SSA asked the Health and Medicine Division of the National Academies of Sciences, Engineering, and Medicine to convene a committee of relevant experts to provide findings and conclusions regarding the collection of information and assessment of functional abilities relevant to work requirements. The committee's Statement of Task is presented in Box S-1.

STUDY APPROACH AND SCOPE

The committee conducted an extensive review of the literature pertaining to functional assessment of physical and mental abilities relevant to work requirements, as well as the literature specific to assessment of function and impairment trajectories in individuals with back disorders, cardiac impairments, depression, and traumatic brain injury (TBI). The committee chose to address TBI in addition to the three conditions listed in the Statement of Task (see Box S-1) because of its prevalence and the associated high rates of cognitive impairment and work disability. In addition, the committee held three public meetings and one public teleconference to hear from invited experts in areas pertinent to its charge. The committee also commissioned three papers: on assessment of (1) hearing, (2) speech and language, and (3) vision in the context of work requirements. Collectively, these sources informed the committee's findings and conclusions.

The information gathered by the committee falls into four overlapping areas. The first is background information on the concepts of disability, function, and functional assessment, along with the types, sources, and quality of functional information; properties of assessment measures; and potential threats to validity in assessments—information that provided context for the committee's task. Having gathered this information, the committee developed a framework based on the *International Classification of*

BOX S-1
Statement of Task

An ad hoc committee will

1. Identify and describe ways to collect information about an individual's physical and mental (cognitive and noncognitive) functional abilities relevant to work requirements as defined by the U.S. Bureau of Labor Statistics' Occupational Requirements Survey (ORS), such as sitting/standing/walking, lifting/carrying, vision, communication, decision making, and adaptability;
2. Discuss the types of information that support findings of limitations in functional abilities relevant to work requirements as defined by the U.S. Department of Labor for the Occupational Information System (OIS); and
3. Provide findings and conclusions regarding the collection of information and assessment of functional abilities relevant to work requirements.

As a guide for literature review, information and data gathering, public sessions, discussions, deliberations, and report development, including findings and conclusions, the committee shall consider the following specific topics:

1. **Identify and describe ways to collect information about an individual's physical and mental (cognitive and noncognitive) functional abilities relevant to work requirements as defined by the U.S. Bureau of Labor Statistics' Occupational Requirements Survey (ORS), such as sitting/standing/walking, lifting/carrying, vision, communication, decision making, and adaptability.**
   - Provide an overview of the functional assessment processes in at least three similar benefit programs that assess disability or vocational capabilities (national and state government programs, private-sector programs, and foreign programs as applicable);
   - Provide examples of forms, tools, guides, examinations, and other resources used by benefit programs that assess functional aspects of disability and vocational capabilities;
   - Identify activities of daily living that correlate with the physical and mental (cognitive and noncognitive) demands of work; and
   - Provide examples of how to collect information on activities of daily living that correlate with the physical and mental (cognitive and noncognitive) demands of work.
2. **Discuss the types of information that support findings of limitations in functional abilities relevant to work requirements as defined by the U.S. Department of Labor for the Occupational Information System (OIS).**
   - Describe the laboratory findings, signs, or symptoms of impairments that support findings of limitations in functional abilities relevant to work requirements;
   - Explain what information, including that which pertains to level of severity and duration, may be found in medical or other evidence to support a finding that a person is unable to sustain physical and mental work

continued

BOX S-1 Continued

activities on an ongoing and independent basis in the context of functional limitations; and

- Identify any quantifiable limitations that may preclude certain levels of work (including sedentary) and give examples of the evidence to demonstrate such limitations.

3. **Provide findings and conclusions regarding the collection of information and assessment of functional abilities relevant to work requirements.**
    - Explain how limitations in functional abilities relevant to work requirements are more or less associated with particular mental or physical impairments;
    - Identify particular medical specialties and allied health fields that are likely to have the training and expertise to perform functional assessments related to work requirements;
    - Identify tools that signify a functional assessment was performed, and how likely those reports are to be valid representations of a claimant's functional limitations;
    - In the context of disability assessment, describe the spectrum of changes to functional abilities relevant to work requirements related to the progression of common disease processes in example impairments. These could include, but are not limited to, back disorders, cardiac impairments, or depression.
        - Identify where along the spectrum an individual's ability to perform functions relevant to work requirements is affected;
        - Describe whether the Social Security Administration (SSA) could expect improvement, no improvement, or progressive worsening in the example impairments;
        - Describe when significant changes in functional abilities relevant to work requirements may occur through the aging process for these examples, such as for adults with common age-related physical and mental impairments;
        - Describe the efficacy of medications and other treatments on an individual's ability to perform functional abilities relevant to work requirements for these examples, and whether that treatment causes its own subset of medical and/or psychological problems that negatively affect an individual's functioning and how SSA could request an appropriate assessment of functional changes; and
        - Describe how the examples are similar to or different from other impairments;
    - Discuss the advantages and disadvantages of generic functional assessment questionnaires that address a broad range of impairments and functional abilities relevant to work requirements, and targeted impairment-specific questionnaires, along with considerations in their use; and
    - Describe the best ways to determine the accuracy and validity of self-reported functional abilities, for example asking for input over the course of the claimant's interactions with SSA and comparing for consistency.

Functioning, Disability and Health (ICF) (see Figure S-1). The framework illustrates a structure and a hierarchy for moving from functional assessment of an individual to his or her capacity to perform work. Some components are influenced more by the person (left of the dotted line), while others are influenced more by work factors (right of the dotted line). These factors can overlap and interact. "Interrupters" are factors associated with the individual's health condition and its treatment (e.g., medication effects, fluctuations in symptoms) that may limit his or her ability to perform work activities on a sustained basis. In addition, environmental and organizational contextual factors may act as barriers to or facilitators of work performance. In terms of facilitators, modifications to community and work environments permit work participation by many individuals who otherwise would be unable to do so. For example, public transportation, including wheelchair-accessible buses, permits travel to work for many individuals with mobility impairments. Likewise, workplace accommodations and worker adaptation programs offered by employers and employee assistance programs can help workers navigate challenges posed by their conditions and retain employment.

The committee's framework provides a way of conceptually organizing various sources of information and specific types of tools for assessing function. For example, medical records may provide information about an individual's particular health condition and its manifestation in body function and structure, while computer adaptive testing may yield information

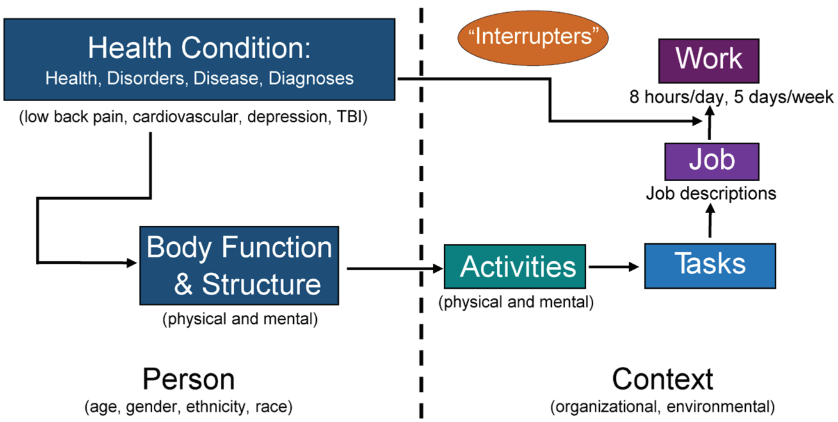

FIGURE S-1 A conceptual framework for functional assessment of an individual's capacity for work.
NOTE: TBI = traumatic brain injury.

on the person's ability to complete goal-based tasks, such as those required for work, that incorporate both mental and physical activities. Collectively, this material demonstrates the theoretical complexity of assessing an individual's functioning with respect to work. Box S-2 contains definitions for many of the terms used in this discussion.

Second, the committee gathered information on instruments used to assess the integrated effect of individuals' impairments on general daily life and participation and/or on work-related function, the relationship between instruments used to assess activities of daily living and the physical and mental demands of work, and instruments used to assess limitations in work activity due to health conditions. Also reviewed were a variety of instruments for assessing specific physical and mental functional abilities relevant to work requirements.

Third, the committee explored the spectrum of changes in work-related functional abilities that may occur during the progression of the four selected impairments (back disorders, cardiac impairments, depression, and TBI), as well as the effects of treatment on a person's ability to perform

BOX S-2
Definitions

Abilities: "Enduring attributes of the individual that influence performance" (O*NET Resource Center, 2019, p. 1).

Activity: "The execution of a task or action by an individual. It represents the individual perspective of functioning" (WHO, 2001, p. 213).

Capacity: "An individual's ability to execute a task or an action" (WHO, 2001, p. 123).

Functional abilities relevant to work: Refers to physical and mental activities such as sitting/standing/walking, lifting/carrying, vision, communication, decision making, and adaptability (see the committee's Statement of Task in Box S-1).

Functioning: "An umbrella term encompassing all body functions, activities, and participation" (WHO, 2001, p. 3).

Participation: "A person's involvement in a life situation. It represents the societal perspective of functioning" (WHO, 2001, p. 213).

Performance: "The execution of an action" (*Merriam-Webster*, 2019).

Task: A set of mental and physical activities in which an individual engages to accomplish a specific goal at or by a specific time.

work-related functions. The focus on specific impairments illustrated the complexity of assessing function with respect to work.

Fourth, in accordance with its Statement of Task (see Box S-1), the committee reviewed functional assessment processes in selected public and private disability programs that provide monetary benefits.

OVERALL CONCLUSIONS

In addition to findings and conclusions specific to the topics outlined above, the committee formulated five overall conclusions.

Relationship of Functional Abilities to Work Participation

Current models of disability, such as the ICF model, consider disability to involve the effects (limitations) an individual's health condition places on his or her ability to function and participate fully in society. In keeping with these models, assessment of individuals' functional abilities relevant to work requirements is an important part of determining whether they are able to meet workplace demands and sustain work performance on a regular and continuing basis.

Numerous validated performance-based and self-report instruments are available to assess physical and mental functions and can be used to inform disability determination. However, assessment of functional abilities does not necessarily address an individual's capacity to perform tasks required for work participation. Although an individual may be capable of performing each activity separately, he or she may not be able to coordinate and sequence them effectively. In addition, while an individual may be able to perform work tasks successfully during a single assessment, he or she may be unable to perform required work tasks on a sustained or consistent (day-to-day) basis because of one or more underlying physical and/or mental health conditions. For example, an individual may be able to lift a 50-pound box several times during an assessment, but he or she may not be able to do so repeatedly throughout a workday. It also is important to consider that testing is typically administered in a controlled, quiet environment without extraneous noise, social demands, and other factors that typically occur on a job, which, depending on the individual, can adversely affect the ability to perform work tasks.

Factors associated with an individual's health condition (e.g., treatment demands, side effects) may limit the ability to participate in work on a regular and continuing basis even if the person is able to perform each of the tasks associated with a job. Similarly, environmental factors (e.g., physical [built and natural], social, organizational) may limit an individual's ability to participate in work on a regular and continuing basis even if the person is able to perform the relevant work requirements. An individual's capacity to perform work requirements successfully in one specific work environment does not necessarily indicate the ability to perform the same work in a different setting.

For these reasons, the committee drew the following conclusion:

1. **Individuals' assessed functional abilities relevant to work requirements when assessed outside of actual work settings may be insufficient to establish their capacity to perform full-time work on a regular and continuing basis.**

Multiple Sources of Work-Related Functional Information

The committee determined that no single source is likely to provide all of the information needed to evaluate an individual's ability to work. Professionals in multiple disciplines administer and interpret results of assessments of physical and mental function. Convergence of information from multiple sources increases confidence in its validity, making it important to combine and evaluate the consistency of information from different sources (e.g., self-reports, quantitative measures, medical records, consultative examinations) when evaluating an individual's ability to work. The committee also found that professionals who have responsibility for repeated assessments may render more detailed and accurate evaluations of an individual's physical and/or mental functioning over time relative to medical specialists who have less frequent interactions with the person and less time per encounter during the same observation period.

Although not without limitations, standardized self-report questionnaires are an important source of information regarding the nature and severity of an applicant's functional limitations, especially when used in conjunction with other assessments. Qualitative data provided by applicants, family members, and other key sources who are sufficiently familiar with the applicant's activities, health, and functional status, in combination with a review of medical evidence, complement quantitative information that serves as the basis for disability decisions. The use of measures based on item response theory that can be administered using computer adaptive testing can decrease respondent burden by reducing survey length and administration time while minimizing measurement error. For these reasons, the committee drew the following conclusion:

2. **The validity of the results of work-related functional assessments is enhanced by a comprehensive approach that includes test results and other information about an individual's physical and mental**

functional abilities from multiple sources, as well as relevant social and environmental factors and the full scope of tasks involved in a job and sustained gainful employment.

Integrated Assessment of Work-Related Functional Ability

The committee determined that numerous validated tests are available for measuring physical and mental functional abilities at the impairment and body or organ system level, although currently no single tool, by itself, can reliably and consistently determine the inability or ability to work. Available instruments, whether based on performance, self-report, or proxy, are useful individually, but their value may be increased when different types of instruments are combined to provide a fuller picture of an individual's ability to sustain work on a regular and continuing basis, especially when they can be repeated over time. In addition, "integrated" assessment measures that provide information regarding the integrated effect of individuals' impairments on general daily life and participation are useful for capturing the additive and sometimes multiplicative effects of multiple impairments and comorbid conditions on an individual's functional ability to meet work requirements. For these reasons, the committee drew the following conclusion:

3. **Assessments that integrate information about impairments and abilities, including multiple tests of different types, repeated over time, provide the most useful information about work-related function.**

Challenges for Assessment of Work-Related Functional Abilities

Assessment of individuals' functional abilities with respect to work is more complicated than whether and how long a person can sit, stand, walk, or perform specific physical or cognitive activities. An individual's ability to perform a single work activity needs to be evaluated with respect to the context and practical relevance of his or her ability to perform work tasks effectively and hold a job, including adaptability and work-related personal interactions. The committee's conceptual framework for assessing work capacity (see Figure S-1) demonstrates the complexity and challenges of using functional measures of individuals' ability to perform specific activities and tasks, especially instruments that assess only body structures and function or impairments, to make a determination about their capacity to perform work and to sustain full-time work on a regular and continuing basis. In addition, there are a number of threats to the validity of assessments of functional abilities, including testing of maximal versus typical performance, assessment of episodic activity versus sustained task performance, absence of standardized testing conditions, mixed-motive incentives, compromised test integrity owing to prior use of the test in low-stakes testing applications, and diverse test populations on whom tests may not have been validated. Furthermore, the presence of multiple impairments and comorbidities, including symptoms associated with depression or anxiety, can further impair functioning. For these reasons, the committee drew the following conclusion:

4. **Numerous challenges complicate accurate assessment of an individual's ability to work, including the following:**
    - Measures of physiological, morphological, psychological, or cognitive severity (e.g., laboratory findings, signs, or symptoms of impairments) may not correlate with the severity of functional limitations (i.e., the effect of a condition on an individual's ability to work or conduct daily life).
    - It is simpler to demonstrate inability or limitation to perform a specific activity (e.g., reaching overhead, climbing a ladder) than to demonstrate an individual's ability to perform the combination of activities required for different occupations.
    - Tests of functional abilities often do not measure whether an individual is able to combine functions to perform tasks as needed for work.
    - Successful work performance is more than the sum of the specific tasks and skills required, and the overall limitation to successful work for an individual is often more than the sum of single impairments.
    - Threats to the validity of assessments of functional abilities include testing of maximal versus typical performance, assessment of episodic activity versus sustained task performance, absence of standardized testing conditions, mixed-motive incentives, compromised test integrity owing to prior use of the test in low-stakes testing applications, and diverse test populations on whom tests may not have been validated.
    - Symptoms associated with psychological conditions such as depression and anxiety can affect a person's ability to manage one or more limitations in a work setting. Therefore, it is necessary to consider them when assessing an individual's ability to sustain work on a regular and continuing basis because a person's capacity to work may be overestimated if a psychological comorbidity is present.

Factors Limiting the Quality and Quantity of Information on Functional Ability for an Applicant

The committee determined that functional assessment instruments vary in the degree to which they have been tested or adapted across diverse populations, making it important to consider an instrument's performance across multiple subgroups. Also, assessment instruments developed for research applications may not account for cultural, linguistic, or literacy factors, such as limited English proficiency or low literacy, which may limit access to such assessments. This may mean that few or no assessments are available that can provide valid and reliable information for these populations.

In addition, health care data relevant to disability determinations, such as the results of specific, expensive tests (e.g., certain cardiovascular tests and psychological test batteries) that are valid and potentially useful, may not be readily available because an individual may be uninsured or underinsured, or the tests may be denied by an insurance plan because they are not deemed medically necessary. Disparities in access to care and consequently poor health outcomes can affect not only the quantity of tests conducted in the context of disability determinations but also the quality of the tests and resulting information. Access to health care professionals, including those with expertise in providing information relevant to disability determination, often is limited by lower socioeconomic status and/or geographical location.

Administrative challenges such as acquisition of an applicant's clinical records are another factor that may limit available information for disability determinations. Acquiring this information may be difficult for several reasons: providers' fear of sharing confidential information, the limited capacity of a provider's organization to gather and transmit records, and high administrative costs for record transfer. For these reasons, the committee drew the following conclusion:

5. **A number of factors, including age, gender, lower socioeconomic status, race, ethnicity, cultural group, and geographic location, may limit the quality and quantity of functional information available for a disability applicant.**

REFERENCES

Merriam-Webster. 2019. *Performance*. https://www.merriam-webster.com/dictionary/performance (accessed April 3, 2019).

O*NET Resource Center. 2019. *The O*NET® content model*. https://www.onetcenter.org/dl_files/ContentModel_Summary.pdf (accessed April 3, 2019).

WHO (World Health Organization). 2001. *International Classification of Functioning, Disability and Health*. Geneva, Switzerland: WHO.

1

Introduction

The U.S. Social Security Administration (SSA) provides disability benefits through the Social Security Disability Insurance (SSDI) and the Supplemental Security Income (SSI) programs. The SSDI program, established in 1956, provides benefits to eligible adults with disabilities who have paid into the Disability Insurance Trust Fund, as well as to their spouses and adult children who are unable to work because of severe long-term disabilities. Enacted in 1972, SSI is a means-tested program based on income and financial assets that provides income assistance from U.S. Treasury general funds to adults aged 65 and older, individuals who are blind, and adults and children with disabilities. As of December 2017, SSDI had approximately 10.4 million beneficiaries and SSI about 7.1 million recipients who were classified as blind or disabled (SSA, 2018a,b).

To receive SSDI or SSI disability benefits, an individual must meet the statutory definition of disability, which is "inability to engage in any substantial gainful activity by reason of any medically determinable physical or mental impairment which can be expected to result in death or which has lasted or can be expected to last for a continuous period of not less than 12 months."[1] In determining whether the definition of disability is met for an adult, SSA uses a five-step disability evaluation process, described in detail in Chapter 2,[2] which includes consideration of the individual's functional abilities relevant to work.

[1] 42 U.S.C. 423; see also 20 CFR 404.1505; 20 CFR 416.905.
[2] See 20 CFR 404.1520; 20 CFR 416.920.

STUDY CHARGE AND SCOPE

Seeking to ensure consistency and accuracy in its disability evaluation process, SSA in 2017 asked the Health and Medicine Division of the National Academies of Sciences, Engineering, and Medicine to convene a committee of relevant experts to provide findings and conclusions regarding the collection of information and assessment of work-related functional abilities (see Box 1-1 for the committee's Statement of Task). The 15-member committee included experts in the areas of physical medicine and rehabilitation, occupational medicine, internal medicine, mental health, ergonomics, occupational therapy and vocational rehabilitation, social and behavioral science, disability law and policy, measurement and survey methodology, epidemiology, and biostatistics (see Appendix D for biographical sketches of the committee members).

In carrying out this study, the Committee on Functional Assessment for Adults with Disabilities was asked by the sponsor to perform several specific tasks, including identifying and describing evidence-based methods to collect information about an individual's physical and mental functional abilities relevant to work requirements; discussing the types of information that support findings of limitations in functional abilities relevant to work requirements; and, in the context of disability assessment, describing for functional abilities relevant to work requirements changes related to the progression of common disease processes, including but not limited to back disorders, cardiac impairments, and depression. In addition to these three conditions, the committee chose to address traumatic brain injury because of its prevalence and the associated high rates of cognitive impairment and work disability. As specified in its Statement of Task, the committee was tasked with providing findings and conclusions based on the evidence it gathered. At the committee's first meeting, SSA confirmed that it wanted the committee to provide only findings and conclusions; recommendations were not to be included in this report.

"Functional abilities relevant to work requirements" are defined by the U.S. Bureau of Labor Statistics' (BLS's) Occupational Requirements Survey and the Department of Labor for the Occupational Information System (see Chapter 2). Specific physical and mental functional abilities relevant to work requirements are discussed in Chapters 5 and 6, respectively. The committee also was asked to provide an overview of the functional assessment processes in at least three benefit programs similar to those of SSA that include assessment of disability or vocational capabilities, such as national and state government programs, private-sector programs, and programs based in other countries (see Chapter 8).

BOX 1-1
Statement of Task

An ad hoc committee will

1. Identify and describe ways to collect information about an individual's physical and mental (cognitive and noncognitive) functional abilities relevant to work requirements as defined by the U.S. Bureau of Labor Statistics' Occupational Requirements Survey (ORS), such as sitting/standing/walking, lifting/carrying, vision, communication, decision making, and adaptability;
2. Discuss the types of information that support findings of limitations in functional abilities relevant to work requirements as defined by the U.S. Department of Labor for the Occupational Information System (OIS); and
3. Provide findings and conclusions regarding the collection of information and assessment of functional abilities relevant to work requirements.

As a guide for literature review, information and data gathering, public sessions, discussions, deliberations, and report development, including findings and conclusions, the committee shall consider the following specific topics:

1. **Identify and describe ways to collect information about an individual's physical and mental (cognitive and noncognitive) functional abilities relevant to work requirements as defined by the U.S. Bureau of Labor Statistics' Occupational Requirements Survey (ORS), such as sitting/standing/walking, lifting/carrying, vision, communication, decision making, and adaptability.**
   - Provide an overview of the functional assessment processes in at least three similar benefit programs that assess disability or vocational capabilities (national and state government programs, private-sector programs, and foreign programs as applicable);
   - Provide examples of forms, tools, guides, examinations, and other resources used by benefit programs that assess functional aspects of disability and vocational capabilities;
   - Identify activities of daily living that correlate with the physical and mental (cognitive and noncognitive) demands of work; and
   - Provide examples of how to collect information on activities of daily living that correlate with the physical and mental (cognitive and noncognitive) demands of work.
2. **Discuss the types of information that support findings of limitations in functional abilities relevant to work requirements as defined by the U.S. Department of Labor for the Occupational Information System (OIS).**
   - Describe the laboratory findings, signs, or symptoms of impairments that support findings of limitations in functional abilities relevant to work requirements;
   - Explain what information, including that which pertains to level of severity and duration, may be found in medical or other evidence to support a finding that a person is unable to sustain physical and mental work activities on an ongoing and independent basis in the context of functional limitations; and

continued

BOX 1-1 Continued

- Identify any quantifiable limitations that may preclude certain levels of work (including sedentary) and give examples of the evidence to demonstrate such limitations.

3. **Provide findings and conclusions regarding the collection of information and assessment of functional abilities relevant to work requirements.**
   - Explain how limitations in functional abilities relevant to work requirements are more or less associated with particular mental or physical impairments;
   - Identify particular medical specialties and allied health fields that are likely to have the training and expertise to perform functional assessments related to work requirements;
   - Identify tools that signify a functional assessment was performed, and how likely those reports are to be valid representations of a claimant's functional limitations;
   - In the context of disability assessment, describe the spectrum of changes to functional abilities relevant to work requirements related to the progression of common disease processes in example impairments. These could include, but are not limited to, back disorders, cardiac impairments, or depression.
     – Identify where along the spectrum an individual's ability to perform functions relevant to work requirements is affected;
     – Describe whether the Social Security Administration (SSA) could expect improvement, no improvement, or progressive worsening in the example impairments;
     – Describe when significant changes in functional abilities relevant to work requirements may occur through the aging process for these examples, such as for adults with common age-related physical and mental impairments;
     – Describe the efficacy of medications and other treatments on an individual's ability to perform functional abilities relevant to work requirements for these examples, and whether that treatment causes its own subset of medical and/or psychological problems that negatively affect an individual's functioning and how SSA could request an appropriate assessment of functional changes; and
     – Describe how the examples are similar to or different from other impairments;
   - Discuss the advantages and disadvantages of generic functional assessment questionnaires that address a broad range of impairments and functional abilities relevant to work requirements, and targeted impairment-specific questionnaires, along with considerations in their use; and
   - Describe the best ways to determine the accuracy and validity of self-reported functional abilities, for example asking for input over the course of the claimant's interactions with SSA and comparing for consistency.

STUDY APPROACH

The committee conducted an extensive review of the literature pertaining to functional assessment for adults with disabilities, as well as the literature specific to assessment of function and impairment trajectories in individuals with back disorders, cardiac impairments, depression, and traumatic brain injury. This review began with a search in online databases for U.S. and international English-language literature from 1980 through 2018. This search encompassed Medline (Ovid), PubMed, Web of Science, and Scopus, as well as websites including those of SSA, the U.S. Department of Labor, BLS, and the National Academies Press (see Appendix C). A second search of the same databases was conducted for 1998 through 2018 to capture additional peer-reviewed articles and reviews not captured in the initial search (see Appendix C). A third search targeted peer-reviewed articles and review articles pertaining to specific physical assessment instruments from 1980 through 2018 in Medline (Ovid), PubMed, Web of Science, and Scopus (see Appendix C). Committee members and project staff identified additional literature and information using traditional academic research methods and online searches throughout the course of the study.

The committee used a variety of resources to supplement its review of the literature. Meeting in person five times, the committee held three public workshops to hear from invited experts in areas pertinent to its charge (see Appendix A). Speakers at the workshops included experts in functional assessment of physical and mental abilities relevant to work requirements and functional assessment tools and batteries, including the Patient-Reported Outcomes Measurement Information System (PROMIS), the National Institutes of Health (NIH) Toolbox, and the Work Disability-Functional Assessment Battery. The committee also heard from representatives of the Veterans Benefits Administration, workers' compensation insurance (Chesapeake Employers' Insurance Company, Maryland), and private disability insurance (Prudential and Sun Life Financial), who addressed the use of functional assessment in different benefit programs that assess disability or vocational capabilities, as well as representatives of several stakeholder organizations, including The Arc, Justice in Aging, Legal Services of New Jersey, and the National Organization of Social Security Claimants' Representatives. In addition, the committee commissioned three papers: on the assessment of (1) hearing, (2) speech and language, and (3) vision in the context of work requirements.

The committee's work was further informed by previous reports of the National Academies. These included *Measuring Functional Capacity and Work Requirements: Summary of a Workshop* (IOM and NRC, 1999), *Survey Measurement of Work Disability: Summary of a Workshop* (IOM and NRC, 2000), *The Dynamics of Disability: Measuring and Monitoring*

Disability for Social Security Programs (IOM and NRC, 2002), *Visual Impairments: Determining Eligibility for Social Security Benefits* (NRC, 2002), *Improving the Social Security Disability Decision Process* (IOM, 2007a), *The Future of Disability in America* (IOM, 2007b), *Cardiovascular Disability: Updating the Social Security Listings* (IOM, 2010a), *HIV and Disability: Updating the Social Security Listings* (IOM, 2010b), *A Database for a Changing Economy: Review of the Occupational Information Network (O*NET)* (NRC, 2010), *Psychological Testing in the Service of Disability Determination* (IOM, 2015), *Informing Social Security's Process for Financial Capability Determination* (NASEM, 2016), and *The Promise of Assistive Technology to Enhance Activity and Work Participation* (NASEM, 2017).

REPORT ORGANIZATION

Chapter 2 focuses on disability and function, providing further context for this report and introducing the committee's conceptual framework as well as a description of SSA's collection of the information on function and disability used in determining an individual's eligibility for benefits. Chapter 3 describes the types, sources, and quality of functional information; the properties of assessment measures; and potential sources of bias in assessments. Topics covered in Chapter 4 include integrated assessments of work-related function, the relationship between activities of daily living and the physical and mental demands of work, and instruments designed to assess limitations in work activity due to health conditions. Chapters 5 and 6, respectively, review functional assessments of physical and mental abilities relevant to work requirements. Chapter 7 addresses selected impairments and associated limitations in functional abilities relevant to work. Chapter 8 provides a review of functional assessment processes in selected benefit programs that assess disability or vocational capabilities. Finally, Chapter 9 presents the committee's overall conclusions. The report also includes four appendixes: Appendix A provides the agendas for the three public sessions held for this study; Appendix B is a glossary of terms used in the report; Appendix C gives further detail on the committee's literature searches; and Appendix D contains biographical sketches of the committee members.

REFERENCES

IOM (Institute of Medicine). 2007a. *Improving the Social Security disability decision process*. Washington, DC: The National Academies Press.

IOM. 2007b. *The future of disability in America*. Washington, DC: The National Academies Press.

IOM. 2010a. *Cardiovascular disability: Updating the Social Security listings*. Washington, DC: The National Academies Press.

IOM. 2010b. *HIV and disability: Updating the Social Security listings*. Washington, DC: The National Academies Press.

IOM. 2015. *Psychological testing in the service of disability determination*. Washington, DC: The National Academies Press.

IOM and NRC (National Research Council). 1999. *Measuring functional capacity and work requirements: Summary of a workshop*. Washington, DC: National Academy Press.

IOM and NRC. 2000. *Survey measurement of work disability: Summary of a workshop*. Washington, DC: National Academy Press.

IOM and NRC. 2002. *The dynamics of disability: Measuring and monitoring disability for Social Security programs*. Washington, DC: National Academy Press.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2016. *Informing Social Security's process for financial capability determination*. Washington, DC: The National Academies Press.

NASEM. 2017. *The promise of assistive technology to enhance activity and work participation*. Washington, DC: The National Academies Press.

NRC (National Research Council). 2002. *Visual impairments: Determining eligibility for Social Security benefits*. Washington, DC: National Academy Press.

NRC. 2010. *A database for a changing economy: Review of the Occupational Information Network (O*NET)*. Washington, DC: The National Academies Press.

SSA (U.S. Social Security Administration). 2018a. *Annual statistical supplement, 2018—OASDI current-pay benefits: Summary (5.A)*. https://www.ssa.gov/policy/docs/statcomps/supplement/2018/5a.pdf (accessed April 3, 2019).

SSA. 2018b. *Annual statistical supplement, 2018—SSI summary (7.A)*. https://www.ssa.gov/policy/docs/statcomps/supplement/2018/7a.pdf (accessed April 3, 2019).

2

Disability and Function

This chapter begins with a brief overview of different models of the concept of disability and discussion of the U.S. Social Security Administration's (SSA's) definition of disability and consideration of function within its disability determination process. The latter discussion includes a summary of SSA's collection and use of occupational information that sets the stage for identification of the physical and mental "functional abilities relevant to work requirements" the committee was asked to consider in its Statement of Task. The chapter continues with a description of the committee's ecological framework for functional assessment as it relates to work and ends with findings and conclusions. Appendix B provides definitions for many of the terms used in this discussion.

EVOLUTION OF THE CONCEPT OF DISABILITY

The concept of disability has evolved over the past several decades from a medical to a biopsychosocial model. In the medical model, disability is viewed as a feature of a person that is caused by injury, disease, or other health conditions and managed through medical treatment or modification of an individual's behavior (WHO, 2001, 2002; see also IOM, 1991; Kaplan, 2000). In contrast, the biopsychosocial model views disability as a problem at the societal level; therefore, it is the collective responsibility of society to facilitate the full participation of individuals with disabilities in all aspects of life (Kaplan, 2000; WHO, 2001). The biopsycho-ecological model expands the concept of disability further by recognizing that elements of individuals' physical and social environments may serve as barriers to or facilitators that enhance activity and participation along the disablement pathway (Stineman and Streim, 2010). For example, inaccessible buildings and transportation represent barriers to participation for individuals who use wheelchairs, while modifications such as ramps, electric doors, elevators, and the like and wheelchair-accessible public transportation serve as facilitators (NASEM, 2017).

The World Health Organization's *International Classification of Functioning, Disability and Health* (ICF) portrays human functioning and disability "as a dynamic interaction between health conditions (diseases, disorders, injuries, traumas, etc.) and contextual factors" (WHO, 2001, p. 8). Figure 2-1 depicts the ICF model of functioning and disability. As shown in the middle tier of the figure, functioning and disability consist of an interplay among body functions and structures at the organ level, activities at the person level, and participation at the societal level. The bottom tier of the figure depicts the contextual factors (environmental and personal) that may affect an individual's functioning as barriers or facilitators. Environmental factors may be individual, encompassing one's immediate environment, or societal, encompassing the "formal and informal social structures, services, and overarching approaches or systems in the community or society" that affect one's functioning. Personal factors include gender, race, ethnicity, age, social and educational background, current and past experiences, behavior patterns, and psychological assets (WHO, 2013).

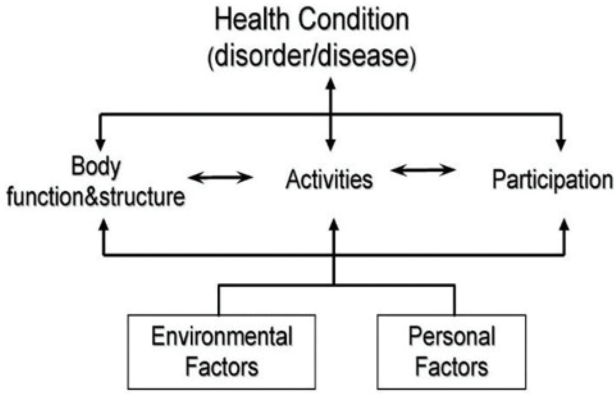

FIGURE 2-1 *International Classification of Functioning, Disability and Health* (ICF) model of functioning and disability.
NOTE: The ICF model refers to deficits in body function and structure as "impairments," deficits in completing activities as "limitations," and reductions in participation as "restrictions."
SOURCE: WHO, 2001, p. 18.

In the ICF model, disability is "an umbrella term for impairments [of body functions or structures], activity limitations and participation restrictions" (WHO, 2001, p. 213). Disability arises from "the interaction between an individual (with a health condition) and that individual's contextual factors (environmental and personal factors)" (WHO, 2001, p. 213).

The ICF model encompasses three levels of functioning (middle tier of Figure 2-1). The first level, *body function and structure*, includes the "physiological functions of body systems, including psychological functions" and the functioning of body structures, or movement of "anatomical parts of the body such as organs, limbs and their components" (WHO, 2001, p. 213). Decrements in these functions are labeled "impairments" in the ICF model. The second level of functioning is *activities*—actions or tasks performed by an individual, such as walking, lifting, keyboarding, or problem solving. The ICF calls decrements in this level of functioning "limitations." The third level of functioning is *participation*, or the performance of tasks in a societal context. Work is an example of participation. A *task* is a set of mental and physical activities in which an individual engages to accomplish a specific goal at or by a specific time. Although tasks are incorporated under activities in the ICF model, the committee recognizes them as a discrete intermediary level of functioning between activities and participation. Decrements in participation are called "restrictions" in the ICF model. Accommodations (e.g., assistive technologies, environmental modifications) are environmental contextual factors that act on ICF domains to enhance an individual's activity and participation.

SOCIAL SECURITY ADMINISTRATION'S CONSIDERATION OF FUNCTION

As stated in Chapter 1, SSA's definition of disability in adults is "inability to engage in any substantial gainful activity [SGA] by reason of any medically determinable physical or mental impairment which can be expected to result in death or which has lasted or can be expected to last for a continuous period of not less than 12 months."[1] SGA is work activity that "involves doing significant and productive physical or mental duties" or activities that are "done (or intended) for pay or profit," regardless of whether a profit is realized.[2] This work activity includes both an individual's previous work, if any, and, considering age, education, and work experience, any other kind of substantial gainful work that exists in the national economy. For SSA, "disability" refers to work disability. More broadly, work disability may be understood as an inability to participate in work on a "regular

[1] 42 U.S.C. 423; see also 20 CFR 404.1505; 20 CFR 416.905.

[2] 20 CFR 404.1510; 20 CFR 416.910; 20 CFR 404.1572; 20 CFR 416.972.

and continuing basis" (8 hours per day, 5 days per week, or an equivalent work schedule [SSA, 2018g]), which would place it under "participation" on the far right side of the ICF model (see Figure 2-1).

SSA uses a five-step evaluation process in determining whether an adult meets the definition of disability.[3]

- In the first step, SSA considers applicants' work activity in the past year. If an applicant is working at SGA, SSA will not proceed with a disability determination. SGA is defined as "work that—(a) involves doing significant and productive physical or mental duties; and (b) is done (or intended) for pay or profit."[4] To be eligible for a disability determination, in 2019 a nonblind individual must not earn more than $1,220 per month after deduction of impairment-related work expenses (SSA, 2019). Examples of impairment-related work expenses that can be deducted include certain attendant care services, medical devices, equipment, and prostheses (SSA, 2015c). For Social Security Disability Insurance applicants, insured status is verified, while for Supplemental Security Income applicants, countable income and assets are verified to be below thresholds.
- In step 2, SSA determines whether applicants have a medically determinable physical or mental impairment or a combination of impairments that meets severity and duration requirements. According to SSA's Program Operations Manual System, "when medical evidence establishes only a slight abnormality or a combination of slight abnormalities which would have no more than a minimum effect on an individual's ability to work, such impairment(s) will be found [to be] 'not severe,' and a determination of 'not disabled' will be made" (SSA, 2012). Applicants will also be denied in step 2 if their impairment neither is "expected to result in death" nor "has lasted or can be expected to last for a continuous period of not less than 12 months."[5]
- In step 3, SSA assesses an applicant's impairment using the *Listing of Impairments* (SSA, 2017b), a regulatory list of medical conditions and criteria created by SSA to assist in disability determination. The *Listing of Impairments* describes, for each of the major body systems, impairments that SSA considers to be severe enough to prevent a person from performing *any* gainful activity, regardless of his or her age, education, or work experience. Step 3 serves as a "screen-in" step. If an applicant's impairment "meets" or "equals"

[3]See 20 CFR 404.1520; 20 CFR 416.920.
[4]20 CFR 404.1510; 20 CFR 416.910.
[5]20 CFR 404.1505; 20 CFR 416.905.

a listing, the applicant is allowed benefits. To meet a listing, an applicant must have a medically determinable impairment that satisfies all of the criteria for that listing (SSA, 2018c). An impairment equals a listing if it is "at least equal in severity and duration to the criteria of any listed impairment" (SSA, 2018d). Applicants proceed to step 4 when their impairment is severe but does not meet or medically equal any listing within the *Listing of Impairments*.

- In step 4, SSA assesses whether applicants' physical or mental residual functional capacity (RFC) allows them to perform past relevant work.[6] SSA defines RFC as "the most [an applicant] can still do despite [his or her impairment-related] limitations" or restrictions[7] on "a regular and continuing basis"[8] (SSA, 2017a). Applicants able to perform past relevant work are denied benefits, while applicants unable to do so proceed to step 5.
- At step 5, SSA considers applicants' RFC and such vocational factors as age, education, and work experience, including transferrable skills in determining whether they can perform other work in the national economy. Applicants determined to be unable to adjust to performing other work are allowed benefits, while those who are determined able to adjust are denied.

In determining whether an applicant's impairment(s) is severe and meets or medically equals a listing or assessing an applicant's RFC, SSA gathers various information. This includes functional information from the applicant, relevant health care providers, and third parties about the applicant's impairment-related symptoms, such as pain, that may affect what he or she can do in a work setting. SSA considers function at several points in the five-step evaluation process. Functional criteria are built into a number of the adult listings in the *Listing of Impairments* SSA considers at step 3 (SSA, 2017b) (see Annex Table 2-1). These listings include both impairments and limitations in activity, using the ICF terminology. SSA also considers function in its assessment of applicants' physical and/or mental RFC at steps 4 and 5. In some cases, RFC, particularly mental RFC, is a refinement of activity performance taken from the *Listing of Impairments*.

SSA uses different forms to collect functional information. The job activity section of the adult disability report form (SSA-3368-BK) collects functional information related to the number of hours an employee

[6] "Past relevant work" refers to "work that you have done within the past 15 years, that was substantial gainful activity, and that lasted long enough for you to learn to do it" (20 CFR 404.1560).

[7] These are functional limitations in activities according to the ICF approach, but they may also include data on restrictions on participation in work.

[8] 20 CFR 404.1545; 20 CFR 416.945.

performed such tasks as walking, standing, sitting, climbing, stooping, kneeling, crouching, crawling, handling large objects, and reaching (SSA, 2015a). The information collected on lifting and carrying includes detail on the amount of weight lifted and how frequently (one-third to two-thirds of the workday) and how far the object had to be carried.

A second form, the function report form (SSA-3373-BK),[9] asks how the applicant's illness, injury, or condition limits his or her ability to work (SSA, 2015b). In addition, the form asks detailed questions about various daily activities. The applicant also is asked whether his or her illness, injury, or condition affects the ability to dress, bathe, care for hair, shave, self-feed, and use the toilet. Questions about meals, house- and yardwork, getting around, shopping, money, hobbies and interests, and social activities are asked as well. Also collected is information about which physical and mental abilities—including lifting, walking, memory, concentration, understanding, and following instructions—have been affected by the applicant's injury/illness.

A third form, the psychiatric review technique form (SSA-2506-BK),[10] is completed by the disability examiner or the medical consultant (MC) or the psychological consultant (PC). The MC or PC has overall responsibility for the assessment of medical severity and for the content of the form. The form includes the criteria in the mental disorders listings, and is used to record the presence or absence of the listing criteria and the rating of the degree of functional limitation.[11]

Two other forms address physical and mental RFC and are described in the following section.

Residual Functional Capacity

An RFC assessment

- is based primarily on **medical evidence** but may also include **observation or description** of limitations (e.g., lay evidence, including the claimant's statement);
- describes **what an individual is able to do,** despite functional limitations resulting from a medically determinable impairment(s) and impairment-related symptoms; and

[9]Additional forms include SSA-3380-BK, which collects the same information but from third-party sources, and SSA-3385-BK, which is used to collect information from a former employer about an applicant's functioning in a work setting. This text has been revised since prepublication release.

[10]Form SSA-2506-BK (01-2017) UF, obtained via personal communication with Joanna Firmin, Social Security Administration, February 23, 2018.

[11]This text has been revised since prepublication release.

- is an **administrative determination** of an individual's capacity to perform work-related physical and mental activities (excerpted from SSA, 2018e).

In terms of content, an RFC

a. Describes **work-related functions** a person can do on a sustained basis.
b. Addresses all **functional capabilities** and provides a **written analysis** of how the evidence in file supports or refutes the following:
   - claimant **allegations** (including allegations of symptoms, such as pain, and their effects on functioning).
   - **descriptions, observations, and all conclusions** of all treating and examining sources.
   - other pertinent **medical and nonmedical evidence.**
c. **Resolves all issues** of functional capacity pertinent to a determination of ability to do past relevant work or other work.
d. Provides **conclusions** as to functional capabilities (with a citation of pertinent evidence) and the **reasoning** to support these findings (excerpted from SSA, 2018e).

SSA MCs and PCs evaluate physical and mental RFC, respectively, although the disability examiner may assist in completion of the RFC assessment forms in some cases (SSA, 2018e). SSA uses the information about applicants' RFC to make a determination at step 4 or 5 of the disability evaluation process. If a determination is appealed, a new RFC assessment is made at each adjudicative level. Typically, RFC is assessed by the MC/PC at the initial and reconsideration level. An administrative law judge assesses the RFC at the hearing level, and an administrative appeals judge assesses RFC at the Appeals Council level.[12]

Physical RFC is based on all the relevant evidence in the applicant's case record, such as medical signs and laboratory findings, effects of treatment, reports of daily activities, lay evidence, recorded observations, medical source statements, effects of symptoms, evidence from attempts to work, need for a structured living environment, and work evaluations (SSA, 2018f). Exertional limitations on the physical RFC assessment form (SSA-4734-BK) (SSA, 2004) refer to the weight an individual can lift and/or carry, including upward pulling, and with what frequency. "Occasionally" means "from very little up to one-third of an 8-hour workday (cumulative, not continuous)," and "frequently" means "for one-third to two-thirds of an 8-hour workday (cumulative, not continuous)" (SSA, 2004). Weights range

[12]20 CFR 404.1546. This text has been revised since prepublication release.

from less than 10 pounds to 100 pounds or more (occasionally) and less than 10 pounds to 50 pounds or more (frequently). Other exertional limitations include the length of time an applicant can stand and/or walk and sit, as well as any limitation in pushing and/or pulling, including the operation of hand and/or foot controls. Postural limitations assessed include climbing (ramp/stairs, ladder/rope/scaffolds), balancing, stooping, kneeling, crouching, and crawling. Manipulative limitations include reaching in all directions, including overhead; handling (gross manipulation); fingering (fine manipulation); and feeling (skin receptors). Visual limitations are assessed for near and far acuity, depth perception, accommodation, color vision, and field of vision. Communicative limitations include hearing and speaking. Environmental exposure limitations are noted as well, for extreme cold or heat; wetness; humidity; noise; vibration; fumes, odors, dusts, gases, and poor ventilation; and hazards, such as machinery or heights.

Mental RFC again is based on all the relevant evidence in the applicant's case record. The mental RFC assessment form (SSA-4734-F4-SUP) (SSA, n.d.) asks for the PC's summary conclusions derived from the evidence in an applicant's file. Twenty mental activities in four categories (understanding and memory, sustained concentration and persistence, social interaction, and adaptation) are summarized in terms of the individual's ability to sustain them over the course of an 8-hour workday and a 5-day workweek. Each mental activity is rated as "no evidence of limitation," "not significantly limited," "moderately limited," "markedly limited," or "not ratable on available evidence." The form also contains a section for the PC to elaborate on the summary conclusions in narrative form, including any clarifying information and explanations of any conclusions that "differ from those of treating medical sources or from the individual's allegations."

Occupational Information

In addition to RFC, SSA must consider occupational information at steps 4 and 5 of the sequential evaluation process. As described previously, SSA must determine whether individuals' RFC allows them to perform past relevant work at step 4 or, along with such vocational factors as age, education, and work experience, including transferrable skills, allows them to adjust to any other work in the national economy at step 5.

SSA's primary source of information about work in the national economy is the U.S. Department of Labor's *Dictionary of Occupation Titles* (DOT), originally developed in 1938 (SSA, 2018a). The U.S. Department of Labor stopped updating the DOT in 1991 and in 1998 replaced it with the Occupational Information Network (O*NET) (SSA, 2018a), a database of occupation characteristics and worker attributes for almost 1,000 occupations (O*NET Resource Center, 2018a).

Despite the wealth of information contained in O*NET, numerous reports have concluded that the database does not fully meet SSA's disability adjudication needs (GAO, 2002a,b; IOM, 1998, 2002; NRC, 2010; Social Security Advisory Board, 2001). Accordingly, in addition to using information from O*NET, SSA currently is developing a new Occupational Information System (OIS) to serve as the main source of occupational information for its disability adjudication process (SSA, 2018a). The OIS will include occupation-related data from multiple sources[13] and will provide SSA with updated occupational information that is measured and defined in a way tailored to meet SSA's specific program needs (SSA, 2018a). As described by SSA, "the OIS will define work as it is generally performed in the national economy; it will describe the ways in which most workers carry out the typical tasks associated with the critical functions of their occupations" (SSA, 2018a).

In 2012, SSA entered into an interagency agreement with the U.S. Bureau of Labor Statistics (BLS) to collect updated occupational information for use in the OIS (SSA, 2018a). To collect this information, BLS developed the Occupational Requirements Survey (ORS), which includes data elements that capture the physical demands and environmental conditions of work, as well as the necessary vocational preparation. Although the OIS will mirror many of the data elements from the DOT, it also will provide more detail about occupational requirements (SSA, 2018b). In addition to assigning exertional levels to occupations similar to those found in the DOT, for example, the OIS will contain information about the amount of standing or walking required by jobs and whether jobs require driving, using a keyboard, and reaching overhead (SSA, 2018b). In addition, the ORS eventually will collect data on the mental and cognitive requirements of occupations, although collection of these data elements is still being refined (SSA, 2018a).[14] Some DOT elements, such as color vision and balancing, will be excluded (SSA, 2018b).

FUNCTIONAL ABILITIES RELEVANT TO WORK REQUIREMENTS

The committee's Statement of Task (see Box 1-1 in Chapter 1) includes providing information about the assessment of physical and mental "functional abilities relevant to work requirements." The Statement of Task specifically mentions functional abilities relevant to work requirements "as

[13]Sources of data include the U.S. Bureau of Labor Statistics (BLS), O*NET, and the Military Occupational Classification.

[14]The July 2017 version of the Occupation Requirements Survey (ORS) Collection Manual (DOL, 2017) included a set of "cognitive elements" that informed the committee's writing of this report. These cognitive elements were updated subsequently in an August 2018 version of the ORS Collection Manual (DOL, 2018).

defined by the U.S. Bureau of Labor Statistics' Occupational Requirements Survey (ORS)" and "as defined by the U.S. Department of Labor for the Occupational Information System (OIS)."

Occupational Requirements Survey (ORS) Data Elements

The data elements on physical and cognitive[15] demands in the ORS make up the "functional abilities relevant to work requirements as defined by the … ORS" referred to in the committee's Statement of Task.[16] BLS defines "job demands" as "the knowledge, cognitive abilities, and physical actions required to perform critical tasks, as well as environmental conditions experienced while completing critical job tasks" (DOL, 2018, p. 6).[17] Job demands include observable physical activities, such as standing, lifting, reaching, typing, and driving, as well as unobservable behaviors, such as learning and applying knowledge, perception, and problem solving (DOL, 2018, p. 6). In addition to driving, the ORS collects data on 24 physical job demands in 10 categories: sitting versus standing/walking; lifting/carrying; pushing/pulling; reaching; manipulation; keyboarding; stooping, crouching, kneeling, and crawling; climbing; speaking/hearing requirements; and vision (DOL, 2018, p. 88). The 2017 ORS Collection Manual also identifies five cognitive demand data elements: decision making, work review, pace, adaptability, and work-related personal interactions (DOL, 2017, p. 52).[18] Of these, work review (i.e., the frequency with which an employee's work is checked to ensure performance standards are being met) is not a worker characteristic or functional ability, whereas the remaining four depend on characteristics or abilities of the worker.

[15]The ORS Collection Manual (DOL, 2017, 2018) uses the term "cognitive" to refer to elements the committee considers more broadly to be mental demands of work.

[16]As previously noted, the committee's work was informed by the July 2017 version of the ORS Collection Manual (DOL, 2017), which subsequently was updated in August 2018 (DOL, 2018). References to the ORS Collection Manual have been updated to the 2018 version wherever doing so would not result in a substantive change to the report text. Because the ORS cognitive elements were substantially revised, references to the 2017 ORS Manual have been retained for those elements.

[17]Critical job tasks are activities "workers must perform to carry out their critical job function(s). A task is considered critical when it is a primary and required component of the critical job function(s)" (DOL, 2018, p. 3).

[18]The 2018 ORS Collection Manual identifies the following cognitive elements: problem solving, work review, pace, personal contacts (verbal interactions and people skills), and interactions with the general public/crowds and telework (DOL, 2018, p. 63).

Occupational Information Network (O*NET) Data Elements

Because O*NET is an additional source of data for the OIS, the abilities captured in O*NET, as well as the data elements captured by the ORS, are relevant to the committee's consideration of "functional abilities relevant to work requirements as defined … for the OIS." O*NET contains occupational data organized into six domains that encompass attributes, characteristics, and requirements of occupations and workers (O*NET Resource Center, 2018b). The three job-oriented domains are occupational requirements, including work activities and work context; workforce characteristics; and occupation-specific information. The three worker-oriented domains are worker characteristics, including abilities, occupational interests, and work values and styles; worker requirements, including knowledge, skills, and education; and experience requirements, including experience and training. Of these categories, worker abilities are of particular relevance to this study. Abilities, defined as "enduring attributes of the individual that influence performance," include cognitive, psychomotor, physical, and sensory abilities (O*NET Resource Center, 2018b).

Work Participation

The physical and mental functional abilities relevant to work requirements addressed by the committee are described in Chapters 5 and 6, respectively. It is important to note that whereas the committee's Statement of Task focuses on the functional assessment of people, the ORS, O*NET, and ultimately the OIS contain information about the physical and mental demands of jobs. For example, while the worker abilities captured in O*NET are characteristics of people, the data collected pertain to the requirements of specific occupations in terms of which abilities are required for a given job. Although the worker abilities in O*NET and the physical and proposed cognitive demands collected in the ORS may be affected by physical or mental impairments and are to some extent amenable to functional assessment, many instruments used to assess function do not necessarily correlate with individuals' ability to perform work-related activities.

Certain physical demands of jobs, such as sitting, standing, walking, lifting, and climbing, may correlate more directly than mental/cognitive demands with activities that are amenable to functional assessment. For example, a functional capacity evaluation can help determine an individual's (in)ability to perform specific physical activities (e.g., sitting, standing, walking, lifting, climbing) safely, if not the ability to perform the activities on a sufficiently sustained basis to maintain work participation on a regular and continuing basis. Similarly, individuals' visual acuity and hearing can be measured using standardized instruments. These are among the physical activities or characteristics of individuals that may also be identified as physical demands or requirements of jobs. But the correlation between an individual's measurable mental/cognitive abilities and, for example, the four relevant cognitive job demands defined in the ORS is less straightforward. The terms or concepts used to describe the cognitive demands of jobs differ from those typically used in the functional assessment of individuals' mental abilities. The 2017 ORS Collection Manual specifies three levels of decision making[19] and directs the collection of information on "the highest level of independent judgment a worker is expected to use to perform the critical tasks of the occupation" (DOL, 2017, p. 53). Individuals' ability to use independent judgment as specified in the three levels of decision making defined in the ORS is not something health professionals specifically assess when they conduct cognitive assessments. Similarly, assessment of individuals' functional abilities with respect to adaptability and work-related personal interactions is more complex than assessment of whether and how long an individual can sit, stand, or walk. The less direct the correlation between assessment measures for physical and mental functional abilities and the work requirements captured by the ORS for the OIS, the greater will be the challenge for cross-walking between the functional assessment of individuals' abilities and the demands of jobs.

Another challenge is the extrapolation from assessment of functional abilities ("activities" in ICF parlance) to the ability to perform tasks or to meta-task as required for work participation. A person may be capable of performing each activity or even each task separately, but not be able to coordinate and sequence them together. In addition, most functional assessments capture a person's functional abilities at a particular point in time. A person may successfully pass a variety of episodic assessments but, as discussed further below, be unable to engage in repetitive or continued performance over time. The relationship among activities, tasks, and ability to perform work is developed further in the following section.

[19] "1. Employee uses independent judgment to select from a limited number of predetermined actions.... 2. Employee uses independent judgment to determine the most appropriate course of action in situations that do not have set responses.... 3. Employee uses independent judgment to make decisions by choosing from a large number of possibilities in situations where a high degree of uncertainty or complexity may exist" (DOL, 2017, p. 53).

AN ECOLOGICAL FRAMEWORK FOR FUNCTIONAL ASSESSMENT FOR WORK

Activities, Tasks, and Ability to Perform Work

Evaluation of the ability to perform a single work activity needs to be viewed with respect to the context and practical relevance of an individual's being able to perform work and hold a job. For example, an individual might successfully pass a number of functional tests for a given, single activity, such as lifting a 10-pound box in a controlled and coached environment; however, this same individual might prove unable to engage in continued performance of this activity because of his or her health condition (e.g., cardiac condition). Environmental factors, both physical (e.g., light, noise, air quality) and organizational (e.g., work policy, psychosocial factors), also may make individuals incapable of engaging in continuous performance of work-related activities because of their heath condition. For example, an individual with multiple sclerosis may be able to do work when it is cool but unable to do it in heat (air-conditioned versus outside), or a person with significant distractibility or attention problems may be able to work in a quiet setting but not in a highly stimulating setting.

SSA is interested in an individual's ability to perform sustained work-related physical and mental activities in a work setting on a "regular and continuing basis," meaning 8 hours per day, 5 days per week, or an equivalent work schedule (SSA, 2018g).[20] As a result, information from functional assessments needs to be considered within the context of completing a job.

Hierarchy of Job and Tasks Analyses

The hierarchy of job and task analyses, together with considerations of ergonomics and human factors, provides a framework for describing how single functional assessments fit into the context of sustained work (see Figure 2-2). Jobs are comprised of tasks, which in turn are comprised of activities. At the highest level, job analysis provides information on the overall job requirements and the work performed. The analysis creates a list of tasks that are the requirements or responsibilities for the job. These tasks are analogous to the critical job tasks identified for different occupations in the OIS. For example, the job analysis of a worker managing the drive-through window of a limited-service restaurant may list such tasks as take and enter the order of the customer, collect the items of the order, process payment for the order (cash, credit card), and deliver the order to

[20]It should be noted that many jobs in reality require a longer workweek, which could affect a worker's ability to return to his or her job at the prior level of work.

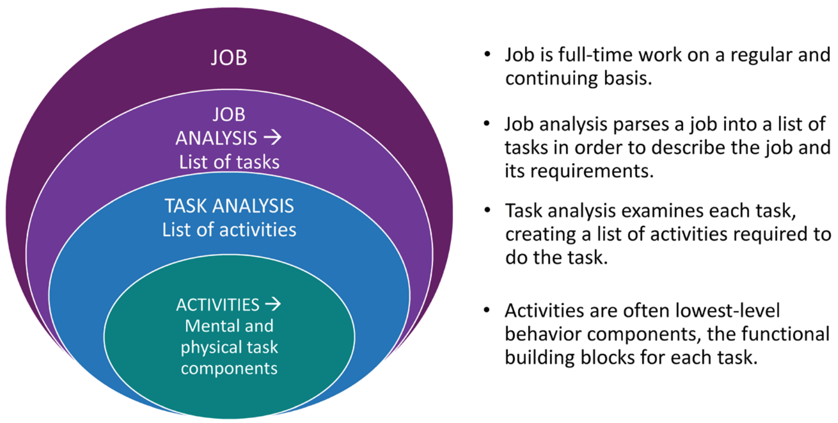

FIGURE 2-2 The context of work described through the hierarchy of job and task analyses.

the customer.[21] Task analysis then creates a list of activities (mental and physical) required to accomplish the goal of a specific task. Continuing with the above example, collecting the items of the order will include remembering the components of the order, walking to the food station (drink, sandwich, or fryer), lifting and carrying the items, and visually recognizing and inspecting the items.

Applying the *International Classification of Functioning, Disability and Health* (ICF) Disability Model

The concepts of the contextual factors surrounding work and the hierarchy of job and task analyses can be applied to the ICF model of functioning and disability presented earlier in Figure 2-1, providing a conceptual or ecological framework for functional assessment for work (see Figure 2-3). The framework provides a way to organize various sources of information and specific types of tools to assess function as described in subsequent chapters. As previously discussed, the major components of the ICF model include health conditions, body function and structure, and activities and participation. Within the context of the committee's charge, the specific health conditions include back disorders, cardiac impairments, and depression, to which the committee added traumatic brain injury (see Chapter 7). The

[21]A limitation of this job analysis example is that the list of tasks omits the sequencing of these tasks for each customer, that the sequence for one customer overlaps with the sequences for other customers, and that larger cognitive management tasks are involved in keeping the orders of different customers organized.

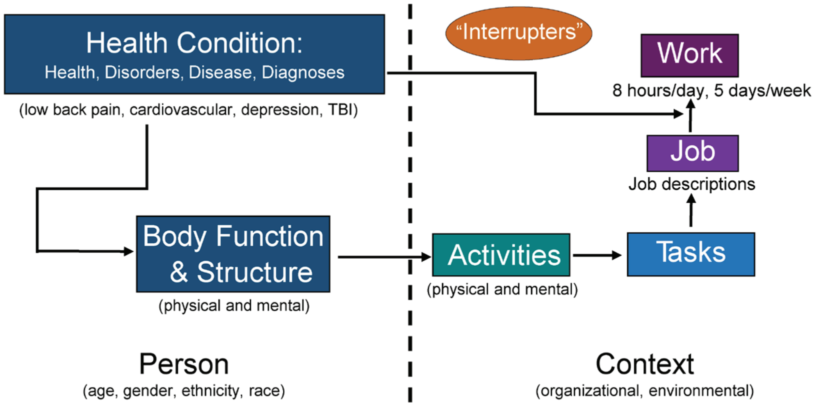

FIGURE 2-3 A conceptual framework for functional assessment of an individual's capacity for work.
NOTES: The framework provides a structure and a hierarchy for moving from functional assessment of an individual to his or her capacity to perform work. Some components are influenced more by the person (left of the dotted line), while others are influenced more by work factors (right of the dotted line). These factors can overlap and interact. "Interrupters" are factors associated with the individual's health condition and its treatment (e.g., medication effects, fluctuations in symptoms) that may limit his or her ability to perform work activities on a sustained basis. TBI = traumatic brain injury.

committee's framework hence considers both mental and physical components for body function and structure and for specific work-related activities. In accordance with the hierarchy of job and task analyses discussed above, the committee added *tasks*—the combination of both mental and physical activities required to accomplish a specific goal (subsumed under *activities* in the ICF model)—as a separate component of its framework (see the further discussion below). For purposes of this report, because SSA's focus is work disability, participation refers to participating in work and being able to hold a job (at least at the level of SGA). In turn, participating in work and being able to hold a job refer to an individual's ability to perform sustained work-related physical and mental activities in a work setting on a "regular and continuing basis" (8 hours/day, 5 days/week [SSA, 2018g]). The arrows linking tasks to work in Figure 2-3 capture the concepts of task coordination, task sequencing, and other meta-task processes as described previously.

To capture the context of work, the committee's framework adds the hierarchy of job and task analyses between function and work. As noted, the ICF model considers tasks to be part of activities, where activity is an individual's execution of a task or action (Jette, 2006, 2009). However, distinguishing task from activity may be appropriate for the work setting to accord with the functional assessment tools, as well as some of the ORS and O*NET categories. The distinction between task and activity is not absolute and mutually exclusive, however. The concepts and uses of the terms overlap throughout the literature. The two concepts do, however, provide a structure and hierarchy bridging functional assessment of an individual and his or her capacity to perform work. Table 2-1 shows the relationship among terms and concepts from the committee's conceptual framework, SSA, and the ORS.

To further capture the context of work, specifically pertaining to the ability to perform sustained work-related physical and mental activities in a work setting on a regular and continuing basis, the framework adds "interrupters." These are factors associated with the individual's health condition and its treatment that interfere with the ability to perform sustained work activities on a regular and continuing basis. For example, an individual with end-stage renal disease may be quite able to perform all of the physical and mental activities and tasks associated with a job, but require dialysis treatment that interrupts his or her ability to work for 8 hours per day, 5 days per week. Other examples include medications or fluctuations in a condition that prevent or limit the ability to perform sustained work activities and tasks. Another way to think about interrupters is that the health condition and the demands of its treatment conflict with the demands of

TABLE 2-1
Terms from the Committee's Conceptual Framework Cross-Linked with Terms and Concepts of the Social Security Administration and the Occupational Requirements Survey

Conceptual Framework	Social Security Administration	Occupational Requirements Survey
Health condition; body function and structure	*Listing of Impairments*	—
Activities	*Listing of Impairments* (activity limitations) Residual functional capacity (functional abilities relevant to work requirements)	Physical and cognitive demands
Tasks	Ability to perform elements of a job (not observed)	Task lists (critical tasks)
Job/work	Ability to maintain a job at the level of substantial gainful activity or above (observed only for those that work)	Job descriptions (critical job functions)

work. Such conflict is similar to the concept of work–family conflict,[22] in which familial responsibilities conflict with and add stress to a worker's employment. Stressors may interfere with individuals' ability to sustain work performance especially when they lack adequate stress management skills. Similarly, the imbalance in the work–health interface resulting from an interrupter can create additional, conflicting demands and stress that limit one's participation in sustained day-to-day work activities.

Individual and work environment factors can contribute to the associations among the framework's components. Some components are influenced more by the person (left of the dotted line in Figure 2-3), while others are influenced more by work factors (right of the dotted line). These two sets of factors can overlap and interact—hence a dotted line. An increasing number of studies demonstrate that environmental, physical, psychosocial, and organizational factors add to the variability of a person's capacity to participate in work (DOL, 2019; Shaw et al., 2017). Modifications to community and work environments permit work participation by many individuals who otherwise would be unable to do so (NASEM, 2017). For example, public transportation, including wheelchair-accessible buses, permit travel to work for many individuals with mobility impairments, an issue in areas where public transportation is infrequent or nonexistent (Bezyak et al., 2017). Likewise, worker adaptation programs offered by employers and employee assistance programs can help workers navigate challenges posed by their conditions and retain employment.

Conversely, an individual's capacity to hold a particular job in one environment may not extend to the ability to hold the same job in a different environment (e.g., because of the variability in workplace or office location, company, management, and co-workers). In addition, such environmental factors can change over time within the same job. For example, an autistic individual who can successfully perform a computer programming job in one location may not be able to do so in another location where he or she finds the background noise and/or lighting distracting, or actually painful. Likewise, an individual with a respiratory condition who can successfully perform the tasks of a security guard checking IDs in the lobby of a clean office building may not be able to perform those same tasks in a factory that manufactures dusty or dirty substances that negatively affect his or her respiratory condition.

[22]Work–family conflict occurs when an individual experiences incompatible demands of work and family roles, causing participation in both roles to become more difficult (Greenhaus and Beutell, 1985).

Parallels with SSA and ORS Terms and Concepts

As shown in Table 2-1, the committee's conceptual framework has parallels in the terms used by SSA in its current approaches for assessing disability. Both health condition and body function and structure correlate with SSA's *Listing of Impairments*, with many of the listings including limitations in activities along with impairment criteria (see Annex Table 2-1), as previously noted. The list of activities is associated with activity limitations in the *Listing of Impairments* and SSA's assessment of RFC. The concept of tasks is related to SSA's assessment of the capacity to perform elements of a job, which for the most part is not observed, although there may be some computer-assisted evaluations that act as proxies for work tasks. Finally, work correlates with the ability to maintain a job at the level of SGA or above, which is observed only for individuals who work.

The components of the committee's conceptual framework also parallel components of the ORS (DOL, 2017, 2018) (see Table 2-1). Many of the cognitive and physical job demand data elements in the ORS (DOL, 2017, p. 52ff, 2018, p. 88ff) fall in the area of mental and physical activities in the framework. In addition, several of the cognitive elements, such as pace, adaptability, and personal interactions, also relate to the meta-tasking components of work.

The ORS Collection Manual describes classifying jobs according to O*NET codes and then provides task lists (DOL, 2018, p. 18ff). For example, a critical job function for a hairdresser is to provide beauty services relating to clients' requests. The job tasks associated with this job function include shampooing, cutting, coloring, and blow drying hair for men, women, and children; recommending styling products; perming hair; waxing eyebrows and facial hair; and creating up-dos for special occasions, such as weddings or proms (DOL, 2018, p. 30). Job task lists such as this fall under the area of job descriptions in the committee's framework.

Finally, environmental components of work, including temperature, humidity, hazardous contaminants, moving machinery and mechanical parts, heights, vibration, and noise (DOL, 2018, p. 133ff), are captured under "context" in the committee's framework. These components are limited to the physical work environment, however, and do not encompass other environmental and organizational factors often associated with injury and disability inside or outside of the work setting (Sorensen et al., 2016).

FINDINGS AND CONCLUSIONS

Findings

2-1. Since the statutory language for the determination of disability by the U.S. Social Security Administration (SSA) was formulated, models, nomenclature, and language concerning disability have evolved, and they continue to do so.

2-2. Current models of disability, such as the *International Classification of Functioning, Disability and Health* (ICF) model, consider disability to involve the effects (limitations) an individual's health condition places on his or her ability to function and participate fully in society.

2-3. SSA considers functional information at several points in the disability determination process.

2-4. Although the worker abilities in the Occupational Information Network (O*NET) and the physical and proposed cognitive demands collected in the Occupational Requirements Survey (ORS) may be affected by physical or mental impairments and are to some extent amenable to functional assessment, many instruments used to assess function do not necessarily correlate with individuals' ability to perform work-related activities. In addition, certain physical demands of jobs, such as sitting, standing, walking, lifting, and climbing, may correlate more directly than mental/cognitive demands with activities that are amendable to functional assessment.

2-5. Assessment of individuals' functional abilities with respect to adaptability and work-related personal interactions is more complicated than assessment of whether and how long an individual can sit, stand, or walk.

2-6. Extrapolation from assessment of functional abilities ("activities" in ICF parlance) to the ability to perform tasks or to meta-task as required for work participation is a challenge.

2-7. To capture the context of work, the committee's conceptual framework for functional assessment for work adds the hierarchy of job and task analyses between function and work and takes account of personal and contextual (organizational and environmental) factors that influence individuals' capacity to perform sustained work activities.

2-8. The committee's conceptual framework includes "interrupters," factors associated with an individual's health condition and its treatment that limit the ability to perform sustained work activities on a regular and continuing basis.

2-9. Assessment of the capacity of an individual to work and to sustain full-time work on a regular and continuing basis encompasses many factors that often go beyond whether the person can complete specific individual physical and mental activities or tasks.

2-10. The nature of job requirements has changed over time as the types of jobs available in the national economy have evolved. The Occupational Information System currently in development for use by SSA reflects these changes in jobs and job requirements and is designed to be updated regularly.

2-11. Compared with the committee's conceptual framework and SSA definitions, the ORS includes more detailed lists of critical tasks and job functions that go beyond impairments (body function and structure and health conditions) and activities considered in isolation.

Conclusions

2-1. In keeping with current models of disability, assessment of individuals' functional abilities relevant to work requirements is an important part of determining whether they are able to meet workplace demands and sustain work performance on a regular and continuing basis.

2-2. The evolution of models, nomenclature, and language concerning disability since the statutory language required for use by SSA was formulated makes it difficult to reconcile differences, recognize commonalities, and integrate the conceptual changes into the disability determination process. Yet, SSA's disability determination process includes consideration of functional information at several points.

2-3. The committee's conceptual framework for assessing work capacity demonstrates the complexity and challenges of functional assessments, especially the use of instruments that assess only body and structure function or impairment, in extrapolating from individuals' ability to perform specific activities and tasks to their capacity to perform work and to sustain full-time work on a regular and continuing basis.

REFERENCES

Bezyak, J. L., S. A. Sabella, and R. H. Gattis. 2017. Public transportation: An investigation of barriers for people with disabilities. *Journal of Disability Policy Studies* 28(1):52–60. doi: 10.1177/1044207317702070.

DOL (U.S. Department of Labor). 2017. *Occupational Requirements Survey (ORS): Collection manual*, 3rd ed. Washington, DC: U.S. Bureau of Labor Statistics. https://www.bls.gov/ors/about/information-for-survey-participants/pdf/occupational-requirements-survey-collection-manual-072017.pdf (accessed April 3, 2019).

DOL. 2018. *Occupational Requirements Survey (ORS): Collection manual*, 4th ed. https://www.bls.gov/ors/about/information-for-survey-participants/pdf/occupational-requirements-survey-collection-manual-082018.pdf (accessed April 4, 2019).

DOL. 2019. *S@W/R2W research & RETAIN demonstration projects: Research & publications*. https://www.dol.gov/odep/topics/SAW-RTW/research-publications.htm (accessed April 3, 2019).

GAO (U.S. General Accounting Office). 2002a. *SSA and VA disability programs: Re-examination of disability criteria needed to help ensure program integrity.* http://www.gao.gov/new.items/d02597.pdf (accessed April 3, 2019).

GAO. 2002b. *SSA disability programs: Fully updating disability criteria has implications for program design.* http://www.gao.gov/new.items/d02919t.pdf (accessed April 3, 2019).

Greenhaus, J. H., and N. J. Beutell. 1985. Sources and conflict between work and family roles. *Academy of Management Review* 10(1):76–88.

IOM (Institute of Medicine). 1991. *Disability in America: Toward a national agenda for prevention.* Washington, DC: National Academy Press.

IOM. 1998. *The Social Security Administration's decision process: A framework for research, 2nd interim report.* Washington, DC: National Academy Press.

IOM. 2002. *The dynamics of disability: Measuring and monitoring disability.* Washington, DC: The National Academies Press.

Jette, A. M. 2006. Toward a common language for function, disability, and health. *Physical Therapy* 86(5):726–734.

Jette, A. M. 2009. Toward a common language of disablement. *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences* 64(11):1165–1168. doi: 10.1093/gerona/glp093.

Kaplan, D. 2000. The definition of disability: Perspective of the disability community. *Journal of Health Care Law and Policy* 3(2):352–364.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2017. *The promise of assistive technology to enhance activity and work participation.* Washington, DC: The National Academies Press.

NRC (National Research Council). 2010. *A database for a changing economy: Review of the Occupational Information Network (O*NET).* Washington, DC: The National Academies Press.

O*NET (Occupational Information Network) Resource Center. 2018a. *About O*NET.* https://www.onetcenter.org/overview.html (accessed April 3, 2019).

O*NET Resource Center. 2018b. *The O*NET® content model.* https://www.onetcenter.org/dl_files/ContentModel_Summary.pdf (accessed April 3, 2019).

Shaw, W., R. J. Gatchel, J. Christian, and L. Toms Barker. 2017. Improving pain management and support for workers with musculoskeletal disorders: Policies to prevent work disability and job loss, Vol. 1 – Policy Action Paper. *SAW/RTW Policy Collaborative*, U.S. Department of Labor, Office of Disability Employment Policy.

Social Security Advisory Board. 2001. *Charting the future of Social Security's disability programs: The need for fundamental change.* http://www.ssab.gov/Details-Page/ArticleID/153/Charting-the-Future-of-Social-Securitys-Disability-Programs-The-Need-for-Fundamental-Change-January-2001 (accessed April 3, 2019).

Sorensen, G., D. L. McLellan, E. L. Sabbath, J. T. Dennerlein, E. M. Nagler, D. A. Hurtado, N. P. Pronk, and G. R. Wagner. 2016. Integrating worksite health protection and health promotion: A conceptual model for intervention and research. *Preventive Medicine* 91(October):188–196. doi: 10.1016/j.ypmed.2016.08.005.

SSA (U.S. Social Security Administration). 2004. *Physical residual functional capacity assessment form SSA-4734-BK.* https://secure.ssa.gov/apps10/poms/images/SSA4/G-SSA-4734-U8-1.pdf (accessed April 3, 2019).

SSA. 2012. *POMS DI 24505.001: Individual must have a medically determinable severe impairment*. https://secure.ssa.gov/poms.nsf/lnx/0424505001 (accessed April 3, 2019).

SSA. 2015a. *Disability report adult form SSA-3368-BK*. https://www.ssa.gov/forms/ssa-3368.pdf (accessed April 3, 2019).

SSA. 2015b. *Function report adult form SSA-3373-BK*. https://www.ssa.gov/forms/ssa-3373-bk.pdf (accessed April 3, 2019).

SSA. 2015c. *POMS DI 10520.001: Impairment-Related Work Expenses (IRWE)*. https://secure.ssa.gov/poms.nsf/lnx/0410520001 (accessed April 3, 2019).

SSA. 2017a. *Adult functional assessment in Social Security disability claims*. Presentation to the Committee on Functional Assessment for Adults with Disabilities, Washington, DC, December 7.

SSA. 2017b. *Listing of impairments—Adult listings (Part A)*. https://www.ssa.gov/disability/professionals/bluebook/AdultListings.htm (accessed April 3, 2019).

SSA. 2018a. *Occupational information system project*. https://www.ssa.gov/disabilityresearch/occupational_info_systems.html (accessed April 3, 2019).

SSA. 2018b. *OIS project frequently asked questions*. https://www.ssa.gov/disabilityresearch/ois_project_faqs.html (accessed April 3, 2019).

SSA. 2018c. *POMS DI 24508.005: Impairment meets a listing*. https://secure.ssa.gov/apps10/poms.nsf/lnx/0424508005 (accessed April 3, 2019).

SSA. 2018d. *POMS DI 24508.010: Impairment or combination of impairments equaling a listing—Medical equivalence*. https://secure.ssa.gov/apps10/poms.nsf/lnx/0424508010 (accessed April 3, 2019).

SSA. 2018e. *POMS DI 24510.001: Residual functional capacity (RFC) assessment—Introduction*. https://secure.ssa.gov/poms.nsf/lnx/0424510001 (accessed April 3, 2019).

SSA. 2018f. *POMS DI 24510.006: Assessing residual functional capacity (RFC) in initial claims (SSR 96-8P)*. https://secure.ssa.gov/poms.nsf/lnx/0424510006 (accessed April 3, 2019).

SSA. 2018g. *SSR 96-8P: Policy interpretation ruling Titles II and XVI: Assessing residual functional capacity in initial claims*. https://www.ssa.gov/OP_Home/rulings/di/01/SSR96-08-di-01.html (accessed April 3, 2019).

SSA. 2019. *Substantial gainful activity*. https://www.ssa.gov/oact/cola/sga.html (accessed April 3, 2019).

SSA. n.d. *Mental residual functional capacity assessment form SSA-4734-F4-SUP*. http://ssaconnect.com/tfiles/SSA-4734-F4-SUP.pdf (accessed April 3, 2019).

Stineman, M. G., and J. E. Streim. 2010. The biopsycho-ecological paradigm: A foundational theory for medicine. *Physical Medicine and Rehabilitation* 2(11):1035–1045. doi: 10.1016/j.pmrj.2010.06.013.

WHO (World Health Organization). 2001. *International classification of functioning, disability and health*. Geneva, Switzerland: WHO.

WHO. 2002. *Towards a common language for functioning, disability and health: ICF*. Geneva, Switzerland: WHO.

WHO. 2013. *How to use the ICF: A practical manual for using the International Classification of Functioning, Disability and Health*. Geneva, Switzerland: WHO.

ANNEX TABLE 2-1
Examples of Functional Criteria in the Social Security Administration's Adult *Listing of Impairments*

Listing	Examples of Functional Criteria
1.00 Musculoskeletal System	• "inability to ambulate effectively" • "inability to perform fine or gross movements effectively"
2.00 Special Senses and Speech	• "inability to produce by any means speech that can be heard, understood, or sustained"
4.00 Cardiovascular System	• "very serious limitations in the ability to independently initiate, sustain, or complete activities of daily living" • "inability to perform on an exercise tolerance test at a workload equivalent to 5 METS or less"
7.00 Hematological Disorders	• Marked level of limitation in one of — "activities of daily living" — "maintaining social functioning" — "completing tasks in a timely manner due to deficiencies in concentration, persistence, or pace"
8.00 Skin Disorders	• "inability to function outside of a highly protective environment"
10.00 Congenital Disorders That Affect Multiple Body Systems	• "evidence demonstrating ... function at a level consistent with non-mosaic Down syndrome"
11.00 Neurological Disorders	• "disorganization of motor function in two extremities ... resulting in an extreme limitation ... in the ability to stand up from a seated position, balance while standing or walking, or use the upper extremities" • Marked limitation in — "physical functioning" — "understanding, remembering, or applying information" — "interacting with others" — "concentrating, persisting, or maintaining pace" — "adapting or managing oneself" • "ineffective speech or communication"
12.00 Mental Disorders	• "significant cognitive decline from a prior level of functioning in one or more of the cognitive areas" — complex attention — executive function — learning and memory — language — perceptual-motor — social cognition • "extreme limitation of one, or marked limitation of two, of the following areas of mental functioning" — "understand, remember, or apply information" — "interact with others" — "concentrate, persist, or maintain pace" — "adapt or manage oneself" • "minimal capacity to adapt to changes in ... environment or to demands that are not already part of [individual's] daily life"

continued

ANNEX TABLE 2-1
Continued

Listing	Examples of Functional Criteria
14.00 Immune System Disorders	• "inability to ambulate effectively" • "inability to perform fine or gross movements effectively" • Marked level of limitation in one of — "activities of daily living" — "maintaining social functioning" — "completing tasks in a timely manner due to deficiencies in concentration, persistence, or pace"

SOURCE: SSA, 2017b.

3

Collection of Information on Function and Disability

The U.S. Social Security Administration (SSA) makes a determination about whether an individual is disabled after considering all of the relevant evidence in the applicant's case record.[1] SSA currently defines evidence as any information related to an individual's application that is submitted by the applicant or anyone else for consideration by SSA, as well as information the agency obtains while developing the claim (SSA, 2018a). The categories of evidence are objective medical evidence (signs, laboratory findings, or both), medical opinion, other medical evidence from medical sources, evidence from nonmedical sources, and prior administrative medical findings of state or federal Disability Determination Services medical and psychological consultants (SSA, 2018b). The applicant and individuals connected to him or her bear the burden of proof of a medically determinable impairment and associated limitations in activities. When no relevant records are available from a treating medical source, SSA has an obligation to obtain one or more consultative examinations. The extent and types of medical evidence in an applicant's file likely will be affected by the availability and cost of specific tests.

This chapter begins with an overview of the types, sources, and quality of information about individuals' functioning that SSA may encounter in its consideration of evidence in an applicant's case record. Next is a discussion of psychometric and other properties of measurement instruments, which is followed by sections on professionals with the training and expertise to

[1]20 CFR 404.1520; 20 CFR 416.920.

assess function and potential threats to the validity of functional assessment. The final section presents findings and conclusions.

Various sources of information and specific types of tools to assess function can be mapped to the committee's conceptual framework presented in Chapter 2. For example, medical records typically contain information about the individual's specific health issue and its manifestation in body function and structure. The results of X-rays and various types of scans, for example, provide information on body structure, and performance-based tests and instruments typically provide information on body function and ability to perform physical and mental activities or tasks. Self-report instruments can provide information on symptoms (e.g., pain, fatigue, anxiety); performance of physical and mental activities and tasks; and factors, including "interrupters," that interfere with an individual's ability to perform work on a regular and continuing basis.

TYPES, SOURCES, AND QUALITY OF FUNCTIONAL INFORMATION

Collecting information about individuals that is useful for evaluating function and work disability is a complex process requiring the careful application of rigorous, modern techniques and selection of the best information collection methods, measurement instruments, and performance measures. This section provides an overview of these tools and related considerations, including the strengths and weaknesses of different information collection methods; the pitfalls of conducting clinical interviews; the use of proxy respondents; the exploitation of clinical records, including electronic health records (EHRs); the conduct of clinical and functional observations; and the direct measurement of function. The committee notes that the collection and handling of clinical information has important ethical dimensions (e.g., informed consent for data release in accordance with the provisions of the Health Insurance Portability and Accountability Act) (NRC, 2007), but it is assumed here that SSA has substantial experience with these issues.

The Building Blocks of Health Measurement

The practice of clinical medicine and the clinical records that ensue employ a vocabulary (nomenclature) used to describe patient complaints and health reports and express the diagnostic conclusions of health professionals, as reported in the clinical records. This vocabulary includes standard terms used for the building blocks of health measurement, such as symptoms (abnormal bodily perceptions), signs (visible abnormalities), abnormal bodily movements or more complex behaviors (observed by patients, clinicians, and others), abnormalities in function (disabilities reported by patients or other observers), and physiological measures usually made by professionals in the health care system (e.g., biomarkers, imaging procedures, and other physiological measures). The findings in these areas are then gathered, as actual or provisional diagnoses, into known or predicted clusters (i.e., syndromes) or formally designated conditions. While the focus of the following discussion is on the assessment of function and disability, it is important to note that essentially all of the aforementioned building blocks of health measurement are part of standard nomenclature systems, such as the *International Classification of Diseases* or the many resources of the U.S. National Library of Medicine, that can be important in evaluating medical records with digital tools employing machine-reading techniques (e.g., text mining, natural language processing,[2] data mining), as discussed later in the chapter. It is also important to note the need to focus on the coexistence and interaction of multiple disabling conditions as part of the disability evaluation process (Oni et al., 2014; Ubalde-Lopez et al., 2016).

General Considerations in Information Collection

Although the accuracy of the information collected for disability evaluation depends primarily on the reliability and validity of the collection method used, as described below, some general considerations apply to nearly all collection methods related to function. First, characteristics of the individual may affect the accuracy of the information collected. Differences in gender, race, ethnicity, and culture can affect individuals' perceptions of illness and their reporting of relevant health information (Anglin et al., 2008; Carpenter-Song et al., 2010; Forestier et al., 2019; Fuentes and Aranda, 2018; Zdunek et al., 2015). People also have different levels of general and health literacy, educational attainment and language experience, and cognitive or speech impairments (e.g., dyslexia), for example, which may affect their ability to convey information about their condition. Some difficulties in individuals' reporting of information may also be related to the underlying condition being evaluated, whether physical or psychiatric, while others may be related to the adverse effects of various medical treatments, particularly medications with psychotropic properties. In addition, individuals may be subject to social or other stressors, whether related to work or not, that can impact the quality of the information they provide. In general, even when validated information collection methods

[2] In simple terms, natural language processing is "a branch of artificial intelligence that helps computers understand, interpret and manipulate human language" (SAS Institute, 2019a), such as the free text in clinical health records.

are employed, it is often difficult to determine the causes of errors in applying so-called standard methods.

Second, the setting in which information is collected is important. Standardized laboratory settings with published protocols are usually necessary for performance testing, to ensure the reproducibility and accurate interpretation of the test results. In many ways, the same is true for the collection of information through interviews, and particularly for psychological testing. Test settings require good lighting, absence of noise, and freedom from important distractions, as well as special accommodations for persons with disabilities that can impede testing procedures. Characteristics of respondents and data acquisition methods that may alter the findings of information collection are called "response variables" in survey research.

Types of Functional Assessment Measures and Sources of Information

Sources of information for assessment of function include (1) clinical records, (2) performance-based measures, and (3) self- or proxy-reported measures. Each of these three sources has strengths and weaknesses, and the results of one are often used to validate those of another (Oude Voshaar et al., 2019; Schalet et al., 2015). The use of all three sources entails certain assumptions, such as the stability of the environment used for diagnostic testing and conditions in that environment that are conducive to the testing process, and the use of a measurement instrument with sufficiently strong and established psychometric properties to meet conventional standards. In addition, for performance-based and self- or proxy-report measures, it is assumed that the respondent has sufficient cognitive and reading skills to understand and follow directions and that the language and references in the performance-based and self- or proxy-report measures are appropriately compatible with the respondent.

Clinical Records

Among the information found in an individual's health record are the results of diagnostic testing. In addition to the results of physiological measures, such as blood pressure measurements, blood tests and imaging procedures, and any functional tests performed, clinical records provide information about diagnoses and treatments, including prescribed medications, appliances, and devices. Professionally witnessed physical or behavioral events or characteristics of patients also may be documented (e.g., seizures or spasticity), although occasional episodic events may not occur during a clinical examination. Clinical records, including results of directed laboratory testing, may include information on biomarkers that are physiologically or biochemically relevant to an individual's health and functional status and disability-related medical condition. Clinical records also may provide information over extended periods of time, yielding insight into clinical and functional trajectories of the person's condition.

Despite their importance as an information source, clinical records have many actual or potential limitations that need to be recognized. The records may, for example, be derived from many institutional sources and have varying formats, content, and authors. At times, clinical findings in the records may be unclear or contradictory, a problem exacerbated by differences in clinical observers and the ways in which assessments are performed and documented. And many elements of a clinical record, particularly the medical history, are by definition based on patient self-report, so they may be subjective and unverified. Other potential problems with clinical records are worth noting. Diagnoses may be incomplete or not sufficiently accurate. Some records contain only tentative diagnoses, pending further evaluation. Some types of health considerations are known to be frequently missing, such as safety events that occur in hospitals and adverse drug reactions. A common problem is that records of consultations that occur outside of the parent institution, such as for rehabilitation or psychiatric treatment, may not be present.

Also, the availability of specific tests (e.g., certain cardiovascular tests and psychological batteries) that are valid and potentially useful to disability evaluations may be limited by costs and individuals' access to specialized health care. Relevant health care data may not be easily available because an individual may lack insurance coverage or be underinsured, or the means of obtaining the information needed may be denied by insurance as not medically necessary. Health and care disparities can have a significant impact on the collection of health information available to inform disability determinations. In the United States, lower socioeconomic status is associated with less access to high-quality care and health care professionals (AHRQ, 2018; IOM, 2001, 2003), including those with expertise in providing information on functional status relevant to disability determinations. Thus, disparities in access to care and health outcomes can affect not only the quantity of assessments conducted in the context of disability determinations but also the quality of the assessments that are conducted and the resulting information (IOM, 2015; NASEM, 2016). Disability applicants who are uninsured or underinsured are less likely to have a well-developed body of health data, including the results of expensive, specialized tests, to demonstrate evidence of disability.

In addition, the acquisition of clinical records may be difficult for several reasons: providers' fear of sharing confidential information, the limited capacity of a provider's organization to gather and transmit records, and high administrative costs for record transfer. If clinical records are in electronic format, they can be shared and transmitted more easily, and digital applications can enhance the types of information available, as described later. However, most EHRs are supported by commercial vendors with differing digital systems, and a lack of interoperability may make reading and manipulating such records complex and costly. Still another difficulty is that information and representations concerning function and disability are often not standardized, although some systems, such as the publicly available Patient-Reported Outcomes Measurement Information System (PROMIS), are progressing toward standard taxonomy-driven published measures (HealthMeasures, 2018). In addition, EHRs typically contain a substantial amount of free text, which may impede easy analysis and summarization, although natural language processing software is available for this purpose. Finally, use of EHRs makes it easy for providers to "cut and paste" previous entries, potentially affecting accuracy.

Performance-Based Measures

Performance-based measures require that the individual being assessed perform a set of functional tasks so that his or her ability to execute them can be ascertained. Examples of such measures include assessments of gait, balance, and lifting in the physical realm and cognition in the mental realm. In addition, a number of instruments are available for integrated assessment of physical and mental function, often entailing evaluation of specific activities of daily living (ADLs) and instrumental activities of daily living (IADLs). ADL and IADL measures can be applied reasonably well but have the problem of not being directly applicable to specific workplace settings and activities. Potential weaknesses of integrated and impairment-specific functional assessments also include social interaction tasks required in certain workplaces and the speed and endurance of various work tasks. In general, performance-based measurement requires formal and substantial control of assessment conditions and the information as it is collected. When the intent is to assess function based on maximal effort, observer-reported (as opposed to self-reported) function may itself affect the information collected.

There are two general approaches to performance-based functional assessment: direct measurement and in situ observation of task/activity performance.

Direct Direct measurement involves testing of relevant, perhaps general or stereotyped tasks in a clinical "laboratory" setting, often in rehabilitation or occupational medicine facilities. Direct performance testing for physical and neurocognitive functional abilities is well developed for various common illnesses and conditions and defined injury-related impairments. Such testing typically is used to assess common disease-specific deficits and to monitor functional increments or decrements over time. Such measures may be useful for tracking the progress of those diseases, but they are not necessarily generalizable to other potentially disabling conditions.

Work-related tasks are tested using best approximations of the work settings being simulated. Physiological measurements are often performed, and the data collected may require some translation by the observer as to whether the client is likely to be capable of carrying out the work tasks under real-life conditions. These testing procedures can be valuable but may require substantial resources, including client transport to and from the testing facility.

In situ observation of task/activity A second approach is to test performance abilities by observing the client in a work setting that is identical, or nearly so, to that in which the tasks/activities in question would be carried out. In situ observation has some advantages over direct measurement. Assuming that the client being evaluated can be taught the procedural requisites and safety procedures related to performing in that particular workplace, this approach has the distinct advantage of representing both many of the challenges of the tasks/activities and the availability of potential adaptations to accommodate the client's condition. Nevertheless, there are disadvantages to in situ observation as well. It is more logistically complex, time and resource intensive, and costly than direct measurement. Additional challenges include simulating social and cognitive tasks; assessing the client's endurance in performing these tasks/activities in the particular work setting over a 40-hour workweek; and anticipating variation in the tasks/activities as they evolve over time, as well as variations and fluctuations in the individual's symptoms over time. It should be noted that job coaches in the context of supported employment can provide useful information about their clients' performance in the specific job setting, which takes account of some these challenges.

Self- or Proxy-Reported Measures

Self- or proxy-reported measures are those that require an individual being assessed or a third party to complete a questionnaire asking about the overall ability of the individual to perform a specific set of functional tasks. Patient-reported outcome measures, ADL questionnaires, and some types of psychological tests are examples of such measures. Note that self-report is not the same as self-administration, and different modes of administration (e.g., mail, computerized surveys, "pencil and paper," automated audio telephone surveys) may alter the nature of self-reported findings. An advantage of self-administered instruments is that they can be completed over longer periods of time relative to the other two methods, providing a time perspective on function and potentially improving accuracy by allowing for more response verification.

Self- and proxy-reported instruments may be based on either classical test theory methods or item response theory (IRT) methods.

Instruments based on classical test theory methods Instruments based on classical test theory are designed to be completed in toto via either self-report, proxy (surrogate) report, or interview. Self-report is feasible when an instrument's reading level accommodates respondents with low literacy and limited concentration. Proxy respondents are helpful when instruments focus on observable behaviors. Personal interview administration is helpful when respondents need assistance staying on task or would benefit from the rapport that can be established during an interview. If the primary respondent is cognitively impaired, for example, assistance may improve the accuracy of the information obtained.

Health, functional, and disability assessments may be enhanced by acquiring information from third-party respondents. Relevant third-party respondents may include, for example, friends and family members, health care and social professionals, and workplace colleagues and employers. Such individuals can be particularly helpful for providing ancillary information on health and behavioral matters, physical and mental functioning, and workplace performance, sometimes supported by written documents. Applicants who are injured or ill can benefit in particular from appropriate third-party observers. Yet, while such reports can provide valuable information, it is important to understand the nature of respondents' relationship to the individual being evaluated. They may not be skilled in the types of observations needed, may not be suitably familiar with the person, or may have biases or interests of their own that could affect the accuracy of information they provide (Gill et al., 2002; Lum et al., 2005). Collection of information about the length and nature of a third-party respondent's relationship with the individual being assessed can help in interpreting the information gathered. It is important to note that tests assessing beliefs, attitudes, moods, and other internal states are not suitable for proxy respondents (Dorman et al., 1997; Duncan et al., 2002; Mathias et al., 1997; Oczkowski and O'Donnell, 2010; Pickard and Knight, 2005; Poulin and Desrosiers, 2008).

In general, personally administered interviews will provide greater accuracy and potentially obtain more complex responses relative to self- and proxy reports obtained by other means. Reading a survey to respondents with language or literacy problems will likely improve the quality of the information provided. Interviewers can also determine the most acceptable pace of data collection, ensure completion of all desired items, sometimes elicit more sensitive information, and explain items otherwise not fully understood. At the same time, however, administered interviews consume more resources, and administration of a survey of self-reported function by an interviewer may actually reduce the accuracy of the information collected. Also of note, both self- and interviewer-administered survey instruments (as well as computer-administered instruments, discussed below) may allow for "adaptive interviews," where, for example, certain items may be omitted or added if they are redundant or require respondent-specific information.

Whether surveys are self- or interviewer administered, the accuracy of survey information is based mainly on respondent characteristics, including abilities, knowledge, motivations, and competing burdens. The accuracy of self-reported information can be affected, intentionally or unintentionally, by the respondent. For example, some individuals who want their condition or the magnitude of their perceived distress to be taken seriously may overestimate their difficulty in performing various tasks. Conversely, other individuals may overestimate their abilities out of a desire to please the interviewer or to maintain independence or not appear weak. In addition, certain individuals, for example, some with traumatic brain injury or stroke, may have poor self-awareness or an inability to assess their limitations accurately because of a neurological deficit (e.g., anosognosia). The use of instruments or test batteries that include validity measures[3] can help testers determine the validity of the results obtained (IOM, 2015). Another consideration is gender, racial, ethnic, and cultural variation in individuals' perception of illness and symptoms and whether relevant self-report measures have been assessed for equivalency of scores in different populations.

Instruments based on item response theory methods Measurement tools may be built on IRT and computer adaptive testing (CAT). IRT is used to establish an individual's position on the continuum of a trait of interest by asking him or her a series of questions. In the case of functional assessment, for example, the questions would be calibrated to a scale covering the range of function in one dimension, such as mobility. CAT instruments can be used to administer a selected sample of questions from an IRT-calibrated "item bank," choosing questions based on how the respondent answered the previous questions (Chan, 2018). CAT instruments are highly efficient, typically involving shorter administration times and requiring respondents to answer fewer questions than would be required by a conventional test (Cheville et al., 2012; Fliege et al., 2009; Ware et al., 2003). Such instruments also may include embedded validity measures. Under optimal

[3]Validity measures are used to provide information about an individual's effort on tests of maximal performance, such as cognitive tests, or information about the consistency and accuracy of an individual's self-report of symptoms he or she is experiencing (IOM, 2015).

circumstances, computer-based surveys have several additional advantages: they can provide helpful prompts if a respondent indicates a question is unclear or provides a response that is "out of range"; questions can be revisited on request; and real-time response editing can indicate that a particular response cannot be accurate and needs to be reconsidered.

CAT has generally shown its enhanced value in a number of applications relevant to disability situations. For example, CAT has increased the precision of discriminating persons who fall more frequently relative to conventional balance measures (Pardasaney et al., 2014), and similarly discriminating those with disabling back pain compared with standard testing (Choi, 2015). Use of a CAT approach to evaluating PROMIS-based upper-extremity function showed better psychometric properties compared with a non-CAT comparator instrument (Tyser et al., 2014), as well as improved sensitivity to functional change after hip and knee prosthetic procedures (McDonough et al., 2016). The use of measures developed using IRT that can be administered using CAT can also decrease respondent burden by reducing survey length and administration time while minimizing measurement error.

In addition, computerized surveys can provide immediate document structuring and formatting of the interview, supplying text versions of responses for immediate examination by professionals. A further potential advantage of electronically administered interviews is that the perceived privacy may allow the acquisition of more sensitive information than might otherwise be obtained. It should also be noted that certain types of cognitive and psychological testing protocols can be administered online.

Regardless of the sources and types of information, convergence of the information is important in weighing the validity of the evidence presented in a claim of disability. Divergent evidence erodes confidence, whereas convergence adds confidence that the reported information is accurate.

Potential Digital Applications of Clinical Records and Related Information

One important consequence of the use of digital, computerized clinical records and other digitized information is the ability to take advantage of techniques increasingly being applied in clinical settings. Many of these techniques could be considered decision-support tools, and while each technique that exploits big data (i.e., extremely large datasets) and predictive analytics[4] (Shah et al., 2018) may have weaknesses or suffer from incom-

[4] "Predictive analytics" is "the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data" (SAS Institute, 2019b).

plete development, these techniques have the potential to become useful aids in the collection of information to inform disability determinations. For example, techniques such as text and data mining, machine learning, and natural language processing—some of which fall under the heading of clinical information extraction applications (Wang et al., 2018)—can help in summarizing the content, quality, and completeness of collected information. The process of reading, organizing, and interpreting clinical information has improved with the use of machine reading (artificial intelligence) programs that have emerged in the past several years.

The modeling of predictive analytics can offer reasonably accurate estimates of clinical outcomes for various conditions over the ensuing months and years. As of 2012, for example, there were about 800 predictive risk models for outcomes of cardiovascular conditions (Wessler et al., 2015). In another important application, the U.S. Veterans Health Administration has developed a detailed risk score (the Care Assessment of Need Score), based on medical records from hospitalized patients, that predicts 1-year mortality and the risk of rehospitalization (Ruiz et al., 2018). With appropriate available information, it may be useful to evaluate these information technology techniques and others as aids in predicting future health and functional trajectories and outcomes of individuals applying for disability benefits. Predictive modeling also has progressed with respect to the functional outcomes of mental conditions (Koutsouleris et al., 2018). An example of an emerging digital technology in this realm is the use of linguistic analysis (of recorded natural speech) to help assess neurological and psychiatric conditions (deBoer et al., 2018).

PROPERTIES OF MEASUREMENT INSTRUMENTS

When evaluating the utility of a functional assessment instrument for SSA's adjudication process, it is important to collect and consider the available evidence on that instrument's design and performance. To guide potential users of the PROMIS and National Institutes of Health (NIH) Toolbox instruments, NIH (2018) suggests evaluating the available evidence related to the eight key instrument properties described by the Scientific Advisory Committee of the Medical Outcomes Trust (2002). The present committee adapted these properties for use in selecting the functional assessment instruments considered in this report:

- conceptual model and measurement approach,
- reliability,
- validity,
- sensitivity to change and responsiveness,

- interpretability of results (e.g., self-report and trained observer rating),
- administrative and respondent burden,
- alternative modes of administration, and
- cultural and language adaptations (e.g., translations).

Conceptual Model and Measurement Approach

The *conceptual model* on which an instrument is based provides "the rationale for and description of the concepts" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 198) the instrument is intended to measure and the "populations [it] is intended to assess" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 198). The underlying *measurement approach* operationalizes the conceptual model and involves a range of decisions made during the instrument's design and testing, such as the use of a scale or set of scales, the corresponding measurement units, modes of collection, scoring procedures, and empirical strategies (e.g., principal components analysis, confirmatory factor analysis). Brandt and colleagues (2011), for example, describe (1) SSA's five-step sequential review process in the context of contemporary conceptualizations of disability and (2) the potential of using an IRT-CAT assessment tool for measuring the multiple dimensions of disability—both of which would guide the development of the Work Disability-Functional Assessment Battery (Meterko et al., 2015, 2018). In the context of functional assessment, an underlying conceptual model is one piece of evidence supporting the overall credibility of the measurement results and subsequent decisions.

Reliability

Reliability denotes "the degree to which an instrument is free from random error" at a point in time (Joint Commission on Accreditation of Healthcare Organizations, 2003, p. 29). Several types of reliability are typically investigated in developing an instrument: *internal consistency, interrater reliability,* and *test-retest reliability*.

Internal consistency reliability is "the precision of a scale, based on the homogeneity (intercorrelations) of the scale's items" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 196). For example, the developers of the Kessler Psychological Distress Scale (K10)—a self-report questionnaire with 10 brief items used to identify psychological distress—established internal consistency reliability using the responses of individuals to whom the scale was administered to estimate Cronbach's alpha, which indicates the correlation among items (Kessler et al., 2002).

Greater correlation suggests that the items are identifying similar phenomena, which may be consistent with the assumed conceptual model. In the context of functional assessment, internal consistency reliability strengthens the credibility of results by minimizing the potential for collecting contradictory data.

Interrater and test-retest reliability are, respectively, the degree to which different raters are consistent in their observations and scoring at one point in time and the degree to which the results of a test are consistent over time. Wittchen and colleagues (1991), for example, compared the diagnoses of the same sample of patients made by clinicians and nonclinicians using the Composite International Diagnostic Interview, which is designed to identify mental disorders, and estimated Cohen's kappa, which indicates the extent of agreement between the two types of raters. In the context of functional assessment, it is critically important to establish interrater and test-retest reliability, since it is important to obtain the same score when an instrument is used by different clinicians or when the same claimant is retested, all else being held constant.

Validity

The *validity* of an instrument is "the degree to which the instrument measures what it purports to measure" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 200). Three forms of validity are typically investigated when an instrument is developed. The first, *content validity*, is the degree to which "the domain of an instrument [i.e., what it purports to measure] is appropriate relative to its intended use" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 200). *Construct validity* is the degree to which the "proposed interpretation of scores [is] based on theoretical implications associated with the constructs being measured" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 200). And *criterion validity* is the "extent to which scores of the instrument are related to a criterion measure" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 200). Criterion measures are "measures of the target construct that are widely accepted as scaled, valid measures of that construct" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 200). Kessler and colleagues (2002), for example, compared results from the K10 self-report scale with the results of clinical assessments, the latter of which are assumed to be closer to the true level of psychological distress. Establishing the criterion validity of functional assessments is critical to their use in disability claims adjudication and to the underlying accuracy and defensibility of subsequent claims decisions.

Sensitivity to Change and Responsiveness

Although conceptually similar and sometimes used interchangeably, sensitivity to change and responsiveness have been defined differently (Corzillius et al., 1999; Pardasaney et al., 2012). *Sensitivity to change* is defined as the ability of an instrument to detect change in a state over time reliably, regardless of whether the change is meaningful, whereas *responsiveness* is defined as the ability of an instrument to detect a meaningful or clinically relevant change over time that is reproducible against an alternative measure or criterion (Corzillius et al., 1999; Pardasaney et al., 2012). For example, an assessment instrument might detect a significant increase in an individual's leg muscle strength (sensitivity), but the changes might not be detectable using an alternative measure or criterion, such as the ability to stand unassisted. In the context of functional assessment, the ability to detect clinically relevant changes over time is relevant to assessing functional decline and recovery.

Interpretability of Results

Interpretability refers to "the degree to which one can assign easily understood meaning to an instrument's quantitative scores" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 202). An instrument's interpretability "is facilitated by information that translates a quantitative score or change in scores to a qualitative category or other external measure that has a more familiar meaning" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 202). Different types of information contribute to the interpretation of scores, including their relationship to clinical conditions or significant life events (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 202). With regard to disability determination, interpreting scores in the context of job requirements is crucially important.

Administrative and Respondent Burden

The burden involved in using an instrument is considered from the perspective of the person administering the instrument (*administrative burden*) and the person responding to the instrument (*respondent burden*). Burden encompasses "time, effort, and other demands" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 202), including the "amount of training and level of education or professional expertise and experience needed" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 203) by the person administering the instrument. Respondents with cognitive disabilities may experience additional burden resulting from the cognitive load required to complete an instrument. Respondents with fine motor control limitations may encounter difficulties with reporting formats that require fine motor control of fingers. Respondents who are distractible and have difficulty remaining on task may require reminders or redirection to attend to test materials.

Alternative Modes of Administration

The burden of an instrument is related to the available *alternative modes of administration*, which include "self-report, interviewer-administered, trained observer rating, computer-assisted interviewer-administered, [and] performance-based measures" (Scientific Advisory Committee of the Medical Outcomes Trust, 2002, p. 203). In the context of consultative examinations, understanding the administrative aspects of an assessment instrument is important to the cost and timeliness of an assessment.

Cultural and Language Adaptations

Lastly, an instrument may be adapted or translated for use with populations that differ culturally and linguistically from those for which the instrument was initially developed. The *cultural and language adaptations* or translations of an instrument involve two primary steps: (1) assessment of conceptual and linguistic equivalence, and (2) reevaluation of the seven priorities described above. Üstün and colleagues (2010), for example, compared the results of psychometric testing of the World Health Organization (WHO) Disability Assessment Schedule 2.0 across multiple countries. In the context of disability claims adjudication, it is important that any cultural and linguistic barriers experienced by the applicant be noted and that inferences from assessment results be appropriately qualified.

PROFESSIONALS WITH THE TRAINING AND EXPERTISE TO ASSESS FUNCTION

An important consideration when evaluating the validity, reliability, and usefulness of an instrument is the educational, professional experience, and training requirements for those who can administer the instrument and, by extension, for those who can interpret the instrument's results. When an instrument is developed, validated, and/or normed on a population, such requirements typically are specified or at least recommended. The developers may also specify requirements for additional training before individuals are certified as qualified to administer the test and/or interpret the results.

The types of professionals who are qualified to administer instruments and/or interpret their results for purposes of functional assessment of physical and mental abilities are as varied as the conditions they represent, the multiple disciplines of the medical and allied health workforce, and the current state of the scientific literature with respect to the assessment target (e.g., major depression, ischemic heart disease, fibromyalgia, traumatic brain injury). To illustrate, a trained laboratory technician is qualified to perform neuroimaging tests using sophisticated diagnostic equipment, but is not qualified to interpret, diagnose, and report the results of those tests. Conversely, in the case of real-world applications of screening and assessment in which the goal is to identify treatments, interventions, and practices, tools may have been adapted for administration and scoring by a more diverse workforce. An example is depression screening tools (e.g., Patient Health Questionaire-9), which are widely available and come with easy-to-follow instructions on how to administer and score them and interpret their results. Information on requirements for persons who are qualified to administer and/or interpret the results of selected instruments for assessing physical and mental abilities relevant to work is provided in Chapters 5 and 6 (and associated annex tables), respectively.

Of particular note is that in the last decade or so, community health workers have been assuming responsibility for administering and/or interpreting the results of instruments that previously were considered largely the purview of more highly trained or certified specialists or assessors. These community health workers fill an important gap, especially for underresourced communities and service programs experiencing workforce shortages, freeing up skilled health care providers to perform more complex health care tasks. They typically are recruited because of their unique knowledge of and ability to navigate patient, family, and community expectations and norms around health, functional well-being, and access to care (Crigler et al., 2011; Hartzler et al., 2018).

Another consideration pertaining to the identification of professionals with the training and expertise to perform functional assessments related to work requirements is the importance of the balance between the snapshot of an examinee's performance provided by a particular instrument at a single point in time and the understanding gained from repeated assessments or observations over time by professionals who have frequent interactions with patients by nature of their role and responsibilities in a clinical or rehabilitative setting or other system of care. Licensed clinical social workers, occupational therapists, physical therapists, and other professionals may administer ongoing assessments in their respective roles on a multidisciplinary team. They may have responsibility for repeated assessments using standardized assessment tools and procedures, and thus may render more detailed and accurate evaluations of an individual's physical and/or mental functioning over time than can be provided by medical specialists who have less frequent interactions with the person and less time per encounter during the same observation period.

Also important is identifying which professionals may be best suited to evaluating an applicant for disability benefits. For example, physicians who are skilled, trained, and experienced in determining impairments, such as occupational medicine physicians and rehabilitation medicine physicians (physiatrists), may be best qualified to perform these evaluations by virtue of their training and expertise in understanding not only the physical aspects of impairment, but also the work environment and how the abilities of impaired individuals may match that environment. Accordingly, they may provide information that is most relevant and useful to the disability determination process. Other clinicians and health care providers, such as physician assistants and nurse practitioners, who have followed the person being assessed for an extended period may be best suited to providing information on the individual's functional abilities over time, regardless of whether they have expertise in evaluating the person's specific impairment. Other clinicians with experience in evaluating and treating impairments—such as occupational therapists; speech-language pathologists; and physical therapists—also may be well qualified to conduct these evaluations.

Beyond identifying those professionals with appropriate expertise to perform functional assessments, acquiring information helpful in determining individuals' functional abilities relevant to work is facilitated by asking clear and specific questions that target the information of greatest use in making a disability determination. To this end, forms with such questions can be provided to relevant professionals. For example, asking one item per question imposes less cognitive load on the professional, while asking as many discrete questions as possible rather than open-ended questions yields specific responses that may be more useful to the adjudicator than the responses to open-ended questions, which may be more or less useful depending on how the professional responds.

One useful approach is to front-load the short-answer—specific and straightforward—questions, such as: How long have you known the individual? What is/are the diagnoses for which you are seeing the person? What are the individual's main symptoms? and Approximately how long has the individual had these symptoms? An expert in questionnaire development and/or psychometrics may be helpful in formulating questions that will be least ambiguous for the responding professionals and yield responses that the adjuster will find most helpful. This may be important given that providers often have limited time to evaluate an individual, allowing for a more efficient process for both the provider and the adjudicator.

Also helpful would be providing the clinician with information that is as specific and detailed as possible about why the applicant is seeking disability benefits, as well as the type of information that would be useful to the adjudicator for making the determination. Providing as much guidance as possible to the clinician would contribute to a more efficient evaluation. A model for consideration is the recent Accreditation Council for Graduate Medical Education requirement that residents have a formal process for transferring information used in the care of a patient, known as "hand-off," because it has been shown that discontinuity creates opportunities for error and miscommunication (PSNet, 2018; Riebschleger and Philibert, 2011).

POTENTIAL THREATS TO THE VALIDITY OF FUNCTIONAL ASSESSMENT

SSA is particularly interested in information that pertains to individuals' capacity to sustain physical and mental work activities on an ongoing and independent basis in the face of functional limitations. However, information in applicants' health records typically is gathered for other purposes (e.g., treatment, rehabilitation) and so does not speak unambiguously to an individual's capacity to sustain work-related physical or mental activities for 8 hours per day, 5 days per week. The information obtained is often influenced by the purpose for which it was originally gathered, which makes it difficult to draw inferences from that information for a different purpose. Assessments of work disability often require an inferential leap because they are based in part on information gathered for other purposes. The committee identified six primary threats to the validity of assessments of functional abilities: (1) testing maximal versus typical performance, (2) assessment of episodic activity versus sustained task performance, (3) absence of standardized testing conditions, (4) mixed-motive incentives, (5) compromised test integrity in high-stakes testing, and (6) diversity in the test population.

Testing Maximal Versus Typical Performance

In some cases, functional assessments are performed under conditions that best resemble maximal rather than typical performance, which by definition implies continuous and independent performance. For instance, the controlled settings in which physical and cognitive activities often are assessed typically fail to replicate the actual conditions under which such activities are performed at work. Specifically, such variables as social pressure (e.g., irate customers), hostile environmental conditions (e.g., temperature, humidity, noise), and continuous repetition over extended time periods and in a variety of settings are not always well replicated in the settings in which assessments of functional capacity are conducted.

Presentations to the committee by several stakeholder representatives raise similar concerns (Ford et al., 2018; Liebkemann, 2016; Liebkemann and Lang, 2017). Specifically, some functional assessments conducted in controlled conditions may overestimate an individual's capacity to perform work activities independently on a sustained basis. Certain medical conditions (e.g., cardiac, mental health) may interact with the context in which these assessments are carried out, so that a person is shown to be capable of performing work activities in a single episode in the presumably less adverse conditions of a controlled setting but proves unable to perform those activities in an ongoing and independent manner as demanded in an actual work setting. Research on typical versus maximal performance suggests that the antecedents of maximal performance on certain tasks do not always coincide with those of typical performance on the same tasks (DuBois et al., 1993; Salgado et al., 2015). Therefore, the individual qualities that facilitate success on assessments conducted under controlled conditions may not always ensure success in sustained performance in contexts subject to constant change. In the physical realm, the difference between peak capacity and sustained performance can be quantified specifically for aerobic functional capacity by means of cardiopulmonary exercise testing, but similar assessments of function are currently not available for all organ systems.

Assessment of Episodic Activity Versus Sustained Task Performance

Clearly, sustained and independent performance of a job involves more than the ability to perform each of the work activities separately. In other words, a job is more than the sum of its tasks; rather, it entails a series of coordination and integration processes involving the frequency, sequence, and duration of those tasks. We define meta-task processes as those concerning the coordination, integration, and sequencing of tasks because such processes require a clear understanding of task interdependence and criticality or the consequences of an error in task performance.

Unfortunately, meta-task processes are not necessarily evident in assessments of a single work activity performed in a controlled setting. In such cases, it may be necessary to make a difficult inferential leap from evidence of the individual's ability to perform some of the work activities involved in a job in isolation, and often in ecologically unrealistic contexts that do not accurately represent actual working conditions, to the individual's ability to sustain the full range of job activities and meta-task processes demanded in those conditions. Therefore, assessments ideally would be representative and encompass both the full range of tasks involved in the job and the ability to meta-task as necessary to perform the job. Consider, for example, the case of someone who performs successfully in a variety of episodic assessments targeting several activities or even broader tasks of a job. This same individual might prove unable to engage in continued performance of those activities or tasks because of his or her inability to decide on their sequencing as the result of a health condition (e.g., mental illness) or the inability to sustain required physical activities over the necessary period of time. In other words, an activity-by-activity assessment of capacity, even if it covers the entire scope of activities and tasks involved in a job, may not provide a valid prediction of the individual's ability to sustain performance of the job over time in an independent manner.

Absence of Standardized Testing Conditions

Standardization is a basic precept of valid and reliable assessment. Assessments are sometimes administered under varied conditions, such as test administrators who go out of their way to encourage applicants or who allow examinees to rest between tasks. The results of such assessments would not be comparable to those performed in the absence of encouragement or recovery periods. Variations in testing conditions are likely to be most problematic when nonstandardized instruments or methods (e.g., clinical judgment based on a potentially unrepresentative sample of behavior) are employed.

In some cases, customized assessments focused on the performance of activities unique to a certain job type (e.g., dexterity for a wet-bench lab worker where pipettes are used) may be appropriate. Such customized assessments are less generalizable and not as rigorous as standardized ones, but they allow assessments with a higher degree of "fidelity" to the job and, therefore, possess better face validity (i.e., more likely to be perceived as relevant by applicants) (Salgado et al., 2015).

Understanding the conditions under which assessments are administered is therefore of critical importance in interpreting assessment results. For this reason, it is important to gather as much information as possible on the standardized conditions under which assessments were conducted, as well as a detailed description of any exceptions to or deviations from these conditions.

Mixed-Motive Incentives

Also important is understanding the potential influence of test administrators and other third parties operating under mixed-motive incentives, in which those conducting assessments operate under conflicting external pressures that motivate them to both adjudicate and not adjudicate disability benefits. For instance, test administrators, whether consciously or not, may at times conduct assessments in a manner that provides the examinee with much greater encouragement to perform work activities than is typically encountered in an actual work setting, thereby rendering results unrepresentative of the actual conditions under which the work activities being assessed need to be performed. Because of their background, for example, this may be the case for rehabilitation specialists and other professionals engaged in therapeutic interventions. By virtue of their training, these professionals are interested in helping their patients succeed, and so might display an unduly optimistic and cheerful demeanor during the assessment. Individuals being assessed by such professionals may be able to complete certain work activities in these unusually motivating circumstances, but be unable to sustain those same activities on an ongoing basis and in the independent manner demanded by the job. The literature suggests that low-ability individuals may be particularly susceptible to this type of motivational stimulation and benefit more from motivational interventions (e.g., goal setting) relative to high-ability individuals (Kanfer and Ackerman, 1989). Thus, it is not unreasonable to think that low-ability individuals assessed under unusually motivating circumstances (e.g., a cheerful therapist) would show better results than they would be capable of achieving in a less supportive or encouraging environment.

Whenever possible, then, it is advisable to declare and formally document the mixed motives of test administrators so that the extent to which they may have motivated the examinee can be assessed, regardless of whether they intended to alter the examination conditions. Potential ways to address this issue include enacting standardized protocols and interpretive guidelines that regulate or codify the conditions of assessments and the manner in which they should be conducted by third parties, as well as communicating the goals and purpose of the assessments to those charged with administering them.

It is also important to understand the purpose for which an assessment was conducted in interpreting its results. As previously mentioned, test results gathered for one purpose (e.g., therapeutic or rehabilitation), where the objective may be to determine the maximum extent to which an individual can perform an activity even if he or she cannot sustain that activity over time, may not provide accurate information for a different purpose (e.g., assessment of work disability), where the objective is to learn the maximum extent to which the person can sustain that activity safely on a regular and continuing basis. It is likely, for instance, that some of the physical functional assessments documented in an individual's health records were conducted by occupational and physical therapists in the context of interventions aimed at helping or motivating the person to perform at maximum capacity, rather than determining his or her capacity to sustain work activities independently over time. For this reason, it is important to understand the purpose of the original assessment when using its results for a different purpose.

Compromised Test Integrity in High-Stakes Testing

A high-stakes test is one whose results constitute the basis of a major decision, typically one involving the individual taking the test. Given the significant consequences of the adjudication of disability benefits, assessments conducted in the context of disability determination are a case of high-stakes testing. In light of the personal and social significance of adjudication decisions, the motivation to skew an assessment in the desired direction is rather high. Therefore, the use of assessment instruments developed for relatively low-stakes applications, such as research and teaching purposes—often available to a wide range of professionals for many years—is problematic. In addition, these instruments may be administered in less than fully standardized conditions and through a variety of platforms that are not always secure. In some cases, use of these instruments in less than secure conditions and over extended periods of time may result in their becoming publicly available, which is likely to compromise their integrity (AERA et al., 2014). If the content of a test became widely available publicly (e.g., on the Internet), prospective examinees could potentially preview the test questions and prepare accordingly, undermining the validity of the results. Therefore, it is important to know the integrity of assessment instruments used to inform disability determinations to the extent possible.

Diversity in the Test Population

The literature has long highlighted the role of race, ethnicity, and culture in assessments of physical and mental abilities, and to this day is punctuated by extensive debates over the validity and cultural relevance of procedures, tests, and assessments across racial, ethnic, and cultural groups, as well as groups identified by age and gender (Alonso et al., 2013; Baird et al., 2007; CNPAAEMI, 2016; Wild et al., 2005; Wong et al., 2000).

Indeed, patient-reported symptomatology measures and clinician/observer-rendered assessments vary in the degree to which they have been tested or adapted across diverse racial, ethnic, and cultural populations. Cross-cultural adaptations and validations of assessments in different cultural contexts and languages are predicated on the notion that such efforts take into account distinct groups' experiences and meanings of health, behaviors, illness, symptomatology, and disability and help-seeking behaviors (Forestier et al., 2019; Fuentes and Aranda, 2018; Odole et al., 2016; Tennant et al., 2004; Zdunek et al., 2015).

Moreover, assessment instruments developed in research and training settings may not account for cultural, linguistic, or literacy factors, such as limited English proficiency or low literacy, that may limit access to such assessments. As a result, few or no assessments may be available that can capture valid and reliable administration and scoring information for these populations. In sum, when evaluating the utility of a functional assessment instrument for informing disability determinations, it is important to consider the instrument's performance across multiple subgroups (e.g., age, gender, socioeconomic status, race, ethnicity, cultural group) as a principle of good practice (Wild et al., 2005).

FINDINGS AND CONCLUSIONS

Findings

3-1. A variety of methods can be used to collect functional information (e.g., diagnostic testing, performance-based measures, self- or proxy-report measures), each of which has strengths and weaknesses, and the results of one are often used to validate those of another. Each method can yield instruments with satisfactory psychometric properties that allow their implementation in disability decision making.

3-2. It is important to consider eight properties when evaluating the quality of functional assessment instruments:
- conceptual model and measurement approach,
- reliability,
- validity,
- sensitivity to change and responsiveness,
- interpretability of results (e.g., self-report and trained observer rating),
- administrative and respondent burden,
- alternative modes of administration, and
- cultural and language adaptations (e.g., translations).

3-3. The validity of functional assessment tests is enhanced when the test users administer them for the purpose and in the context for which they were designed (e.g., target population).

3-4. Assessment instruments developed for use in research and training settings may not account for cultural, linguistic, or literacy factors, such as limited English proficiency or low literacy, that can limit access to such assessments.

3-5. Direct performance testing of physical and neurocognitive functional abilities is well developed and typically is used to assess common disease-specific deficits and monitor functional increments or decrements over time. Such testing may be useful for tracking the progress of those diseases, but they are not necessarily generalizable to other disabling conditions.

3-6. The accuracy of self-reported information can be affected, intentionally or unintentionally, by the respondent, who may either under- or overestimate his or her ability to perform different tasks.

3-7. The use of instruments or test batteries that include validity measures can help testers determine the validity of the results obtained.

3-8. Third-party sources (e.g., friends and family members, health care and social service professionals, workplace colleagues and employers) who are suitably familiar with the applicant's activities, health, and functional status can be particularly helpful in providing ancillary information on health and behavioral matters, physical and mental functioning, and workplace performance, sometimes supported by written documents. Such reports are at times influenced by such factors as self-interest, mixed motives, or inaccurate observations. Tests assessing beliefs, attitudes, moods, and other internal states are not suitable for proxy respondents.

3-9. Threats to the validity of assessments of functional abilities include testing of maximal versus typical performance, assessment of episodic activity versus sustained task performance, absence of standardized testing conditions, mixed-motive incentives, compromised test integrity owing to prior use of the test in low-stakes testing applications, and diverse test populations in whom tests may not have been validated.

3-10. Information obtained as evidence is often influenced by the purpose for which it was originally gathered (e.g., treatment, rehabilitation), which makes it difficult to draw inferences from that information for a different purpose (e.g., determination of work disability).

3-11. A variety of professionals representing multiple disciplines of the medical and allied health workforce are qualified to administer and interpret results of assessments of physical and mental function and have the capacity and experience to provide valuable information regarding individuals' functional abilities.

3-12. Community health workers have assumed responsibilities for administration and/or interpretation of instruments that previously were typically considered the purview of more highly trained or certified specialists or assessors. In so doing they have filled an important gap, especially for underresourced communities and service programs experiencing workforce shortages.

3-13. Health care data relevant to disability determinations, such as the results of specific, expensive tests (e.g., certain cardiovascular tests and psychological batteries) that are valid and potentially useful, may not be easily available because an individual may lack insurance coverage or be underinsured, or the means of administering the tests may be denied by insurance because the tests are not considered medically necessary.

3-14. Lower socioeconomic status is associated with less access to high-quality care and health care professionals, including those with expertise in providing information relevant to disability determination.

3-15. Patient-reported symptom measures and clinician/observer-rendered assessments vary in the degree to which they have been tested or adapted across diverse racial, ethnic, and cultural populations.

Conclusions

3-1. The use of measures based on item response theory that can be administered using computer adaptive testing can decrease respondent burden by reducing survey length and administration time while minimizing measurement error.

3-2. Professionals with responsibility for repeated assessments using standardized assessment tools and procedures may render more detailed and accurate evaluations of an individual's physical and/or mental functioning over time relative to medical specialists who have less frequent interactions with the person and less time per encounter during the same observation period.

3-3. It is important to understand the nature of a proxy informant's relationship to the individual being assessed because a proxy may not always be suitably familiar with the person or may have biases or interests that may affect the accuracy of information provided. Collecting information about the length and nature of a proxy respondent's relationship with the individual can help in interpreting the information gathered.

3-4. It is important to collect information about the nature and original purpose of an assessment instrument, as well as the conditions and context in which it was administered, to help in understanding the results with respect to potential limitations on their generalizability.

3-5. When evaluating the utility of a functional assessment instrument for informing disability determinations, it is important to consider the instrument's performance across multiple subgroups (e.g., age, gender, socioeconomic status, race, ethnicity, cultural group) as a principle of good practice.

3-6. Disparities in access to care and health outcomes can affect not only the quantity of assessments conducted in the context of disability determinations but also the quality of the assessments that are conducted and the resulting information.

REFERENCES

AERA, APA, and NCME (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education). 2014. *Standards for educational and psychological testing*. Washington, DC: AERA, APA, and NCME.

AHRQ (Agency for Healthcare Research and Quality). 2018. *2017 National healthcare quality and disparities report*. Rockville, MD: AHRQ. http://www.ahrq.gov/research/findings/nhqrdr/nhqdr17/index.html (accessed April 3, 2019).

Alonso, J., S. J. Bartlett, M. Rose, N. K. Aaronson, J. E. Chaplin, F. Efficace, A. Leplège, A. Lu, D. S. Tulsky, H. Raat, U. Ravens-Sieberer, D. Revicki, C. B. Terwee, J. M. Valderas, D. Cella, C. B. Forrest, and the PROMIS International Group. 2013. The case for an international Patient-Reported Outcomes Measurement Information System (PROMIS®) initiative. *Health and Quality of Life Outcomes* 11:210.

Anglin, D. M., P. M. Alberti, B. G. Link, and J. C. Phelan. 2008. Racial differences in beliefs about the effectiveness and necessity of mental health treatment. *American Journal of Community Psychology* 42(1–2):17–24. doi: 10.1007/s10464-008-9189-5.

Baird, A. D., M. Ford, and K. Podell. 2007. Ethnic differences in functional and neuropsychological test performance in older adults. *Archives of Clinical Neuropsychology* 22(3):309–318.

Brandt, D. E., A. J. Houtenville, M. T. Huynh, L. Chan, and E. K. Rasch. 2011. Connecting contemporary paradigms to the Social Security Administration's disability evaluation process. *Journal of Disability Policy Studies* 22(2):116–128.

Carpenter-Song, E., E. Chu, R. E. Drake, M. Ritsema, B. Smith, and H. Alverson. 2010. Ethno-cultural variations in the experience and meaning of mental illness and treatment: Implications for access and utilization. *Transcultural Psychiatry* 47(2):224–251.

Chan, L. 2018. *The WD-FAB: Development and validation testing*. Presentation to the European Union of Medicine in Assurance and Social Security, March 2. https://www.eumass.eu/wp-content/uploads/2018/03/Leighton-Porcino.pdf (accessed April 3, 2019).

Cheville, A. L., K. J. Yost, D. R. Larson, K. Dos Santos, M. M. O'Byrne, M. T. Chang, T. M. Therneau, F. E. Diehn, and P. Yang. 2012. Performance of an item response theory-based computer adaptive test in identifying functional decline. *Archives of Physical Medicine and Rehabilitation* 93(7):1153–1160.

Choi, B. 2015. Measurement precision for Oswestry Back Pain Disability Questionnaire versus a web-based computer adaptive testing for measuring back pain. *Journal of Back and Musculoskeletal Rehabilitation* 28(1):145–152.

CNPAAEMI (Council of National Psychological Associations for the Advancement of Ethnic Minority Interests). 2016. *Testing and assessment with persons & communities of color*. Washington, DC: American Psychological Association. https://www.apa.org/pi/oema/resources/testing-assessment-monograph.pdf (accessed April 3, 2019).

Corzillius, M., P. Fortin, and G. Stucki. 1999. Responsiveness and sensitivity to change of SLE disease activity measures. *Lupus* 8(8):655–659.

Crigler, L., K. Hill, R. Furth, and D. Bjerregaard. 2011. *Community Health Worker Assessment and Improvement Matrix (CHW AIM): A toolkit for improving CHW programs and services*. Bethesda, MD: U.S. Agency for International Development.

deBoer, J. N., A. E. Voppel, M. J. H. Begemann, H. G. Schnack, F. Wijnen, and I. E. C. Sommer. 2018. Clinical use of semantic space models in psychiatry and neurology: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews* 93:85–92. doi: 10.1016/j.neubiorev.2018.06.008.

Dorman, P. J., F. Waddell, J. Slattery, M. Dennis, and P. Sandercock. 1997. Are proxy assessments of health status after stroke with the EuroQol questionnaire feasible, accurate, and unbiased? *Stroke* 28(10):1883–1887.

DuBois, C. L., P. R Sackett, S. Zedeck, and L. Fogli. 1993. Further exploration of typical and maximum performance criteria: Definitional issues, prediction, and white-black differences. *Journal of Applied Psychology* 78(2):205–211. doi: 10.1037/0021-9010.78.2.205.

Duncan, P. W., S. M. Lai, D. Tyler, S. Perera, D. M. Reker, and S. Studenski. 2002. Evaluation of proxy responses to the Stroke Impact Scale. *Stroke* 33(11):2593–2599.

Fliege, H., J. Becker, O. B. Walter, M. Rose, J. B. Bjorner, and B. F. Klapp. 2009. Evaluation of a computer-adaptive test for the assessment of depression (D-CAT) in clinical application. *International Journal of Methods in Psychiatric Research* 18(1):23–36.

Ford, M., K. Lang, K. Liebkemann, and B. Silverstone. 2018. *Discussion with stakeholder representatives*. Presentation to the Committee on Functional Assessment for Adults with Disabilities, Washington, DC, February 26.

Forestier, B., E. Anthoine, Z. Reguiai, C. Fohrer, and M. Blanchin. 2019. A systematic review of dimensions evaluating patient experience in chronic illness. *Health and Quality of Life Outcomes* 17:19.

Fuentes, D., and M. P. Aranda. 2018. Disclosing psychiatric diagnosis to close others: A cultural framework based on older Latin@s participating in a depression trial in Los Angeles County. *Aging & Mental Health* 1–9. doi: 10.1080/13607863.2018.1506738.

Gill, T. M., S. E. Hardy, and C. S. Williams. 2002. Underestimation of disability in community-living older persons. *Journal of the American Geriatrics Society* 50(9):1492–1497.

Hartzler, A. L., L. Tuzzio, C. Hsu, and E. H. Wagner. 2018. Roles and functions of community health workers in primary care. *Annals of Family Medicine* 16(3):240–245.

HealthMeasures. 2018. *PROMIS*. http://www.healthmeasures.net/explore-measurement-systems/promis (accessed April 4, 2019).

IOM (Institute of Medicine). 2001. *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.

IOM. 2003. *Unequal treatment: Confronting racial and ethnic disparities in health care*. Washington, DC: The National Academies Press.

IOM. 2015. *Psychological testing in the service of disability determination*. Washington, DC: The National Academies Press.

Joint Commission on Accreditation of Healthcare Organizations. 2003. *Improving the quality of pain management through measurement and action*. Oakbrook Terrace, IL: Joint Commission on Accreditation of Healthcare Organizations.

Kanfer, R., and P. L. Ackerman. 1989. Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology* 74(4):657–690.

Kessler, R. C., G. Andrews, L. J. Colpe, E. Hiripi, D. K. Mroczek, S. L. T. Normand, E. E. Walters, and A. M. Zaslavsky. 2002. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine* 32(6):959–976.

Koutsouleris, N., L. Kambeitz-Ilankovic, S. Ruhrmann, M. Rosen, A. Ruef, D. B. Dwyer, M. Paolini, K. Chisholm, J. Kambeitz, T. Haidl, A. Schmidt, J. Gillam, F. Schultze-Lutter, P. Falkai, M. Reiser, A. Riecher-Rössler, R. Upthegrove, J. Hietala, R. K. R. Salokangas, C. Pantelis, E. Meisenzahl, S. J. Wood, D. Beque, P. Brambilla, and S. Borgwardt. 2018. Prediction models of functional outcomes for individuals in the clinical high risk state for psychosis or with recent-onset depression: A multimodal, multisite machine learning analysis. *JAMA Psychiatry* 75(11):1156–1172. doi: 10.1001/jamapsychiatry.2018.2165.

Liebkemann, K. 2016. *National Disability Forum: Developing and assessing medical evidence for extreme limitations in the ability to focus on tasks*. https://www.ssa.gov/ndf/documents/Panelist%20-%20Kevin%20Liebkemann.pdf (accessed April 4, 2019).

Liebkemann, K., and K. Lang. 2017 (unpublished). Letter to the Committee on Functional Assessment for Adults with Disabilities on behalf of the National Coalition of Social Security and Supplemental Security Income Advocates, December 5.

Lum, T. Y., W. C. Lin, and R. L. Kane. 2005. Use of proxy respondents and accuracy of minimum data set assessments of activities of daily living. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 60(5):654–659.

Mathias, S. D., M. M. Bates, D. J. Pasta, M. G. Cisternas, D. Feeny, and D. L. Patrick. 1997. Use of the Health Utilities Index with stroke patients and their caregivers. *Stroke* 28(10):1888–1894.

McDonough, C. M., E. Stoiber, I. M. Tomek, P. Ni, Y. J. Kim, F. Tian, and A. M. Jette. 2016. Sensitivity to change of a computer adaptive testing instrument for outcome measurement after hip and knee arthroplasty and periacetabular osteotomy. *Journal of Orthopaedic & Sports Physical Therapy* 46(9):756–767.

Meterko, M., E. E. Marfeo, C. M. McDonough, A. M. Jette, P. Ni, K. Bogusz, E. K. Rasch, D. E. Brandt, and L. Chan. 2015. The Work Disability Functional Assessment Battery (WD-FAB): Feasibility and psychometric properties. *Archives of Physical Medicine and Rehabilitation* 96(6):1028–1035. doi: 10.1016/j.apmr.2014.11.025.

Meterko, M., M. Marino, P. Ni, E. Marfeo, C. M. McDonough, A. Jette, K. Peterik, E. Rasch, D. E. Brandt, and L. Chan. 2018. Psychometric evaluation of the improved Work-Disability Functional Assessment Battery. *Archives of Physical Medicine and Rehabilitation*. Corrected proof available online December 19, 2018. doi: 10.1016/j.apmr.2018.09.125.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2016. *Informing Social Security's process for financial capability determination*. Washington, DC: The National Academies Press.

NIH (National Institutes of Health). 2018. *Selecting a HealthMeasure*. http://www.healthmeasures.net/applications-of-healthmeasures/guidance/selecting-a-healthmeasure (accessed April 4, 2019).

NRC (National Research Council). 2007. *Engaging privacy and information technology in a digital age*. Washington, DC: The National Academies Press.

Oczkowski, C., and M. O'Donnell. 2010. Reliability of proxy respondents for patients with stroke: A systematic review. *Journal of Stroke and Cerebrovascular Diseases* 19(5):410–416.

Odole, A. C., P. O. Ibikunle, and U. Useh. 2016. Culturally sensitive and environment-friendly outcome measures in knee and hip osteoarthritis. *African Journal of Biomedical Research* 19(2):71–77.

Oni, T., N. McGrath, R. BeLue, P. Roderick, S. Colagiuri, C. R. May, and N. S. Levitt. 2014. Chronic diseases and multimorbidity: A conceptual modification to the WHO ICCC model for countries in health transition. *BMC Public Health* 14:575. doi: 10.1186/1471-2458-14-575.

Oude Voshaar, M. A. H., H. E. Vonkeman, D. Courvoisier, A. Finckh, L. Gossec, Y. Y. Leung, K. Michaud, G. Pinheiro, E. Soriano, N. Wulfraat, A. Zink, and M. A. F. J. van de Laar. 2019. Towards standardized patient reported physical function outcome reporting: Linking ten commonly used questionnaires to a common metric. *Quality of Life Research* 28(1):187–197.

Pardasaney, P. K., N. K. Latham, A. M. Jette, R. C. Wagenaar, P. Ni, M. D. Slavin, and J. F. Bean. 2012. Sensitivity to change and responsiveness of four balance measures for community-dwelling older adults. *Physical Therapy* 92(3):388–397.

Pardasaney, P. K., P. Ni, M. D. Slavin, N. K. Latham, R. C. Wagenaar, J. Bean, and A. M. Jette. 2014. Computer-adaptive balance testing improves discrimination between community-dwelling elderly fallers and nonfallers. *Archives of Physical Medicine and Rehabilitation* 95(7):1320–1327.

Pickard, A. S., and S. J. Knight. 2005. Proxy evaluation of health-related quality of life: A conceptual framework for understanding multiple proxy perspectives. *Medical Care* 43(5):493–499.

Poulin, V., and J. Desrosiers. 2008. Participation after stroke: Comparing proxies' and patients' perceptions. *Journal of Rehabilitation Medicine* 40(1):28–35.

PSNet (Patient Safety Network). 2018. *Handoffs and signouts*. https://psnet.ahrq.gov/primers/primer/9/Handoffs-and-Signouts (accessed April 4, 2019).

Riebschleger, M., and I. Philibert. 2011. New standards for transitions of care: Discussion and justification. In *The ACGME 2011 duty hour standard: Enhancing quality of care, supervision, and resident professional development,* edited by I. Philibert and S. Amis, Jr. Chicago, IL: Accreditation Council for Graduate Medical Education. Pp. 57–59.

Ruiz, J. G., S. Priyadarshni, Z. Rahaman, K. Cabrera, S. Dang, W. M. Valencia, and M. J. Mintzer. 2018. Validation of an automatically generated screening score for frailty: The Care Assessment Need (CAN) score. *BMC Geriatrics* 18:106. doi: 10.1186/s12877-018-0802-7.

Salgado, J. F., S. Moscoso, J. I. Sanchez, P. Alonso, B. Choragwicka, and A. Berges. 2015. Validity of the five-factor model and their facets: The impact of performance measure and facet residualization on the bandwidth-fidelity dilemma. *European Journal of Work and Organizational Psychology* 24(3):325–349.

SAS Institute. 2019a. *Natural language processing: What it is and why it matters*. https://www.sas.com/en_us/insights/analytics/what-is-natural-language-processing-nlp.html (accessed April 4, 2019).

SAS Institute. 2019b. *Predictive analytics: What it is and why it matters*. https://www.sas.com/en_us/insights/analytics/predictive-analytics.html (accessed April 4, 2019).

Schalet, B. D., D. A. Revicki, K. F. Cook, E. Krishnan, J. F. Fries, and D. Cella. 2015. Establishing a common metric for physical function: Linking the HAQ-DI and SF-36 PF subscale to PROMIS® physical function. *Journal of General Internal Medicine* 30(10):1517–1523. doi: 10.1007/s11606-015-3360-0.

Scientific Advisory Committee of the Medical Outcomes Trust. 2002. Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research* 11(3):193–205.

Shah, N. D., E. W. Steyerberg, and D. M. Kent. 2018. Big data and predictive analytics: Recalibrating expectations. *JAMA* 320(1):27–28. doi: 10.1001/jama.2018.5602.

SSA (U.S. Social Security Administration). 2018a. *POMS DI 22505.001: Medical and non-medical evidence*. https://secure.ssa.gov/poms.nsf/lnx/0422505001#a (accessed April 4, 2019).

SSA. 2018b. *POMS DI 24503.005: Categories of evidence*. https://secure.ssa.gov/poms.nsf/lnx/0424503005 (accessed April 4, 2019).

Tennant, A., M. Penta, L. Tesio, G. Grimby, J. L. Thonnard, A. Slade, G. Lawton, A. Simone, J. Carter, A. Lundgren-Nilsson, M. Tripolski, H. Ring, F. Biering-Sørensen, C. Marincek, H. Burger, and S. Phillips. 2004. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: The PRO-ESOR project. *Medical Care* 42(1 Suppl.):I37–I48.

Tyser, A. R., J. Beckmann, J. D. Franklin, C. Cheng, S. D. Hon, A. Wang, and M. Hung. 2014. Evaluation of the PROMIS physical function computer adaptive test in the upper extremity. *The Journal of Hand Surgery* 39(10):2047–2051.

Ubalde-Lopez, M., G. L. Delclos, F. G. Benavides, E. Calvo-Bonacho, and D. Gimeno. 2016. Measuring multimorbidity in a working population: The effect on incident sickness absence. *International Archives of Occupational and Environmental Health* 89(4):667–678. doi: 10.1007/s00420-015-1104-4.

Üstün, T. B., N. Kostanjsek, S. Chatterji, and J. Rehm. 2010. *Measuring health and disability: Manual for the WHO disability assessment schedule*. Geneva, Switzerland: WHO Press.

Wang, Y., L. Wang, M. Rastegar-Mojarad, S. Moon, F. Shen, N. Afzal, S. Liu, Y. Zeng, S. Mehrabi, S. Sohn, and H. Liu. 2018. Clinical information extraction applications: A literature review. *Journal of Biomedical Informatics* 77:34–49. doi: 10.1016/j.jbi.2017.11.011.

Ware, J. E., Jr., M. Kosinski, J. B. Bjorner, M. S. Bayliss, A. Batenhorst, C. G. Dahlöf, S. Tepper, and A. Dowson. 2003. Applications of computerized adaptive testing (CAT) to the assessment of headache impact. *Quality of Life Research* 12(8):935–952.

Wessler, B. S., Y. L. Lai, W. Kramer, M. Cangelosi, G. Raman, J. S. Lutz, and D. M. Kent. 2015. Clinical predictive models for cardiovascular disease: Tufts predictive analytics and comparative effectiveness clinical prediction model database. *Circulation: Cardiovascular Quality and Outcomes* 8(4):368–375. doi: 10.1161/CIRCOUTCOMES.115.001693.

Wild, D., A. Grove, M. Martin, S. Eremenco, S. McElroy, A. Verjee-Lorenz, and P. Erikson. 2005. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in Health* 8(2):94–104.

Wittchen, H. U., L. N. Robins, L. Cottler, N. Sartorius, J. Burke, and D. Regier. 1991. Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). Results of the multicenter WHO/ADAMHA Field Trials. *British Journal of Psychiatry* 159(5):645–653. doi: 10.1192/bjp.159.5.645.

Wong, T. M., T. L. Strickland, E. Fletcher-Janzen, A. Ardila, and C. R. Reynolds. 2000. Theoretical and practical issues in the neuropsychological assessment and treatment of culturally dissimilar patients. In *Handbook of cross-cultural neuropsychology*, edited by E. Fletcher-Janzen, T. L. Strickland, and C. R. Reynolds. Boston, MA: Springer. Pp. 3–18.

Zdunek, M., L. A. Jason, M. Evans, R. Jantke, and J. L. Newton. 2015. A cross cultural comparison of disability and symptomatology associated with CFS. *International Journal of Psychology and Behavioral Sciences* 5(2):98–107.

4

Integrated Assessment of Work-Related Functional Ability

As described in subsequent chapters, there are numerous validated tools for measuring physical and mental functional abilities at the impairment, body part, or organ system level. For many patients, however, limitations arise from more than one condition, and for most jobs, adequate performance requires completing multiple tasks and a series of task coordination and task sequencing processes, as discussed in Chapter 2. Although failure to perform a single work-related function in a testing environment may provide evidence of inability to perform that function, success in one domain is not sufficient to establish the ability (or capacity) to perform the job related to that function on a regular and continuing basis in the actual work setting or a different work setting. As discussed in Chapter 3, the usual test environment does not adequately reproduce or predict the sustained and repeated task performance required for work. Moreover, the successful integration of multiple tasks and skills into an effective day of work is influenced by multiple aspects of physical and mental health, as well as by environmental and interpersonal aspects of the work setting. Because successful work performance is more than the sum of the individual functions required, the degree of limitation to perform work often exceeds the sum of individual limitations. Comorbidities (e.g., depression and low back pain) that may themselves not present measurable limitations frequently exacerbate the impact of recognized limitations and reduce the ability to compensate for them. In addition, older age is associated with progression of most diseases, increasing prevalence of comorbidity, and diminishing resilience. Accurate assessment of an individual's ability to work therefore requires an integrated approach that considers the totality of the person's physical, cognitive, and adaptive conditions and aligns them with the full scope of tasks required for the work and the schedule and environment in which they will be performed (see Figure 2-3 in Chapter 2).

Types of assessment instruments range from specific to integrated (Reiman and Manske, 2011). Many objective tests quantitate specific functions expected to be reduced by a particular condition, such as range of motion and strength of contraction for musculoskeletal disease, aerobic exercise tolerance for cardiovascular conditions, and cognitive function for traumatic brain injury (far left of Figure 2-3). Impairment-specific assessment instruments may provide information on the progression of or recovery from a specific disease process, but the validity of their use for other conditions is unknown, and they are unlikely to capture the effects of multiple impairments on an individual's ability to function. Qualitative assessments by individuals, their health care and rehabilitation providers, and other third parties with knowledge of the individual often yield scores regarding the integrated effect of individuals' impairments on general daily life and participation (e.g., activities of daily living [ADLs]) and/or on the performance of a specific job, the job's tasks, and its mental and physical demands. Such integrated assessments are useful for capturing the additive and sometimes multiplicative effects of multiple impairments and comorbid conditions on an individual's functional ability to meet work requirements. Therefore, the most informative evaluation of function may include integrated assessments in addition to specific assessments of body structures and systems. However, it is easier to identify focused tests that could identify the inability to perform a specific activity relevant to a work requirement (e.g., inability to reach overhead) than to find a general test that demonstrates ability to perform all of the functions required.

This chapter responds to two parts of the committee's charge: discussion of "generic [versus 'impairment-specific'] functional assessment questionnaires" and identification of any "activities of daily living that correlate with the physical and mental … demands of work." Three categories of integrated ("generic") assessment tools are described: (1) those focused on ADLs and instrumental activities of daily living (IADLs), which assess function in terms of goal-based tasks; (2) self-report instruments designed to assess function in terms of physical and mental activities and tasks; and (3) those that measure limitations in work performance activity due to health conditions (see Figure 2-3).

ACTIVITIES OF DAILY LIVING

Given their universality, ADLs are a common focus for integrated assessment of individuals' functional abilities, regardless of underlying medical conditions. ADLs are basic tasks of daily life that typically include personal care and hygiene, dressing, feeding, continence management, and mobility. IADLs are more complex tasks related to independent living in the community, such as navigating transportation options and shopping, preparing meals, managing one's household, managing finances and medications, communicating with others, and providing companionship and mental support. Assessment of ADLs and IADLs is a common way to assess an individual's ability to perform multiple, integrated functions on a day-to-day basis.

Individuals usually are referred for assessment of ADLs or IADLs in the context of impairments in cognitive or physical functioning. Information from these assessments is typically used to inform the type of assistive devices individuals may need to improve their safe performance of these activities or the amount of assistance individuals may need and/or the living situation they require to perform the activities given their impairment level. Assessment of ADLs and IADLs also considers the contribution of the built environment to one's ability to perform his or her ADLs and IADLs, as well as the individual's social context.

The performance of ADLs and IADLs can be assessed by self-report, informant (third-party) report, specific assessments outlined below, and/or direct observation. Most commonly they are assessed through a combination of self-report and direct observation because self-report alone is often not considered valid for those with substantial cognitive impairments or because of other threats to validity. However, assessment by direct observation requires more training to administer relative to self- or third-party report. Thus in many clinical settings, trained occupational therapists, physical therapists, speech-language pathologists, and/or nurses perform assessments—including direct observation—to determine ADL capacity (Mlinac and Feng, 2016). Performance-based IADL assessments are most relevant to assessing a wide range of functional abilities affected by mild changes in cognitive functioning. These performance-based IADL assessments are more commonly used in specific types of clinical settings, such as rehabilitation hospitals, because of the complexity and time intensity of their administration, which in some cases requires an occupational therapist.

An individual's ability to perform ADLs and IADLs depends on his or her motor abilities, cognitive abilities, and perceptual and sensory abilities (Mlinac and Feng, 2016). Individuals must have the cognitive ability to plan and reason and the motor abilities of balance and dexterity to perform these activities. They must not only be able to complete tasks but also to recognize that they need to do so (Mlinac and Feng, 2016). Other factors also affect the performance of ADLs and IADLs, including, for example, the built environment, such as accessible features in the living space; access to assistive technology; and one's social context or circumstances, such as the availability of attendant care and/or assistance from family members and others. Referring to the committee's conceptual framework, these environmental and social factors are within a different context from that of the organizational and work environmental factors described in Chapter 2 (right-hand side of the dotted line in Figure 2-3). These contextual differences may add challenges to translating or mapping limitations in the ADL and the IADL domain to the work domain.

ADLs are largely unaffected by mild cognitive impairment. Jefferson and colleagues (2008) found no difference in ADL function between individuals with mild cognitive impairment and those with no cognitive impairment. They found that as cognitive impairment worsens, the correlation between cognitive function and ADL dependence appears more consistent. In a study of women and men with mild Alzheimer's disease, measures of attention predicted overall ADL scores, executive function predicted both ADL and IADL scores, and language predicted IADL scores (Hall et al., 2011). Gender differences have been found in the domains of learning and memory, as well as the association between specific cognitive functions and different ADLs or IADLs (Hall et al., 2011). In IADL assessments, independence is one of the distinguishing features of normal aging versus mild cognitive impairment and dementia (Gold, 2012). In healthy aging, the ability to perform IADLs usually remains intact until individuals reach their 80s.

A meta-analysis focused on elucidating the cognitive processes that underlie IADLs in community-dwelling older adults, including those with mild cognitive impairment, found that 21 percent of variance in IADL capacity was predicted by cognition. General cognitive functioning was found to be important in multiple studies, with executive functioning and memory accounting for more variance than other cognitive domains (Gold, 2012). It should be noted that this meta-analysis did not control for assessment approach (e.g., self- or informant report versus observation) (Royall et al., 2007). Thus, executive skills and other cognitive domains likely support IADL performance, but it is important to recognize that the relationship between cognition and IADLs depends on how IADLs are measured.

Depression is one factor that limits the performance of ADLs and IADLs, irrespective of physical or cognitive performance issues or age. Meltzer and colleagues (2012) looked at disability as measured by reported difficulties with ADLs and IADLs in people living with depression, using a large national survey of psychiatric morbidity among adults across the age spectrum in the private household population of England. The results showed that disability was associated with depression even after adjustment for physical health issues (Meltzer et al., 2012). In fact, the number of ADL/IADL difficulties reported by subjects was directly related to the likelihood of their having depression (Meltzer et al., 2012). Meltzer and colleagues concluded that limitations in all domains of ADLs and IADLs are significantly associated with depression. They concluded further that the effect is cumulative irrespective of whether the limitation is in personal care or mobility (Meltzer et al., 2012).

In a Swedish study, Boström and colleagues (2014) looked at the association between depression and functional capacity, dependency in performing ADLs, and dependency in performing individual ADL tasks in 392 older adults living in the community and in residential care facilities. They found that while overall ADL performance was not associated with depression, dependency in the ADL tasks of transfers and dressing appeared to be associated with depressive symptoms (Boström et al., 2014).

Assessment Instruments

A literature review of 15 common ADL and IADL assessment instruments conducted by the committee for this study[1] did not indicate or support a correlation or association between specific ADL or IADL measures and the ability to work. Many of these instruments are used with older nonworking adults or those who have experienced strokes, head injuries, and/or psychiatric impairments. None is specifically designed to assess work capabilities. Most have been validated with more than one group of people with specific functional or cognitive limitations or disabling conditions or in various settings (i.e., acute care, rehabilitation, community dwellers). The 15 instruments reviewed are as follows:

- ADL Profile (head injury and stroke) (Dutil et al., 1990),
- ADL-Focused Occupations-Based Neurobehavioral Evaluation (Gardarsdóttir and Kaplan, 2002),
- Assessment of Motor and Processing Skills (Fisher and Bray Jones, 2010),
- Barthel Index (Quinn et al., 2011),
- Bay Area Functional Performance Evaluation (Houston et al., 1989),
- Cleveland Scale of Activities of Daily Living (dementia) (Patterson and Mack, 2001),
- Executive Function Performance Test (Baum et al., 2008),
- Frenchay Activities Index (IADLs) (Schuling et al., 1993),

[1]This literature review focused on assessment instruments commonly available to clinicians, and was conducted using standard databases. During the course of the review, the committee found a compilation of ADL and IADL assessments in a standard reference work on assessments (Asher, 2014), which also was used as a source. Each ADL and IADL assessment instrument was reviewed with respect to the purpose of the assessment and the population for whom it was intended and/or standardized. An additional search was conducted to determine whether any of the instruments identified had been studied specifically for any relationship with the ability to work or return to work.

- Functional Independence Measure (Ottenbacher et al., 1996),
- Katz ADL Scale (elderly and chronically ill) (Katz, 1983; Katz and Akpom, 1976),
- Kohlman Evaluation of Living Skills (IADLs—psychiatric geriatric) (Burnett et al., 2009; Kohlman-Thomson, 1992),
- The Lawton IADL Scale (Graf and Hartford Institute for Geriatric Nursing, 2008),
- Manual Ability Measure (neurological and musculoskeletal conditions) (Chen and Bode, 2010),
- Multiple Errands Test (brain injury, stroke—executive functioning) (Morrison et al., 2013), and
- Performance Assessment of Self-Care Skills (Chisholm et al., 2014).

As previously noted, several articles address the association between ADL and IADL assessments and early cognitive decline and/or dementia (Mlinac and Feng, 2016; Patterson and Mack, 2001; Sikkes et al., 2009). Many of the assessments listed above are used with this population.

Activities of Daily Living and Work

Although little research exists to connect specific assessments of ADLs and IADLs to an individual's ability to return to work, Cancelliere and colleagues (2016) conducted a best-evidence synthesis of 56 systematic reviews judged to have a low risk of bias based on the Scottish Intercollegiate Guidelines Network. They looked at common prognostic factors for return to work across different health and injury conditions in an effort to describe the association of these factors with return-to-work outcomes. Half of these systematic reviews focused on prognostic factors for return to work for musculoskeletal disorders, related primarily to the spine, while the remaining half focused on prognostic factors for return to work for mental health disorders, cardiovascular conditions, stroke, cancer, multiple sclerosis, and other (nonspecified) health conditions. The reviews found that factors commonly associated with positive return to work included "higher education and socioeconomic status, higher self-efficacy and optimistic expectations for recovery and return-to-work, lower severity of the injury/illness, return-to-work coordination, and multidisciplinary interventions that include the workplace and stakeholders" (Cancelliere et al., 2016, p. 1). Common prognostic factors found to be associated with negative return to work included "older age, being female, higher pain or disability, depression, higher physical work demands, previous sick leave and unemployment, and activity limitations" (Cancelliere et al., 2016, p. 1). Cancelliere and colleagues include limited ability to perform ADLs among "activity limitations," which they also refer to as "participation restrictions."

The only ADL assessment showing any association with ability to work was the Assessment of Motor and Process Skills, which found "a moderate correlation between the level of employment and the global scores of the process skills scale" (Haslam et al., 2010). However, the population of this study was limited to 20 individuals with schizophrenia who were engaged either in competitive employment, supported employment, prevocational training, or nonvocational activities. Thus, no ADL or IADL assessments exist that are standardized on a working-age population with limitations across multiple physical and cognitive areas.

However, information from ADL and IADL assessments can contribute information about an individual's ability to function, and further inquiry can provide additional information relevant to the ability to work. For example, an individual who cannot independently get out of bed, bathe or shower, dress, use the toilet, feed him- or herself, take medication, or manage money to navigate transportation to get to work would most likely be unable to work without significant assistance from family members or an attendant. Individuals with such severe limitations may be expected to meet the criteria for disability at step 3 of the determination process.[2] Conversely, individuals who report the ability to perform ADLs and IADLs may nevertheless be unable to work. They may be able to complete such tasks, but doing so may lead to pain, fatigue, and/or other limitations that interfere with the ability to work. Therefore, the collection of information about ADLs and IADLs could be enhanced by asking follow-up questions about context; environmental factors; required actual assistance; and the impact of performing ADLs and IADLs on pain, fatigue, confusion, concentration, and other physical or cognitive factors that can interfere with the performance of work. The inquiry might explore, for example, whether an individual is able to function after performing all of the morning ADL tasks required to get up and go to work mentioned above, or whether after completing those tasks, the individual is too fatigued or in too much pain to function in a work or any other environment.

Minimal evidence in the literature indicates that limited ability to perform ADLs is a prognostic factor associated with poor return-to-work outcomes. Yet, while many ADL and IADL assessment instruments exist, no one standard assessment can predict whether a person with a given condition or impairment will be able to return to work. Various ADL and IADL assessments may provide information about individuals' ability to meet their personal hygiene needs or manage their medications, for example, but no ADL and IADL assessments correlate directly with the ability to work for all conditions and/or impairments. ADL and IADL assessments designed for specific conditions may provide information about a person's level of

[2] See 20 CFR 404.1520; 20 CFR 416.920.

functioning in specific ADL and IADL areas, but no direct association between this level of functioning and the ability to return to work has been demonstrated for many reasons, including differences in contextual factors and the demands of jobs versus those of self-care. Again, then, assessment of individuals' ability to perform ADLs and IADLs needs to include follow-up questions about the context in which they perform those activities; the amount of assistance they have in doing so; and how they function subsequently in terms of pain, fatigue, confusion, and/or limitations.

INSTRUMENTS AND BANKS OF INSTRUMENTS FOR ASSESSING PHYSICAL AND MENTAL FUNCTION

As stated previously, chronic health conditions, along with comorbidities, can manifest differently in terms of work-related functional and disability limitations at the activity and work task levels—differences in function that are not captured by assessment instruments specific to impairments and body parts or organ systems. This section describes several evidence-based instruments and sets of instruments that can provide integrated information about individuals' overall functional capabilities and limitations and help inform determinations of work disability, although only the Work Disability Functional Assessment Battery (WD-FAB) is designed to address work-related function directly. Some of the scales and measures from these instruments and banks are discussed in more detail in subsequent chapters.

Work Disability Functional Assessment Battery (WD-FAB)

WD-FAB is a functional assessment tool developed through an interagency agreement established in 2008 between the U.S. Social Security Administration (SSA) and the Intramural Research Program of the National Institutes of Health (NIH) Clinical Research Center (Chan, 2018a). The instrument was developed through a scientifically rigorous process that used the *International Classification of Functioning, Disability and Health* (ICF) to conceptualize function and included an extensive literature review, consultation with content experts and focus groups consisting of health care providers and individuals with disabilities, cognitive testing of all items to check clarity and comprehension, and administration of items to user groups (Chan, 2018a). The WD-FAB prototype was completed in 2016, and replenishment of items was completed in 2017. Numerous publications have reported on the development and scientific validation of this instrument (Marfeo, 2013a,b,c, 2014, 2015, 2018; Marino et al., 2015; McDonough et al., 2013, 2017, 2018; Meterko et al., 2015, 2018; Ni et al., 2013).

WD-FAB uses item response theory (IRT) and computer adaptive testing (CAT) methods to create an individualized measure that can best measure the "ability" of the person being tested. CAT algorithms customize the selection of items in real time from the more than 300 items banked, thus reducing respondent burden and allowing comprehensive assessment of functional activity in approximately 15 to 20 minutes (Chan, 2018a,b; Meterko et al., 2018). The instrument encompasses two primary domains—physical function and mental health function. The physical function domain consists of basic mobility; upper-body function; fine motor function; and community mobility, which includes driving, public transportation, and wheelchair use (Chan, 2018a,b; Meterko et al., 2018). The mental health function domain consists of communication and cognition, resilience and sociability, self-regulation, and mood and emotions (Chan, 2018a,b; Meterko et al., 2018).

WD-FAB has demonstrated good test-retest reliability in adults with work disability and general adult samples, very high accuracy for physical function, and moderate to strong convergent validity correlations with legacy measures (Chan, 2018a,b). It has shown more variability in accuracy in the mental health domain.

Strengths of WD-FAB include multiple administration modes (in-person, phone, Web-based, paper/pencil via short forms), depending on user needs[3]; creation of functional profiles and the ability to track changes over time; item pools that can be replenished and improved; a standardized and consistent approach to the assessment of function; and establishment of thresholds for minimal detectable differences (Chan, 2018a,b). The investigators' sampling criteria enhance generalizability (Meterko et al., 2018). They recruited three samples from a U.S. national panel: (1) working-age adults matched to the U.S. adult population on age, sex, race, ethnicity, and education; (2) adults in the same age range who reported a permanent disability for a physical condition; and (3) adults who reported a permanent disability for a mental health condition.

A limitation is that WD-FAB outcomes must be linked to workplace demand—a challenge confronted by all disability benefits programs. WD-FAB measures at the activity level according to the ICF, whereas assessment of work disability requires linking activity (whole-person functioning) to participation (work) (as discussed in Chapter 2). Another limitation is the lack of large-scale validation with diverse samples. One potential approach would be to use WD-FAB to develop functional profiles by occupation (Chan, 2018b). Considerable effort has been expended on the development of WD-FAB, but as noted, the link between whole-person functioning and participation in work has not been established and requires

[3]Equivalence of multiple administration modes has been well established in general (Gwaltney et al., 2008) but not specifically for WD-FAB.

further research. Developing functional profiles using WD-FAB for physical function may be useful. In summary, WD-FAB has great potential to enhance SSA's methods of obtaining claimants' perspectives on their function in specific domains that correspond closely to the functional criteria used in disability determination decisions, but further research is required.

World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0)

WHODAS 2.0 can be used to assess overall disability and health with several subscales based on the ICF disability model (see Figure 2-1 in Chapter 2) (WHO, 2018): cognition, mobility, self-care, getting along, life activities, and participation. The last two were added in the 2.0 version of the instrument in accordance with the ICF model. The design intent included capturing cross-cultural constructs of disability that could be applied internationally. There are two versions of WHODAS 2.0: a 36-item version that takes approximately 20 minutes to complete and a 12-item version that takes about 5 minutes. Population norms exist, and the long version can be scored with or without weighting. While the 36-item version of WHODAS 2.0 provides greater detail about the experience of disability, the 12-item version is sufficient when it is important to minimize respondent burden.

The psychometric properties and validation of this instrument have been the focus of numerous studies for specific health issues and languages since it was developed and first described in 2010. Üstün and colleagues (2010) report its properties based on its use with more than 65,000 individuals from across the globe with physical disorders and/or mental health issues or addictions. They report good internal consistency, with a Cronbach's alpha of 0.86; good and stable structure, as determined through principal component analyses; and a high intraclass coefficient of 0.98 for reliability (Üstün et al., 2010). The scales of the instrument also correlated well with other health measures, including the 36- and 12-Item Short-Form Health Survey (SF-36/12), the World Health Organization's (WHO's) Quality of Life Score, the London Handicap Scale, and the Functional Independence Measure. The instrument also responded when clinical interventions were applied, showing good effect sizes. WHODAS is a generic measure of health and disability status that may be thought of as similar to the SF-36. It was developed using classical test theory (unlike the Patient-Reported Outcomes Measurement Information System [PROMIS] or WD-FAB, whose developers used IRT/CAT methods). Although it may be used as a screening device for self-reported symptoms, which may or may not be related to ability to work, it cannot be used as a direct measure of employability.

Patient-Reported Outcomes Measurement Information System (PROMIS)

PROMIS is one of several sets of instruments developed and evaluated with NIH funding (two of the other NIH initiatives are described below). PROMIS is "a set of person-centered measures that evaluates and monitors physical, mental, and social health in adults and children" (NU, 2018e). It can be used with individuals with chronic conditions and the general population and can be administered in three modes—paper, computer, and an app (Amtmann et al., 2011; Cella et al., 2010). Self-reported health measures are arranged according to the domains of physical, mental, and social health (Hahn et al., 2010). Within each domain, subdomains are listed that relate functions, symptoms, behaviors, and affect.

During the first phase of PROMIS (2004–2009), funded through an NIH Roadmap Initiative, the goal was to develop an efficient state-of-the-art assessment tool for self-reported health. This effort resulted in the development of patient-reported outcome measures using large item banks and CAT, allowing for effective assessment of patient-reported outcomes in clinical research (HHS, 2018). The second phase of PROMIS (2009–2014) saw the addition of such features as longitudinal analyses, more sociodemographically diverse samples, increased emphasis on pediatric populations, and evaluation of PROMIS item banks (HHS, 2018). Additional federal and foundation funding supported the development of condition-specific derivatives of PROMIS, including those for neurological disorders (Neuro-QoL) (NU, 2018f), spinal cord injury (SCI-QoL) (Tulsky et al., 2015), traumatic brain injury (TBI-QOL) (Tulsky et al., 2016), and Huntington's disease (HDQLIFE) (Carlozzi et al., 2017a,b). We note that the NIH Roadmap Initiative specifically funded efforts to ensure that all relevant documentation and evidence for PROMIS would be freely available to the public on Web-based platforms. The PROMIS tools were developed primarily for research and clinical tracking of patients' perceptions of their illness and its impact on their lives. In many institutions, these tools have become available as part of the electronic medical record during routine care for assessing such variables as depression, pain interference, social functioning, and the like. Although there is no research to support using these instruments to predict employability per se, they are high-quality measures for estimating functioning in domains thought to be relevant to employability and may contribute to an overall understanding of employment potential.

National Institutes of Health (NIH) Toolbox

The NIH Toolbox, developed with funding from the NIH Blueprint for Neuroscience Research, is "a comprehensive set of neuro-behavioral measurements that quickly assess cognitive, emotional, sensory, and motor functions from the convenience of an iPad" (NU, 2018d; see also Akshoomoff et al., 2014; Carlozzi et al., 2015, 2017c; Denboer et al., 2014; Dikmen et al., 2014; Gershon et al., 2014; Loring et al., 2018; Tulsky et al., 2014; Weintraub et al., 2013). Advanced approaches, such as IRT/CAT, were used in item development, test scoring, and construction. The Toolbox contains two types of measures: performance-based tests of function (objective measures) and self-report and proxy measures (primarily for emotion). Phase I of its development employed such qualitative methods as online requests for information from experts, interviews with clinicians and scientists, and consensus meetings to identify subdomains to be included. In Phase II, candidate measures were pilot tested, and initial evaluations of psychometric properties were performed. Subsequently, a number of validation studies have been performed on people of varying ages and health status (see NU, 2018c).

The cognitive measures in the Toolbox were designed to be completed in 30 minutes; they provide reliable estimates of specific cognitive and language skills, and demonstrate evidence of external validity (Akshoomoff et al., 2014; Weintraub et al., 2013). Administration requires an annual licensing fee to activate an iPad application and verification of a clinical psychologist's involvement in test administration and score interpretation. Potential users should note that research literature on the Toolbox is relatively limited given its recent development. Additional studies are needed to demonstrate its utility and feasibility in clinical and disability determination contexts.

The NIH Toolbox can be used both for individuals in the general population and for those with chronic conditions and consists of four batteries:

- The Cognition Battery focuses on mental processes used to gain knowledge and comprehension, such as thinking, knowing, and remembering (NU, 2018b). It also encompasses language, imagination, perceptions, and the planning and execution of complex behaviors. This battery yields the following summary scores: cognitive function composite score, fluid cognition composite score (includes picture sequence memory, list sorting, and pattern comparison measures), and crystallized cognition composite score (includes picture vocabulary and reading recognition measures).
- The Emotion Battery can be used to assess strong feelings such as joy, sorrow, or fear. The NIH Toolbox includes four major domains of emotion: psychological well-being, stress and self-efficacy, social relationships, and negative affect. The Emotion Battery surveys positive affect, general life satisfaction, emotional support, friendship, loneliness, perceived rejection, perceived hostility, self-efficacy, sadness, perceived stress, fear, and anger.

- The Motor Battery targets the ability to use and control muscles and movements (NU, 2018b). It consists of tests to assess dexterity, grip strength, standing balance, gait speed, and endurance.
- The Sensation Battery addresses "the biomechanical and neurologic process of detecting incoming nerve impulses as nervous system activity" (NU, 2018b). It consists of tests to assess audition, visual acuity, olfaction, taste (ages 12+), and pain (ages 18+).

Overall, the Toolbox serves as a well-developed set of self-report and performance-based measures that are relatively brief to administer. The self-report portion is similar in administration to PROMIS and Neuro-QoL, discussed below, and is customarily completed on a tablet. The performance-based section is similar to and could provide an alternative to tests in standard neuropsychological evaluations, access to which may be limited by their expense and the availability of professionals qualified to administer and interpret them. Examiners must be trained to use the Toolbox instruments, but the training should not represent a significant barrier given that the instruments were designed to be completed by a nonclinician with a baccalaureate degree. To date, there is no evidence to allow direct inference from scores to employability, although domains considered to be relevant to employment may be measured.

Quality of Life in Neurological Disorders (Neuro-QoL)

Neuro-QoL is "a measurement system that evaluates and monitors the physical, mental, and social effects experienced by adults and children living with neurological conditions" (Cella et al., 2012; Gershon et al., 2012; NU, 2018a). Sponsored by the National Institute of Neurological Disorders and Stroke, Neuro-QoL instruments were developed to construct psychometrically sound and clinically relevant health-related quality of life (HRQOL) measurement tools for individuals with neurological conditions or disorders such as stroke, multiple sclerosis, Parkinson's disease, epilepsy (Nowinski et al., 2010; Victorson et al., 2014), amyotrophic lateral sclerosis, Huntington's disease (Carlozzi et al., 2017a,b), and muscular dystrophy (NU, 2018c). The HRQOL domains were identified through an extensive literature review, an online Request for Information, two phases of in-depth expert interviews, patient and caregiver focus groups, and individual interviews with patients and proxies (NU, 2018c). Based on this input, 17 HRQOL domains and subdomains were selected for adults and 11 for children (Gershon et al., 2012). Neuro-QoL measures functions, symptoms, behaviors, and affect. The adult domains include ability to participate in social roles and activities, satisfaction with social roles and activities, anxiety, bowel function, cognitive function, communication, depression, emotional and behavioral dyscontrol, fatigue, lower-extremity function/mobility, positive affect and well-being, satisfaction with social roles and activities, sleep disturbance, sexual function, stigma, upper-extremity function/fine motor skills, and urinary/bladder function (Neuro-QoL investigators, 2015; NU, 2018c).

As is true for the PROMIS measures and HRQOL in general, there is no evidence to support drawing inferences directly from scores on Neuro-QoL instruments to employability; however, domains relevant to employment may be assessed as part of an overall assessment of work-related function. Neuro-QoL is well developed, and many scores on this instrument may be cross-walked with PROMIS.

Summary

There is strong evidence to support using the PROMIS and Neuro-QoL instruments to assess variables thought to influence participation in general, including participation in employment. The NIH Toolbox is an efficient, well-validated set of performance-based and self-report scales for appraising cognitive status, emotion, motor function, and sensory status. There is no evidence to support drawing direct inferences from scores on PROMIS, Neuro-QoL, or the NIH Toolbox to employability, although scores from these instruments may be very useful in understanding the functioning of an applicant. WD-FAB is a new scale with better validity in the physical than in the mental health domain, and research on this instrument is ongoing. Currently, WD-FAB may be most useful for understanding self-reported physical function, but direct inferences from WD-FAB to employability are not warranted at this time.

INSTRUMENTS FOR MEASURING LIMITATIONS IN WORK ACTIVITY DUE TO HEALTH CONDITIONS

For workers, a number of self-report instruments examine interruption of work and quantify functional limitations at work due to health conditions that have been related to worker productivity. Many of these instruments were developed by groups in the United States and Europe and have been translated and used globally, often in the area of occupational health and safety research. Conceptually, such instruments quantify individual health-related reductions in productivity and the costs to employers (Lerner et al., 2003).

The conceptual elements of these instruments are often referred to as absenteeism (time away from scheduled work) and presenteeism (time when workers are at work, even if they are working at reduced productivity). Although these concepts apply to individuals who are currently working, they can provide a measure of residual capability to retain a job. While an individual may be able to work, he or she may be unable to hold a job because of missing too much time or being engaged ineffectively as a result of his or her condition or its treatment. These constructs are helpful for researchers, employers, and other professionals to understand how impaired health interferes with work. The instruments are not disease specific but are frequently validated in populations with a specific disease diagnosis.

These instruments often are used to evaluate and monitor employer-provided health insurance and wellness programs (Pronk et al., 2016). In light of rising health care costs in the United States and elsewhere, companies need methods and tools that enable them to measure the effectiveness of their programs (Sorensen et al., 2016, 2018). Most guidelines for effective workplace health and safety programs, including programs for health promotion and disability management, encompass evaluation (NIOSH, 2008; Sorensen et al., 2018).

Work Limitations Questionnaire (WLQ)

WLQ is a self-report instrument designed to measure the impact of one's health condition in limiting work activity. Specifically, its results provide a measure of how much time in the past 2 weeks respondents' health limited their capability to work with respect to several categories of job demands (Lerner et al., 2001). WLQ has 25-, 16-, and 8-item versions. Respondents rate the amount of difficulty they had in performing physical and mental job activities during the previous 2 weeks. The 25-item version of the instrument has four subscales: time management demands (5 items), physical demands (6 items), output demands (5 items), and mental-interpersonal demands (9 items). The time management subscale asks about the difficulty in time management and scheduling demands. Physical demands include strength, stamina, movement, coordination, and flexibility, while output demands include work quantity and quality. The mental-interpersonal demands subscale includes items related to completing cognitive tasks at work as well as social interaction in the work setting. These subscales have high internal consistency reliability, with Cronbach's alphas greater than 0.89, in both patient and employee populations (Lerner et al., 2001). The WLQ scales correlate with SF-36 measures of physical and mental health and severity in people with depression and osteoarthritis (Adler et al., 2006; Lerner et al., 2002, 2003). Criterion validity has been tested in several settings and for various types of health issues, including low back pain (Denis et al., 2007), cancer (Feuerstein et al., 2007), heart disease, and other chronic illnesses (Munir et al., 2007). The 25-item questionnaire can be completed in 5–10 minutes, and it has been translated into more than 30 official languages (Tufts Medical Center, 2018). Internet, phone, and mail versions are available.

Work Ability Index (WAI)

WAI is a 7-item questionnaire designed to measure the work ability of individuals in an occupational health clinic environment. Results provide an indication of the length of time individuals are able to work at their jobs (Healthy Workplaces, n.d.; Society of Occupational Medicine, 2018). The WAI questionnaire covers the following dimensions of individuals' capability: their current work ability compared with their lifetime best, their work ability in relation to the demands of the job, the number of diagnosed illnesses or limiting conditions from which they suffer, their estimated impairment owing to diseases/illnesses or limiting conditions, the amount of sick leave they have taken during the last year, and their own prognosis of their work ability in 2 years' time. From these responses, a score is calculated, and this score is used to categorize the individual's work capability as poor, moderate, good, or excellent. The test-retest reliability of the questionnaire is acceptable (de Zwart et al., 2002). Validity studies were conducted in the 1990s by the Finnish Institute of Occupational Health (Ilmarinen et al., 1997; Tuomi et al., 1991, 1997, 2001). Overall, WAI has predicted early retirement, work disability, absence due to sickness, and mortality fairly well. Administering the questionnaire is straightforward; however, calculating its scores is more difficult. WAI has been translated into 24 languages.

World Health Organization Health and Work Performance Questionnaire (HPQ)

HPQ can be used to assess several work performance measures, including quality and quantity of work or presenteeism, number of absences, and critical incidents on the job (Harvard Medical School, 2005; Kessler et al., 2003, 2004). It was developed in part to provide employers with information on the amount of loss (in productivity) related to their workers' health, or as the authors put it on their webpage, "to increase the rationality of employer-sponsored health care purchasing" (Harvard Medical School, 2005).

The HPQ presenteeism scale has four items. Only one of these provides an absolute measure of presenteeism in the past 28 days; the other three questions help calculate a relative scale while providing a reference for respondents' answers to the single item measuring absolute presenteeism. For absences, HPQ provides a series of questions that act as a worksheet for determining the number of hours worked and the amount of work missed because of health considerations during the past 28 days. Finally, HPQ can be used to assess critical incidents, specifically workplace accidents, via a single question.

HPQ has been translated into 24 languages, with surveys in 6 languages being available on its webpage (Harvard Medical School, 2005). It has been validated in a number of industries among working populations diagnosed with a variety of health conditions, including mental health disorders (e.g., Kessler et al., 2003, 2010; Scuffham et al., 2014; Sevak et al., 2017; Suzuki et al., 2015). By determining the number of hours of limited work or absenteeism, its results can be roughly associated with costs to employers in terms of lost productivity (Kessler et al., 2004; Scuffham et al., 2014). As with WAI, administering the questionnaire is straightforward, but it is more difficult to calculate the scores. The questionnaire focuses on current employees and therefore is of limited use for individuals not presently working. However, it may be of use for individuals who are attempting to return to work but continue to miss time because of their health conditions.

Work Productivity and Activity Impairment Questionnaire (WPAI)

WPAI is a 6-item questionnaire used to assess both presenteeism and absenteeism at work during the past 7 days (Margaret Reilly Associates, 2013). It has been translated into more than 100 languages. While designed for general health-related issues, it has been modified to ascertain the impact on work of various health conditions, including mental health disorders and musculoskeletal pain and related disorders (e.g., Asami et al., 2015; Reilly et al., 2010). Validation studies are numerous. During the instrument's original development, construct validity was examined in a small (N = 106) group of workers with health problems (Reilly et al., 1993). Construct validity was found to be good, and the validation explained a majority of the variance (54–64 percent) in the WPAI variables (Reilly et al., 1993). Specifically, WPAI had good positive correlations with SF-36 measures and symptom severity. Similar to other instruments of its type, it is used with people who are currently working, with its questions pertaining to absences and productivity at work within the previous 7 days.

Work Limitation Summary

All of the above instruments have been validated through varying processes, making them credible tools for evaluating capacity and functional limitations. Many of these instruments were developed to measure presenteeism and/or absenteeism among currently employed individuals. Only one, WAI, appears to have been developed with the specific goal of measuring and assessing the respondent's capability to perform work. The different instruments vary considerably as measures of costs in lost work, showing a large amount of variability when considering the same populations (Gardner et al., 2016).

FINDINGS AND CONCLUSIONS

Findings

4-1. Specific assessment instruments measure physical and mental functional abilities at the impairment, body part, or organ system level. Integrated assessments can capture the additive and sometimes multiplicative effects of multiple impairments and comorbid conditions on individuals' functional abilities.

4-2. Activities of daily living (ADLs) are well understood in the health care field and provide a common focus for an integrated assessment of functional abilities, regardless of underlying medical conditions.

4-3. ADLs and instrumental activities of daily living (IADLs) are assessed through a combination of self-report, proxy report, and direct observation.

4-4. The ability to perform ADLs and IADLs is affected by a person's motor abilities, cognitive function, and perceptual and sensory abilities, although ADLs are largely unaffected by mild cognitive impairment.

4-5. Depression can limit performance of ADLs or IADLs irrespective of physical or cognitive impairments or age.

4-6. Factors increasing the likelihood of return to work include higher education and socioeconomic status, higher self-efficacy and optimism for returning to work, lower injury severity, and availability of multidisciplinary interventions. Factors associated with a lower likelihood of return to work include older age, female gender, higher pain or severity of disability, depression, higher work demands, previous sick leave and unemployment, and limitations in current activity.

4-7. While many instruments assess the performance of ADLs and IADLs, no ADL or IADL assessments exist that are standardized on a working-age population with limitations across multiple physical and cognitive areas that map to the work context.

4-8. Research is limited on the relationship between assessments of ADL and IADL performance and an individual's ability to return to work.

4-9. There is little evidence that inability to perform ADLs or IADLs predicts poor return-to-work outcomes.

4-10. Several evidence-based instruments and instrument sets are available that can provide integrated information about individuals' overall functional capabilities and limitations.

4-11. The Work Disability Functional Assessment Battery (WD-FAB) demonstrates good reliability in adults with work disability, very high accuracy for physical function, and good construct validity.
4-12. WD-FAB is a flexible tool that allows multiple administration modes, the ability to track changes over time, and the potential to detect small differences among persons with different types of disability.
4-13. The World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) assesses disability based on the widely accepted *International Classification of Functioning, Disability and Health* (ICF) model of cognition, mobility, self-care, social function, life activities, and participation.
4-14. The Patient-Reported Outcomes Measurement Information System (PROMIS) collects information that can be used to evaluate and monitor physical, mental, and social health using self-report. It has been validated with large, general population samples, and the information it collects is included increasingly in electronic health records.
4-15. The National Institutes of Health (NIH) Toolbox uses a tablet computer to collect a comprehensive array of neurobehavioral measurements of cognitive, emotional, sensory, and motor functions with reference to general and specific population norms. It uses item response theory (IRT) and computer adaptive testing (CAT) methods to facilitate test scoring and reporting.
4-16. Quality of Life in Neurological Disorders (Neuro-QoL) collects psychometrically sound and clinically relevant health-related quality-of-life data for adults and children living with neurological conditions.
4-17. A number of self-report instruments examine interruption of work and quantify functional limitations at work related to worker productivity that are due to health conditions.
4-18. The Work Ability Index (WAI) measures the work ability of a person in an occupational health clinic environment and predicts early retirement, work disability, absence due to sickness, and mortality.
4-19. The Work Productivity and Activity Impairment Questionnaire (WPAI) is a well-validated self-report instrument that assesses the effects on work of various health conditions, including mental health disorders, musculoskeletal pain, and related disorders.
4-20. Most instruments used to measure limitations in work activity due to health conditions assess work function (e.g., presenteeism, absenteeism) among people who are working.

Conclusions

4-1. The most informative evaluations of function may include both specific assessments of body structures and systems and integrated assessments that describe the effects of multiple impairments and comorbid conditions.

4-2. Combining ADLs and IADLs provides a useful means of assessing an individual's ability to perform multiple, integrated functions on a day-to-day basis.

4-3. The validity of ADL and IADL performance assessment is improved by including direct observation, especially in persons with substantial cognitive impairments.

4-4. The collection of direct observations regarding ADLs and IADLs often requires input from trained occupational therapists, physical therapists, speech pathologists, and/or nurses.

4-5. No assessments of ADL or IADL performance correlate strongly with the ability to work for all conditions or impairments. However, an individual who cannot independently perform basic ADLs or IADLs most likely would be unable to work without significant assistance from family members or an attendant.

4-6. Stronger evidence is needed to link ADL and IADL performance to work capacity, perhaps by comparing ADL and IADL performance among applicants who are awarded Social Security Disability Income (SSDI) benefits versus those who are denied.

4-7. The utility of information about ADLs and IADLs in the context of disability determination may be enhanced by asking additional questions about context; environmental factors, including use of assistive technologies; required assistance; and the effect of performing ADLs and IADLs on pain, fatigue, confusion, concentration, and other physical or cognitive factors that can interfere with work performance.

4-8. Evidence-based instruments and sets of instruments that provide integrated information about individuals' overall functional capabilities and limitations could provide helpful information for determinations of work disability.

4-9. Although there is no evidence to support drawing direct inferences from scores on PROMIS, Neuro-QoL, or the NIH Toolbox to employability, scores from these instruments may be very useful in understanding the functioning of an applicant.

4-10. The use of WD-FAB with IRT and CAT methods reduces respondent burden by limiting survey length and can assess functional activity comprehensively and efficiently in 15 to 20 minutes. Currently, WD-FAB may be most useful for understanding self-reported physical function, but direct inferences from WD-FAB to employability are not warranted at this time.

4-11. Although many instruments that measure limitations in work activity due to health conditions assess work function among current workers, they may be of use for previously employed individuals and for those attempting to return to work but continuing to miss time as a result of their health conditions.

REFERENCES

Adler, D. A., T. J. McLaughlin, W. H. Rogers, H. Chang, L. Lapitsky, and D. Lerner. 2006. Job performance deficits due to depression. *American Journal of Psychiatry* 163(9):1569–1576.

Akshoomoff, N., E. Newman, W. K. Thompson, C. McCabe, C. S. Bloss, L. Chang, D. G. Amaral, B. J. Casey, T. M. Ernst, J. A. Frazier, J. R. Gruen, W. E. Kaufmann, T. Kenet, D. N. Kennedy, O. Libiger, S. Mostofsky, S. S. Murray, E. R. Sowell, N. Schork, A. M. Dale, and T. L. Jernigan. 2014. The NIH Toolbox Cognition Battery: Results from a large normative developmental sample (PING). *Neuropsychology* 28(1):1–10.

Amtmann, D., K. F. Cook, K. L. Johnson, and D. Cella. 2011. The PROMIS Initiative: Involvement of rehabilitation stakeholders in development and examples of applications in rehabilitation research. *Archives of Physical Medicine and Rehabilitation* 92(10):S12–S19.

Asami, Y., A. Goren, and Y. Okumura. 2015. Work productivity loss with depression, diagnosed and undiagnosed, among workers in an Internet-based survey conducted in Japan. *Journal of Occupational and Environmental Medicine* 57(1):105–110.

Asher, I. E. 2014. *Asher's occupational therapy assessment tools: An annotated index*, 4th ed. Bethesda, MD: American Occupational Therapy Association.

Baum, C. M., L. T. Connor, T. Morrison, M. Hahn, A. W. Dromerick, and D. F. Edwards. 2008. Reliability, validity, and clinical utility of the Executive Function Performance Test: A measure of executive function in a sample of people with stroke. *American Journal of Occupational Therapy* 62(4):446–455.

Boström, G., M. Conradsson, E. Rosendahl, P. Nordström, Y. Gustafson, and H. Littbrand. 2014. Functional capacity and dependency in transfer and dressing are associated with depressive symptoms in older people. *Clinical Interventions in Aging* 9:249–256. doi: 10.2147/CIA.S57535.

Burnett, J., C. B. Dyer, and A. D. Naik. 2009. Convergent validation of the Kohlman Evaluation of Living Skills (KELS) as a screening tool of older adults' capacity to live safely and independently in the community. *Archives of Physical Medicine and Rehabilitation* 90(11):1948–1952. doi: 10.1016/j.apmr.2009.05.021.

Cancelliere, C., J. Donovan, M. J. Stochkendahl, M. Biscardi, C. Ammendolia, C. Myburgh, and J. D. Cassidy. 2016. Factors affecting return to work after injury or illness: Best evidence synthesis of systematic reviews. *Chiropractic & Manual Therapies* 24:32. doi: 10.1186/s12998-016-0113-z.

Carlozzi, N. E., J. L. Beaumont, D. S. Tulsky, and R. C. Gershon. 2015. The NIH Toolbox Pattern Comparison Processing Speed Test: Normative data. *Archives of Clinical Neuropsychology* 30(5):359–368.

Carlozzi, N. E., E. A. Hahn, S. M. Goodnight, A. L. Kratz, J. S. Paulsen, J. C. Stout, S. Frank, J. A. Miner, D. Cella, R. C. Gershon, S. G. Schilling, and R. E. Ready. 2017a. Patient-reported outcome measures in Huntington disease: Quality of Life in Neurological Disorders (Neuro-QoL) social functioning measures. *Psychological Assessment* 30(4):450–458

Carlozzi, N. E., R. E. Ready, S. Frank, D. Cella, E. A. Hahn, S. M. Goodnight, S. G. Schilling, N. R. Boileau, and P. Dayalu, 2017b. Patient-reported outcomes in Huntington's disease: Quality of Life in Neurological Disorders (Neuro-QoL) and Huntington's Disease Health-Related Quality of Life (HDQLIFE) physical function measures. *Movement Disorders: Official Journal of the Movement Disorder Society* 32(7):1096–1102. doi: 10.1002/mds.27046.

Carlozzi, N. E., D. S. Tulsky, T. J. Wolf, S. Goodnight, R. K. Heaton, K. B. Casaletto, A. W. K. Wong, C. M. Baum, R. C. Gershon, and A. W. Heinemann. 2017c. Construct validity of the NIH Toolbox Cognition Battery in individuals with stroke. *Rehabilitation Psychology* 62(4):443–454.

Cella, D., W. Riley, A. Stone, N. Rothrock, B. Reeve, S. Yount, D. Amtmann, R. Bode, D. Buysse, S. Choi, K. Cook, R. Devellis, D. DeWalt, J. F. Fries, R. Gershon, E. A. Hahn, J. S. Lai, P. Pilkonis, D. Revicki, M. Rose, K. Weinfurt, R. Hays, and PROMIS Cooperative Group. 2010. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology* 63(11):1179–1194.

Cella, D., J. S. Lai, C. J. Nowinski, D. Victorson, A. Peterman, D. Miller, F. Bethoux, A. Heinemann, S. Rubin, J. E. Cavazos, A. T. Reder, R. Sufit, T. Simuni, G. L. Holmes, A. Siderowf, V. Wojna, R. Bode, N. McKinney, T. Podrabsky, K. Wortman, S. Choi, R. Gershon, N. Rothrock, and C. Moy. 2012. Neuro-QOL: Brief measures of health-related quality of life for clinical research in neurology. *Neurology* 78(23):1860–1867.

Chan, L. 2018a. *The WD-FAB: Development and validation testing*. Presentation to the European Union of Medicine in Assurance and Social Security, March 2. https://www.eumass.eu/wp-content/uploads/2018/03/Leighton-Porcino.pdf (accessed April 5, 2019).

Chan, L. 2018b. *The Work Disability Functional Assessment Battery (WD-FAB)*. Presentation to the Committee on Functional Assessment for Adults with Disabilities, Washington, DC, February 26. http://nationalacademies.org/hmd/~/media/Files/Agendas/Activity%20Files/SelectPops/Functional%20Assessment/Meeting%202/5_Chan_WD-FAB.pdf (accessed April 5, 2019).

Chen, C. C., and R. K. Bode. 2010. Psychometric validation of the Manual Ability Measure-36 (MAM-36) in patients with neurologic and musculoskeletal disorders. *Archives of Physical Medicine and Rehabilitation* 91(3):414–420.

Chisholm, D., P. Toto, K. Raina, M. Holm, and J. Rogers. 2014. Evaluating capacity to live independently and safely in the community: Performance assessment of self-care skills. *The British Journal of Occupational Therapy* 77(2):59–63. doi: 10.4276/030802214X13916969447038.

de Zwart, B. C., M. H. Frings-Dresen, and J. C. van Duivenbooden. 2002. Test-retest reliability of the Work Ability Index questionnaire. *Occupational Medicine* 52(4):177–181.

Denboer, J. W., C. Nicholls, C. Corte, and K. Chestnut. 2014. National Institutes of Health Toolbox Cognition Battery. *Archives of Clinical Neuropsychology* 29(7):692–694.

Denis, S., H. S. Shannon, J. Wessel, P. Stratford, and I. Weller. 2007. Association of low back pain, impairment, disability & work limitations in nurses. *Journal of Occupational Rehabilitation* 17(2):213–226.

Dikmen, S. S., P. J. Bauer, S. Weintraub, D. Mungas, J. Slotkin, J. L. Beaumont, R. Gershon, N. R. Temkin, and R. K. Heaton. 2014. Measuring episodic memory across the lifespan: NIH Toolbox Picture Sequence Memory Test. *Journal of the International Neuropsychological Society* 20(6):611–619.

Dutil, E., A. Forget, M. Vanier, and C. Gaudreault. 1990. Development of the ADL Profile: An evaluation for adults with severe head injury. *Occupational Therapy in Health Care* 7(1):7–22.

Feuerstein, M., J. A. Hansen, L. C. Calvio, L. Johnson, and J. G. Ronquillo. 2007. Work productivity in brain tumor survivors. *Journal of Occupational and Environmental Medicine* 49(7):803–811.

Fisher, A. G., and K. Bray Jones. 2010. *Assessment of Motor and Process Skills*, 7th ed., Vols. 1 and 2. Fort Collins, CO: Three Star Press.

Gardarsdóttir, S., and S. Kaplan. 2002. Validity of the Arnadóttir OT-ADL Neurobehavioral Evaluation (A-ONE): Performance in activities of daily living and neurobehavioral impairments of persons with left and right hemisphere damage. *American Journal of Occupational Therapy* 56(5):499–508.

Gardner, B. T., A. M. Dale, S. Buckner-Petty, L. Van Dillen, B. C. Amick III, and B. Evanoff. 2016. Comparison of employer productivity metrics to lost productivity estimated by commonly used questionnaires. *Journal of Occupational and Environmental Medicine* 58(2):170–177.

Gershon, R. C., J. S. Lai, R. Bode, S. Choi, C. Moy, T. Bleck, D. Miller, A. Peterman, and D. Cella. 2012. Neuro-QOL: Quality of life item banks for adults with neurological disorders: Item development and calibrations based upon clinical and general population testing. *Quality of Life Research* 21(3):475–486.

Gershon, R. C., K. F. Cook, D. Mungas, J. J. Manly, J. Slotkin, J. L. Beaumont, and S. Weintraub. 2014. Language measures of the NIH Toolbox Cognition Battery. *Journal of the International Neuropsychological Society* 20(6):642–651.

Gold, D. A. 2012. An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology* 34(1):11–34.

Graf, C., and Hartford Institute for Geriatric Nursing. 2008. The Lawton Instrumental Activities of Daily Living (IADL) scale. *MEDSURG Nursing* 17(5):343–344.

Gwaltney, C. J., A. L. Shields, and S. Shiffman. 2008. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: A meta-analytic review. *Value in Health* 11(2):322–333.

Hahn, E. A., R. F. DeVellis, R. K. Bode, S. F. Garcia, L. D. Castel, S. V. Eisen, H. B. Bosworth, A. W. Heinemann, N. Rothrock, D. Cella, and PROMIS Cooperative Group. 2010. Measuring social health in the Patient-Reported Outcomes Measurement Information System (PROMIS): Item bank development and testing. *Quality of Life Research* 19(7):1035–1044.

Hall, J. R., H. T. Vo, R. C. Barber, L. A. Johnson, and S. E. O'Bryant. 2011. The link between cognitive measures and ADLs and IADL functioning in mild Alzheimer's: What has gender got to do with it? *International Journal of Alzheimer's Disease* 2011: Article ID 276734. doi: 10.4061/2011/276734.

Harvard Medical School. 2005. *World Health Organization Health and Work Performance Questionnaire (HPQ)*. https://www.hcp.med.harvard.edu/hpq (accessed December 6, 2018).

Haslam, J., G. Pépin, R. Bourbonnais, and S. Grignon. 2010. Processes of task performance as measured by the Assessment of Motor and Process Skills (AMPS): A predictor of work-related outcomes for adults with schizophrenia? *Work* 37(1):53–64.

Healthy Workplaces. n.d. *Work Ability Index*. https://healthy-workplaces.eu/previous/all-ages-2016/de/tools-and-publications/practical-tools/work-ability-index (accessed April 5, 2019).

HHS (U.S. Department of Health and Human Services). 2018. *Patient-Reported Outcomes Measurement Information System: Program snapshot*. https://commonfund.nih.gov/promis/index (accessed April 5, 2019).

Houston, D., S. L. Williams, J. Bloomer, and W. C. Mann. 1989. The Bay Area Functional Performance Evaluation: Development and standardization. *American Journal of Occupational Therapy* 43(3):170–183.

Ilmarinen, J., K. Tuomi, and M. Klockars. 1997. Changes in the work ability of active employees over an 11-year period. *Scandinavian Journal of Work, Environment & Health* 23(Suppl. 1):49–57.

Jefferson, A. L., L. K. Byerly, S. Vanderhill, S. Lambe, S. Wong, A. Ozonoff, and J. H. Karlawish. 2008. Characterization of activities of daily living in individuals with mild cognitive impairment. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry* 16(5):375–383.

Katz, S. 1983. Assessing self-maintenance: Activities of daily living, mobility, and instrumental activities of daily living. *Journal of the American Geriatrics Association* 31(12):721–727.

Katz, S., and C. A. Akpom. 1976. 12. Index of ADL. *Medical Care* 14(5 Suppl.):116–118.

Kessler, R. C., C. Barber, A. Beck, P. Berglund, P. D. Cleary, D. McKenas, N. Pronk, G. Simon, P. Stang, T. B. Üstün, and P. Wang. 2003. The World Health Organization Health and Work Performance Questionnaire (HPQ). *Journal of Occupational and Environmental Medicine* 45(2):156–174.

Kessler, R. C., M. Ames, P. A. Hymel, R. Loeppke, D. K. McKenas, D. E. Richling, P. E. Stang, and T. B. Üstün. 2004. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *Journal of Occupational and Environmental Medicine* 46(6 Suppl.):S23–S37.

Kessler, R. C., V. Shahly, P. E. Stang, and M. C. Lane. 2010. The associations of migraines and other headaches with work performance: Results from the National Comorbidity Survey-Replication (NCS-R). *Cephalalgia* 30(6):722–734.

Kohlman-Thomson, L. 1992. *Kohlman evaluation of living skills*, 3rd ed. Bethesda, MD: American Occupational Therapy Association.

Lerner, D., B. C. Amick III, W. H. Rogers, S. Malspeis, K. Bungay, and D. Cynn. 2001. The Work Limitations Questionnaire. *Medical Care* 39(1):72–85.

Lerner, D., J. I. Reed, E. Massarotti, L. M. Wester, and T. A. Burke. 2002. The Work Limitations Questionnaire's validity and reliability among patients with osteoarthritis. *Journal of Clinical Epidemiology* 55(2):197–208.

Lerner, D., B. C. Amick III, J. C. Lee, T. Rooney, W. H. Rogers, H. Chang, and E. R. Berndt. 2003. Relationship of employee-reported work limitations to work productivity. *Medical Care* 41(5):649–659.

Loring, D. W., S. C. Bowden, E. Staikova, J. A. Bishop, D. L. Drane, and F. C. Goldstein. 2018. NIH Toolbox Picture Sequence Memory Test for assessing clinical memory function: Diagnostic relationship to the Rey Auditory Verbal Learning Test. *Archives of Clinical Neuropsychology* 34(2):268–276.

Marfeo, E. E., S. M. Haley, A. M. Jette, S. V. Eisen, P. Ni, K. Bogusz, M. Meterko, C. M. McDonough, L. Chan, D. E. Brandt, and E. K. Rasch. 2013a. Conceptual foundation for measures of physical function and behavioral health function for Social Security work disability evaluation. *Archives of Physical Medicine and Rehabilitation* 94(9):1645–1652.

Marfeo, E. E., P. Ni, S. M. Haley, K. Bogusz, M. Meterko, C. M. McDonough, L. Chan, E. K. Rasch, D. E. Brandt, and A. M. Jette. 2013b. Scale refinement and initial evaluation of a behavioral health function measurement tool for work disability evaluation. *Archives of Physical Medicine and Rehabilitation* 94(9):1679–1686.

Marfeo, E. E., P. Ni, S. M. Haley, A. M. Jette, K. Bogusz, M. Meterko, C. M. McDonough, L. Chan, D. E. Brandt, and E. K. Rasch. 2013c. Development of an instrument to measure behavioral health function for work disability: Item pool construction and factor analysis. *Archives of Physical Medicine and Rehabilitation* 94(9):1670–1678.

Marfeo, E. E., P. S. Ni, L. Chan, E. K. Rasch, and A. M. Jette. 2014. Combining agreement and frequency rating scales to optimize psychometrics in measuring behavioral health functioning. *Journal of Clinical Epidemiology* 67(7):781–784.

Marfeo, E. E., P. Ni, L. Chan, E. K. Rasch, C. M. McDonough, D. E. Brandt, K. Bogusz, and A. M. Jette. 2015. Interpreting physical and behavioral health scores from new work disability instruments. *Journal of Rehabilitation Medicine* 47(5):394–402.

Marfeo, E. E., P. Ni, C. McDonough, K. Peterik, M. Marino, M. Meterko, E. K. Rasch, L. Chan, D. Brandt, and A. M. Jette. 2018. Improving assessment of work related mental health function using the Work Disability Functional Assessment Battery (WD-FAB). *Journal of Occupational Rehabilitation* 28(1):190–199. doi: 10.1007/s10926-017-9710-5.

Margaret Reilly Associates. 2013. *WPAI general information*. http://www.reillyassociates.net/WPAI_General.html (accessed April 5, 2019).

Marino, M. E., M. Meterko, E. E. Marfeo, C. M. McDonough, A. M. Jette, P. Ni, K. Bogusz, E. K. Rasch, D. E. Brandt, and L. Chan. 2015. Work-related measures of physical and behavioral health function: Test-retest reliability. *Disability and Health Journal* 8(4):652–657.

McDonough, C. M., A. M. Jette, P. Ni, K. Bogusz, E. E. Marfeo, D. E. Brandt, L. Chan, M. Meterko, S. M. Haley, and E. K. Rasch. 2013. Development of a self-report physical function instrument for disability assessment: Item pool construction and factor analysis. *Archives of Physical Medicine and Rehabilitation* 94(9):1653–1660.

McDonough, C. M., P. Ni, K. Peterik, E. E. Marfeo, M. E. Marino, M. Meterko, E. K. Rasch, D. E. Brandt, A. M. Jette, and L. Chan. 2017. Improving measures of work-related physical functioning. *Quality of Life Research* 26(3):789–798.

McDonough, C. M., P. Ni, K. Peterik, J. D. Hershberg, L. R. Bell, L. Chan, D. E. Brandt, and A. M. Jette. 2018. Validation of the work-disability physical functional assessment battery. *Archives of Physical Medicine and Rehabilitation* 99(9):1798–1804.

Meltzer, H., P. Bebbington, T. Brugha, S. McManus, D. Rai, M. S. Dennis, and R. Jenkins. 2012. Physical ill health, disability, dependence and depression: Results from the 2007 National Survey of Psychiatric Morbidity Among Adults in England. *Disability and Health Journal* 5(2):102–110.

Meterko, M., E. E. Marfeo, C. M. McDonough, A. M. Jette, P. Ni, K. Bogusz, E. K. Rasch, D. E. Brandt, and L. Chan. 2015. The Work Disability Functional Assessment Battery (WD-FAB): Feasibility and psychometric properties. *Archives of Physical Medicine and Rehabilitation* 96(6):1028–1035.

Meterko, M., M. Marino, P. Ni, E. Marfeo, C. M. McDonough, A. Jette, K. Peterik, E. Rasch, D. E. Brandt, and L. Chan. 2018. Psychometric evaluation of the improved Work-Disability Functional Assessment Battery. *Archives of Physical Medicine and Rehabilitation*. In press. Corrected proof available online December 19, 2018. doi: 10.1016/j.apmr.2018.09.125.

Mlinac, M. E., and M. C. Feng. 2016. Assessment of activities of daily living, self-care, and independence. *Archives of Clinical Neuropsychology* 31(6):506–516.

Morrison, M. T., G. M. Giles, J. D. Ryan, C. M. Baum, A. W. Dromerick, H. J. Polatajko, and D. F. Edwards. 2013. Multiple Errands Test–Revised (MET–R): A performance-based measure of executive function in people with mild cerebrovascular accident. *The American Journal of Occupational Therapy* 67(4):460–468. doi: 10.5014/ajot.2013.007880.

Munir, F., J. Yarker, C. Haslam, H. Long, S. Leka, A. Griffiths, and S. Cox. 2007. Work factors related to psychological and health-related distress among employees with chronic illnesses. *Journal of Occupational Rehabilitation* 17(2):259–277.

Neuro-QoL investigators. 2015. *Neuro-QoL technical report: Development and initial validation of patient-reported item banks for use in neurological research and practice*. http://www.healthmeasures.net/images/neuro_qol/Neuro-QoL_Manual_Technical_Report_v2_24Mar2015.pdf (accessed April 24, 2019).

Ni, P., C. M. McDonough, A. M. Jette, K. Bogusz, E. E. Marfeo, E. K. Rasch, D. E. Brandt, M. Meterko, S. M. Haley, and L. Chan. 2013. Development of a computer-adaptive physical function instrument for Social Security Administration disability determination. *Archives of Physical Medicine and Rehabilitation* 94(9):1661–1669.

NIOSH (National Institute for Occupational Safety and Health). 2008. Essential elements of effective workplace programs and policies for improving worker health and wellbeing. *WorkLife*, October. https://www.cdc.gov/niosh/docs/2010-140/pdfs/2010-140.pdf (accessed April 5, 2019).

Nowinski, C. J., D. Victorson, J. E. Cavazos, R. Gershon, and D. Cella. 2010. Neuro-QOL and the NIH Toolbox: Implications for epilepsy. *Therapy* 7(5):533–540.

NU (Northwestern University). 2018a. *Intro to Neuro-QoL*. http://www.healthmeasures.net/explore-measurement-systems/neuro-qol/intro-to-neuro-qol (accessed April 5, 2019).

NU. 2018b. *Intro to NIH toolbox*. http://www.healthmeasures.net/explore-measurement-systems/nih-toolbox/intro-to-nih-toolbox (accessed April 5, 2019).

NU. 2018c. *Neuro-QoL: Measure development & research*. http://www.healthmeasures.net/explore-measurement-systems/neuro-qol/measure-development-research (accessed April 4, 2019).

NU. 2018d. *Why use NIH toolbox?* http://www.healthmeasures.net/explore-measurement-systems/nih-toolbox (accessed April 5, 2019).

NU. 2018e. *Why use PROMIS?* http://www.healthmeasures.net/explore-measurement-systems/promis (accessed April 5, 2019).

NU. 2018f. *Why use Neuro-QoL?* http://www.healthmeasures.net/explore-measurement-systems/neuro-qol (accessed April 5, 2019).

Ottenbacher, K. J., Y. Hsu, C. V. Granger, and R. C. Fiedler. 1996. The reliability of the Functional Independence Measure: A quantitative review. *Archives of Physical Medicine and Rehabilitation* 77(12):1226–1232.

Patterson, M. B., and J. L. Mack. 2001. The Cleveland Scale for Activities of Daily Living (CSADL): Its reliability and validity. *Journal of Clinical Geropsychology* 7(1):15–28.

Pronk, N. P., D. L. McLellan, M. P. McGrail, S. M. Olson, Z. J. McKinney, J. N. Katz, G. R. Wagner, and G. Sorensen. 2016. Measurement tools for integrated worker health protection and promotion: Lessons learned from the SafeWell Project. *Journal of Occupational and Environmental Medicine* 58(7):651–658.

Quinn, T. J., P. Langhorne, and D. J. Stott. 2011. Barthel Index for stroke trials. *Stroke* 42(4):1146–1151.

Reilly, M. C., A. S. Zbrozek, and E. M. Dukes. 1993. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics* 4(5):353–365.

Reilly, M. C., K. L. Gooch, R. L. Wong, H. Kupper, and D. van der Heijde. 2010. Validity, reliability and responsiveness of the Work Productivity and Activity Impairment Questionnaire in ankylosing spondylitis. *Rheumatology (Oxford)* 49(4):812–819.

Reiman, M. P., and R. C. Manske. 2011. The assessment of function: How is it measured? A clinical perspective. *Journal of Manual and Manipulative Therapy* 19(2):91–99.

Royall, D. R., E. C. Lauterbach, D. Kaufer, P. Malloy, K. L. Coburn, and K. J. Black. 2007. The cognitive correlates of functional status: A review from the Committee on Research of the American Neuropsychiatric Association. *The Journal of Neuropsychiatry and Clinical Neurosciences* 19(3):249–265.

Schuling, J., R. de Haan, M. Limburg, and K. H. Groenier. 1993. The Frenchay Activities Index: Assessment of functional status in stroke patients. *Stroke* 24(8):1173–1177.

Scuffham, P. A., N. Vecchio, and H. A. Whiteford. 2014. Exploring the validity of HPQ-based presenteeism measures to estimate productivity losses in the health and education sectors. *Medical Decision Making* 34(1):127–137.

Sevak, P., D. C. Stapleton, and J. O'Neill. 2017. How individual and environmental factors affect employment outcomes. *Journal of Vocational Rehabilitation* 46(2):117–120.

Sikkes, S. A., E. S. de Lange-de Klerk, Y. A. Pijnenburg, P. Scheltens, and B. M. Uitdehaag. 2009. A systematic review of instrumental activities of daily living scales in dementia: Room for improvement. *Journal of Neurology, Neurosurgery & Psychiatry* 80(1):7–12.

Society of Occupational Medicine. 2018. *The Work Ability Index (WAI)*. https://academic.oup.com/occmed/article/57/2/160/1584972 (accessed April 5, 2019).

Sorensen, G., D. L. McLellan, E. L. Sabbath, J. T. Dennerlein, E. M. Nagler, D. A. Hurtado, N. P. Pronk, and G. R. Wagner. 2016. Integrating worksite health protection and health promotion: A conceptual model for intervention and research. *Preventive Medicine* 91:188–196. doi: 10.1016/j.ypmed.2016.08.005.

Sorensen, G., E. Sparer, J. A. R. Williams, D. Gundersen, L. I. Boden, J. T. Dennerlein, D. Hashimoto, J. N. Katz, D. L. McLellan, C. A. Okechukwu, N. P. Pronk, A. Revette, and G. R. Wagner. 2018. Measuring best practices for workplace safety, health, and well-being: The workplace integrated safety and health assessment. *Journal of Occupational and Environmental Medicine* 60(5):430–439.

Suzuki, T., K. Miyaki, Y. Song, A. Tsutsumi, N. Kawakami, A. Shimazu, M. Takahashi, A. Inoue, and S. Kurioka. 2015. Relationship between sickness presenteeism (WHO-HPQ) with depression and sickness absence due to mental disease in a cohort of Japanese workers. *Journal of Affective Disorders* 180:14–20. doi: 10.1016/j.jad.2015.03.034.

Tufts Medical Center. 2018. *Available questionnaires*. https://www.tuftsmedicalcenter.org/Research-Clinical-Trials/Institutes-Centers-Labs/Center-for-Health-Solutions/Available-Questionnaires.aspx (accessed April 5, 2019).

Tulsky, D. S., N. Carlozzi, N. D. Chiaravalloti, J. L. Beaumont, P. A. Kisala, D. Mungas, K. Conway, and R. Gershon. 2014. NIH Toolbox Cognition Battery (NIHTB-CB): List sorting test to measure working memory. *Journal of the International Neuropsychological Society* 20(6):599–610.

Tulsky, D. S., P. A. Kisala, D. Victorson, D. G. Tate, A. W. Heinemann, S. Charlifue, S. C. Kirshblum, D. Fyffe, R. Gershon, A. M. Spungen, and C. H. Bombardier. 2015. Overview of the Spinal Cord Injury–Quality of Life (SCI-QOL) measurement system. *The Journal of Spinal Cord Medicine* 38(3):257–269.

Tulsky, D. S., P. A. Kisala, D. Victorson, N. Carlozzi, T. Bushnik, M. Sherer, S. W. Choi, A. W. Heinemann, N. Chiaravalloti, A. M. Sander, and J. Englander. 2016. TBI-QOL: Development and calibration of item banks to measure patient reported outcomes following traumatic brain injury. *The Journal of Head Trauma Rehabilitation* 31(1):40–51. doi: 10.1097/HTR.0000000000000131.

Tuomi, K., J. Ilmarinen, L. Eskelinen, E. Järvinen, J. Toikkanen, and M. Klockars. 1991. Prevalence and incidence rates of diseases and work ability in different work categories of municipal occupations. *Scandinavian Journal of Work, Environment & Health* 67–74.

Tuomi, K., J. Ilmarinen, R. Martikainen, L. Aalto, and M. Klockars. 1997. Aging, work, life-style and work ability among Finnish municipal workers in 1981–1992. *Scandinavian Journal of Work, Environment & Health* 17(Suppl. 1):58–65.

Tuomi, K., P. Huuhtanen, E. Nykyri, and J. Ilmarinen. 2001. Promotion of work ability, the quality of work and retirement. *Occupational Medicine* 51(5):318–324.

Üstün, T. B., S. Chatterji, N. Kostanjsek, J. Rehm, C. Kennedy, J. Epping-Jordan, S. Saxena, M. von Korff, and C. Pull. 2010. Developing the World Health Organization Disability Assessment Schedule 2.0. *Bulletin of the World Health Organization* 88:815–823.

Victorson, D., J. E. Cavazos, G. L. Holmes, A. T. Reder, V. Wojna, C. Nowinski, D. Miller, S. Buono, A. Mueller, C. Moy, and D. Cella. 2014. Validity of the Neurology Quality-of-Life (Neuro-QoL) measurement system in adult epilepsy. *Epilepsy & Behavior* 31:77–84.

Weintraub, S., S. S. Dikmen, R. K. Heaton, D. S. Tulsky, P. D. Zelazo, P. J. Bauer, N. E. Carlozzi, J. Slotkin, D. Blitz, K. Wallner-Allen, N. A. Fox, J. L. Beaumont, D. Mungas, C. J. Nowinski, J. Richler, J. A. Deocampo, J. E. Anderson, J. J. Manly, B. Borosh, R. Havlik, K. Conway, E. Edwards, L. Freund, J. W. King, C. Moy, E. Witt, and R. C. Gershon. 2013. Cognition assessment using the NIH Toolbox. *Neurology* 80(11 Suppl. 3):S54–S64.

WHO (World Health Organization). 2018. *WHO Disability Assessment Schedule 2.0*. http://www.who.int/classifications/icf/whodasii/en (accessed April 5, 2019).

5

Selected Instruments for Assessment of Physical Functional Abilities Relevant to Work Requirements

This chapter reviews instruments available for measuring physical functional abilities relevant to work requirements. In terms of the conceptual framework described in Chapter 2 (see Figure 2-3), the instruments discussed assess function in terms of ability to perform physical activities (e.g., assessments of musculoskeletal function) as well as body functions (e.g., exercise testing for cardiac function, vision and hearing testing).

The inclusion criteria for instruments reviewed in this chapter were (1) sufficient representation in the scientific literature and/or widespread use; (2) evidence of sound psychometric properties, including (when applicable) construct validity, internal consistency, sensitivity to change, test-retest reliability, and intra- and interrater agreement (including subject/proxy and telephone/in-person administration) (see Chapter 3); (3) normative data; (4) applicability across a range of conditions and functional levels; (5) availability in the public domain; (6) ease of administration; (7) brevity; (8) availability in multiple languages; (9) validation in subpopulations; (10) multiple administration formats (telephone interview versus in-person administration; self-report versus proxy respondent); and (11) availability of alternative forms to minimize the risk of practice effects for performance measures. Some of the instruments reviewed here do not meet all of these criteria, but they are included because they illustrate potential assessment instruments.

The discussion of these instruments begins with an overview of physical functional abilities relevant to work requirements and the two broad types of instruments that can be used to assess them. The review of specific instruments that follows begins with those used for general assessment of physical function. The chapter then reviews in turn selected instruments used to measure musculoskeletal function, pain, visual function, hearing function, and speech and language function. A series of annex tables at the end of this chapter provides information on selected functional assessment tools for physical abilities, including qualifications to administer, how to administer, time to administer, psychometric properties, proprietary considerations, and the populations to which the tools apply.

OVERVIEW

Physical Functional Abilities Relevant to Work Requirements

Physical functional abilities relevant to work requirements include 24 physical activities performed by workers in carrying out critical tasks, which are grouped into 10 categories (DOL, 2018, p. 88) (see Annex Table 5-1). For example, driving is a critical task that involves a number of physical demands, including far visual acuity, peripheral vision, and gross manipulation. In operating a passenger vehicle with automatic transmission and power brakes, driving also includes manipulation (foot/leg controls) (DOL, 2018, p. 123). In addition, several categories of worker abilities (physical, psychomotor, and sensory) (DOL, n.d.) are pertinent to the discussion of physical functions relevant to work requirements. Annex Table 5-2 shows how the functional domains identified by the committee, which correspond to the organization of this chapter, map to the physical demand data elements from the Occupational Requirements Survey (ORS) and to relevant physical, psychomotor, and sensory abilities identified in the Occupational Information Network (O*NET) (see Chapter 2), as well as to the equivalent functional abilities listed in different forms used by the U.S. Social Security Administration (SSA) in the collection and assessment of information relevant to an applicant's capacity to perform work-related activities.

Instruments for Measuring Physical Function

Instruments for measuring physical function fall into two broad categories: self-report questionnaires and performance-based measures. Both are available for the assessment of work ability (or inability) in individuals with musculoskeletal disorders (Wind et al., 2005), as well as those with impaired cardiovascular and/or cardiopulmonary function, vision, and communication (hearing, speech-language).

Self-Report Questionnaires

Self-report questionnaires may be self-administered, administered by a medical provider, or completed by a proxy. Self-administered questionnaires (self-report measures) are commonly used to have individuals assess their pain and function and have been used to measure or approximate individuals' functional ability. Dedicated self-report functional measures exist for nearly all body regions and conditions. Many self-report instruments have been carefully evaluated for reliability, responsiveness, and validity, and much work has been dedicated to the development of these tools (Reiman and Manske, 2011). Proxies may be used to complete these questionnaires when individuals are unable to provide their own information because of their physical or mental condition. (See Chapter 3 for more information on the strengths and limitations of using a proxy or third-party respondents.)

There are four basic types of self-report instruments: (1) integrated, (2) impairment specific, (3) body-part or region specific, and (4) patient specific. Each type of instrument usually has a unique purpose, and each has advantages and disadvantages that have bearing on its potential utility. Impairment-specific instruments may not capture the additive or multiplicative effects of multiple impairments or comorbidities on an individual's ability to function. Body-part– and region-specific instruments are generally used to address only the injury mechanisms and disease states affecting that body part or region. If a patient-specific instrument is used, generalization across patients may be misleading (Martin and Irrgang, 2007; Westaway et al., 1998). In contrast, an example of an integrated assessment instrument is the Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) (Tarlov et al., 1989), which is a widely used and accepted self-report instrument with evidence to support its use for diverse pathological conditions. It is not a tool specifically for functional assessment of physical abilities, but combines assessment of physical and mental or emotional symptoms. As such, interpreting its results with respect to specific musculoskeletal or cardiovascular impairments may prove challenging. However, there is evidence that combining condition-specific outcome measures such as the Owestry Disability Index (ODI) or the Roland Morris Disability Questionnaire (RMDQ) with the SF-36 provides complementary information on the person's status (Ko and Chae, 2017). Given that this chapter focuses on musculoskeletal, cardiovascular, vision, and communication impairments, the instruments discussed here focus primarily on specific body parts or organ systems.

Performance-Based Measures

Functional performance testing is defined as using a variety of physical skills and tests to determine an individual's ability to participate at the desired level in an occupation or to return to work in a safe and timely manner without functional limitations (Reiman and Manske, 2011). Assessment at the functional level thus looks at the functioning of the person as a whole rather than function of a part of the person (Reiman and Manske, 2011).

A comprehensive functional assessment would make it possible to evaluate an individual's ability to put together a series of movements (rather than perform isolated single-joint and planar movements) toward the safe and efficient completion of a task (Reiman and Manske, 2011). Currently, however, there is no single measure of an individual's overall physical functional ability. The assessment tools summarized below and reviewed in detail in the sections that follow, used in combination, can provide only approximations of function. Yet, such a combination of reliable and valid measures of different constructs may improve the ability to predict work participation (Kuijer et al., 2012) and at present will provide the best determination of an individual's current status in this regard (Reiman and Manske, 2011).

Performance-based assessments are "commonly used to determine the physical work abilities of individuals who have sustained musculoskeletal injury" (Gross and Battié, 2004) and/or cardiovascular impairment. A broad range of physical performance-based measures exist, such as trunk endurance testing; tests of movement patterns; excursion reach testing; jumping tests; hopping tests; strength testing; power testing; aerobic endurance testing in multiple planes of movement; lifting tests; balance/proprioceptive testing in multiple planes of movement; and speed, agility, and quickness testing (Reiman and Manske, 2011). Another example is the Purdue Pegboard Test, which is designed to measure uni- and bimanual finger and hand dexterity. It is a test of manipulative dexterity consisting of four subtests: right hand, left hand, both hands, and assembly (Mathiowetz et al., 1986; Tiffin and Asher, 1948). Other performance-based tests include the 3-minute step test, 15-minute stand test, floor-to-waist lift, 1-minute crouch, 2-minute kneel, 5-minute rotation, stepladder/stairs, waist-to-overhead lift, crawling, handgrip, hand coordination, stooping, and bending (Kuijer et al., 2012) and the Functional Gait Assessment. Finally, the Isernhagen Work System Functional Capacity Evaluation (Kuijer et al., 2012) is a performance-based measure that can be used to assess function at multiple parts of the body and is frequently used in the occupational setting. Physical performance-based measures must be administered by specialized personnel, typically a physical or occupational therapist, which increases their cost and limits their availability. Some performance-based measures have been found to be predictive of work participation for individuals with chronic musculoskeletal disorders, irrespective of whether the condition involves complaints of the upper extremities, lower extremities, or low back (Kuijer et al., 2012).

Performance-based measures provide relatively objective assessments of physical functioning. However, their results are subject to several confounders, including age, gender, education, pain duration, pain intensity, pain-related disability, employment status, physical work demand level, and work organizational policies and practices (Kuijer et al., 2012), although evidence suggests that even such potential confounders as pain intensity, work-related recovery expectations, and organizational policies and practices do not diminish the predictive validity of performance-based measures with respect to work participation (Kuijer et al., 2012). Other factors to be considered include the individual's psychological and social status at the time the test is administered (Reiman and Manske, 2011), which has been found to affect functional ability. Motivation may affect the results of these tests (Touré-Tillery and Fishbach, 2014) and age and level of depression also can predict work participation (Vowles et al., 2004). Assessment of an individual's psychological status may therefore help in better determining his or her functional status and work readiness. Understanding the relationship among these different variables is a complex task, one that will most likely require interaction among multiple disciplines (Reiman and Manske, 2011).

Performance-based measures also are sensitive to change (Pepin et al., 2004) in that results may vary among testers and even with the same tester on different days. Nonetheless, some performance-based measures have been shown to be strong predictors of certain outcomes, such as mortality (Goldman et al., 2014), falls, institutionalization, and other causes of utilization of health services (Curb et al., 2006). In short, performance-based measures have been shown to be instructive, but application of their results to an individual's work ability should be interpreted with caution.

Comparison of Self-Report and Performance-Based Measures

Self-report measures, frequently used to assess functional ability, may be less time-consuming to administer than performance-based assessments. More important, however, is that the two approaches differ considerably in the information they provide (Bean et al., 2011; Sager et al., 1992). Self-report measures provide an individual's perspective on his or her change in function (Nielsen et al., 2016), which may help shed light on the person's (perceived) ability to perform work beyond that which range-of-motion assessments might provide, for example. Performance measures involve quantification of output, such as the ability to lift a specific amount of weight, visual acuity, hearing capacity, and the like. In addition, self-report measures are dramatically influenced by an individuals' perception of pain, which can in turn affect their perception of their functional ability (Reiman and Manske, 2011, pp. 101–107). In general, physical performance measures/tests can add important information to that obtained with self-report questionnaires, and it is best to use the two together in assessing an individual's functional ability and gathering information relevant to determining his or her ability to sustain work on a regular and ongoing basis (Kuijer et al., 2012; Reiman and Manske, 2011).

Finally, as noted earlier, although self-report and performance-based physical measures of individual parts of the human body provide important information, they do not serve to measure an individual's overall physical functional ability. Thus, they may not, individually or collectively, measure function adequately to predict an individual's ability to work. Indeed, given the complexity of measuring function and the multidimensional nature of work participation (WHO, 2001), one cannot expect a single instrument to allow for a complete assessment (Kuijer et al., 2012), and it is important to be aware of the range of physical assessment measures that are currently available, how to interpret them, and how much weight to assign them. Having said that, a self-report measure that appears to be a useful indicator of physical activity levels in people with chronic pain, arthritis, renal failure, and various neurological and cardiorespiratory conditions, as well as in healthy "older" people, is the Human Activity Profile (HAP) (Davidson and de Morton, 2007), which may serve as an example of an exception to this generalization. The HAP is a self-measure of energy expenditure or physical fitness (Davidson and de Morton, 2007) and has been found useful in estimating fitness level when standard exercise testing is not feasible (Bilek et al., 2008).

The following sections describe some of the most commonly used self-report and performance-based instruments for assessing physical function relevant to work participation. General assessments used to approximate function for the full body are described first, followed by those used to assess musculoskeletal function, pain, cardiovascular function, visual function, hearing function, and speech and language function. Data on the psychometric properties of these instruments are presented in Annex Table 5-3, along with information on respondent and administrative burden and any cultural and language adaptations. Note that the instruments discussed are those used frequently in clinical practice, and should not be regarded as an exhaustive list.

GENERAL ASSESSMENTS OF PHYSICAL FUNCTION

Instruments used for general assessment of physical function include functional capacity evaluations (FCEs) and the physical function scales of the Work Disability Functional Assessment Battery (WD-FAB). (See Annex Table 5-3 for a selected listing of these instruments.)

Functional Capacity Evaluations

An FCE, also termed a functional capacity assessment, physical capacity evaluation, or work capacity evaluation (Genovese and Galper, 2009), is a commonly used performance-based measure of physical function. FCEs have been described as systematic, comprehensive, and objective measures of an individual's maximum physical ability to perform tasks involved in activities of daily living and occupational activities (Jahn et al., 2004). They are often used to determine an individual's ability to return to work or to identify appropriate work modifications that would make it possible to return to work at the level of full duty or modified/restricted duty, depending on the situation, and have demonstrated predictive value for this purpose (Kuijer et al., 2012; Soer et al., 2008). Other uses of FCEs include tailoring required occupational tasks to the functional deficits of an individual, screening for physically demanding occupations, and evaluating functional progress in a rehabilitation therapy program.

Unlike most other physical performance-based measures, FCEs can be used to assess more than one musculoskeletal impairment or muscle group; thus they can potentially provide more information than those performance-based measures that focus on specific body regions/parts or conditions. In addition, since FCEs are designed to evaluate an individual's capacity to perform work activities related to his or her participation in employment (Soer et al., 2008), such factors as diagnosed impairments and functional difficulties, as well as job task requirements, are taken into account. Administered by a physical or occupational therapist, an FCE can help determine not only an individual's maximum capacities and ability to perform work-related activities during a designated time period, but also, potentially, an individual's level of effort. Ideally, the assessments are conducted in a standardized and reproducible manner (Genovese and Galper, 2009). As discussed later, FCEs are not without limitations, including with respect to both validity and reliability, as the results can be interrater dependent and affected by a myriad of confounders, including willingness to return to work (Ansuategui Echeita et al., 2018; Gross, 2006; Oesch et al., 2012).

An individual's performance during an FCE may improve following a physical rehabilitation treatment program (Fore et al., 2015), as demonstrated by comparing pre- and postprogram performance. Thus, an FCE may be administered in an occupational medicine setting before and after an individual has undergone a period of "work hardening" or "work conditioning," an expensive and time-consuming process carried out by physical or occupational therapists. During this process, job tasks the individual is expected to perform during a typical workday are simulated, and the intensity of the simulated work is increased in a stepwise manner until it approximates the job to which the individual will return. The individual is evaluated doing these tasks. Once individuals are able to perform the essential functions of their job tasks, they are released to return to work. This process is often carried out in the occupational setting if resources are available, especially if return to work appears likely.

There are approximately 10 commonly utilized FCE instruments (Chen, 2007): Blankenship, ERGOS Work Simulator, Ergo-Kit variation, WorkWell

Systems (formerly known as Isernhagen Work Systems), Hanoun Medical, Physical Work Performance Evaluation (Ergoscience), WEST-EPIC, Key, ERGOS, and ARCON. Evidence is limited as to the reliability and validity of the various FCE instruments. Gouttebarge and colleagues (2004) reviewed studies of the reliability and validity of four instruments. They found that the interrater reliability and predictive validity of Isernhagen Work Systems were good, while the procedures used to examine intrarater reliability (test-retest) were not sufficiently rigorous to allow any conclusions. The studies reviewed did not demonstrate the concurrent validity (i.e., the accuracy of the evaluation, whether the test measures what it intends to measure) of the ERGOS Work Simulator and Ergo-Kit, and the authors found no study on their reliability or on the reliability and validity of the Blankenship System.

Rustenburg and colleagues (2004) measured the concurrent validity of the ERGOS Work Simulator and the Ergo-Kit with respect to maximal lifting capacity. In this study, 25 male firefighters were subjected to tests designed to evaluate upper- and lower-extremity lifting capacities. The authors found the concurrent validity of the ERGOS Work Simulator and the Ergo-Kit to be poor for dynamic lifting.

De Baets and colleagues (2017) evaluated the evidence for the reliability and validity of multiple FCE instruments. The Baltimore Therapeutic Equipment work simulator showed moderate predictive validity, while Ergo-Kit showed high inter- and intrarater reliability and high convergent validity. However, the concurrent validity of Ergo-Kit and the ERGOS Work Simulator varied from low to moderate. Moderate to high test-retest and inter- and intrarater reliability was found for the Isernhagen Work Systems FCE, although the predictive validity of the instrument was low. The Physical Work Performance Evaluation showed moderate test-retest reliability and moderate to high interrater reliability, while its predictive validity was high (De Baets et al., 2017).

The WorkWell Systems FCE (formerly Isernhagen Work Systems) consists of 29 items related to five work performance categories (weight handling and strength, posture and mobility, locomotion, balance, and upper-extremity coordination). A systematic review of the reliability of this FCE showed an acceptable level of reliability for 96 percent of the test-retest reliability measures for weight handling and strength, 67 percent for posture and mobility, and 56 percent for locomotion (Bieniek and Bethge, 2014). The reliability of the extracted test-retest measure for balance was acceptable. In addition, 89 percent of the interrater reliability measures and all of the intrarater reliability measures showed acceptable levels of reliability (Bieniek and Bethge, 2014).

Gross and colleagues (2004) evaluated the prognostic value (predictive validity) of the Isernhagen Work Systems FCE in determining recovery among workers' compensation claimants with low back pain and their eventual return to work. Workers with low back pain underwent this FCE at least 6 weeks following the date of their accident. The workers evaluated were considered to have met or exceeded the expected healing time before undergoing the FCE. Additional medical treatment and rehabilitation were not expected to improve the individual workers' clinical and functional ability. This study found that FCE performance was a poor predictor of time to claim closure (actual return to work). Only 4 percent of the workers met or exceeded the FCE criteria for return to work (Gross et al., 2004), whereas at 1-year follow-up, nearly 95 percent of the workers had returned to work. There was a weak association between better FCE performance and earlier return to work.

In the follow-up study, Gross and Battié (2004) determined that better FCE performance did not correlate with a decreased risk of injury recurrence. Workers' compensation claimants with back pain were followed after undergoing an FCE. The number of failed tasks and performance on the floor-to-waist lift task were used as indicators of FCE performance. Sustained recovery was evaluated according to whether disability benefits had been restarted, a previous claim for back pain had been reopened, or a new claim had been initiated. Workers with the best FCE performance (lowest number of failed FCE tasks) consistently showed a higher risk of recurrent back injury (Gross and Battié, 2004).

Gross and Battié (2005) also found that FCE performance did not predict sustained return to work in workers with chronic back pain. In this study, better FCE performance was mildly associated with faster return to work. However, FCE performance was a poor predictor of recurrent back problems or self-reported disability. The Isernhagen Work Systems FCE protocol was utilized for this study. Higher weight on the floor-to-waist lift and a lower number of failed FCE tasks (i.e., better FCE performance) were weakly associated with faster suspension of benefits and claim closure (return to work). However, FCE performance was not significantly correlated with self-reported outcomes of work status, pain intensity, and disability. The authors conclude that FCE performance should not be relied on to forecast the ability of injured workers with back pain to return to work safely (Gross and Battié, 2005).

FCEs were initially envisioned as an objective measure of functional ability that could be used to assess an individual's ability to perform the tasks of a particular occupation or used as a tool to identify specific occupational restrictions and/or limitations. Unfortunately, while potentially helpful in particular settings, FCEs have been shown to lack the reliability and validity necessary to fulfill these objectives.

Physical Function Scales of the Work Disability Functional Assessment Battery (WD-FAB)

WD-FAB is a self-report instrument that uses item response theory and computer adaptive testing (CAT) to assess physical and mental functioning. The WD-FAB physical function scales currently include four multi-item scales: Basic Mobility (56 items), Upper Body Function (34 items), Fine Motor Function (45 items), and Community Mobility (11 items) (Meterko et al., 2018). CAT algorithms customize the selection of items based on previous responses to avoid the need to administer all the items in the item banks (Meterko et al., 2018). The items are agreement based or ability based. Agreement-based items ask the respondent to “specify your level of agreement” on a 4-point Likert-type scale ranging from “strongly agree” to “strongly disagree.” Some items also include an option for the respondent to indicate “unable to do.” Ability-based items provide 5-point response options for the question “are you able to”; response options range from “yes, without difficulty” to “unable to do.” Both types of items include an opt-out response of “I don’t know” (Meterko et al., 2018). A study of three groups of working-age adults (aged 21 to 66 years)[1] showed “substantial support” for the instrument’s reliability and construct validity (Meterko et al., 2018). The WD-FAB is also discussed in Chapter 4.

MUSCULOSKELETAL ASSESSMENTS

This review of musculoskeletal assessments is organized by body parts: upper extremities and hands and fingers, back and neck, and lower extremities and feet. Instruments used for musculoskeletal assessment are listed in Annex Table 5-4.

Upper Extremities and Hands and Fingers

Disabilities of the Arm, Shoulder and Hand Questionnaire (DASH)

The DASH is a self-report questionnaire that consists of 30 questions designed to measure physical function and symptoms in patients with any or several musculoskeletal disorders of the upper limbs. Questions ask about the individual’s symptoms, as well as his or her ability to perform certain activities. The questionnaire was designed both to help describe the disability experienced by people with upper-limb disorders and to monitor

[1]The groups represented individuals who were “(1) unable to work because of a physical condition (n=375); (2) unable to work because of a mental health condition (n=296); [and] (3) general United States working age sample (n=335)” (Meterko et al., 2018).

changes in symptoms and function over time when administered during successive visits to a provider's office (IWH, n.d.).

The DASH gives clinicians and researchers a single reliable instrument that can be used to assess any or all joints in the upper extremities (IWH, n.d.). It performs well in both these roles. It can detect and differentiate small and large changes in disability over time following surgery in patients with upper-extremity musculoskeletal disorders and show treatment effectiveness after surgery for subacromial impingement and carpal tunnel syndrome (Wong et al., 2007). Both the DASH and the QuickDASH (an 11-question abbreviated version) have been found to be effective in measuring functional status after traumatic hand injury (Wong et al., 2007). QuickDASH has been shown to provide the same information as the full DASH in less time, and is completed more often (Aasheim and Finsen, 2014).

The DASH is available in 27 languages and has been validated in French, Persian, and Japanese, as well as English. Many measurement properties of QuickDASH have been evaluated in multiple studies and across most measurement properties (Kennedy et al., 2013).

Patient-Reported Outcomes Measurement Information System (PROMIS) Upper-Extremity Questionnaire

As described in Chapter 4, PROMIS comprises a bank of questionnaires containing highly reliable, precise measures of self-reported health status developed under the auspices of the National Institutes of Health (HHS, 2018). The PROMIS Upper-Extremity Questionnaire is intended to evaluate the mobility of the upper extremities. It may be used as an upper-extremity disability measure and correlates with the QuickDASH questionnaire (Overbeek et al., 2015).

Patient-Rated Elbow Evaluation (PREE)

The 20-item PREE questionnaire is designed to measure elbow pain and disability in activities of daily living (Vincent et al., 2015). Patients rate their levels of elbow pain and disability from 0 to 10 on two subscales: the pain subscale, which consists of five items for which individuals rate pain from 0 to 10 (0 = no pain, 10 = worst ever); and the function subscale, which consists of 11 items for which individuals rate the level of difficulty, again from 0 to 10 (0 = no difficulty, 10 = unable to do) (MacDermid, 2010). Additionally, a total score can be computed that measures both pain and functional problems weighted equally, on a scale of 0 (no disability) to 100, and a higher score indicates more pain and functional disability (MacDermid, 2010).

Patient-Rated Wrist Evaluation (PRWE)

The PRWE is a 15-item questionnaire designed to assess pain, disability, and functional difficulties in activities of daily living resulting from injuries affecting the wrist joint area (MacDermid et al., 1998). It is used to evaluate pain in the affected wrist and activities affected by the wrist injury, as well as the frequency and intensity of pain and when it occurs—for example, at rest; upon repeated movements; or during lifting of heavy objects, household tasks, fine hand movements, self-care, and/or toileting. It also can be used to assess the individual's ability to participate in household, occupation-related, and recreational activities. The PRWE has two subscales: the pain subscale consists of 5 items, with responses ranging from 0 (no pain) to 10 (worst pain ever); the function subscale consists of 10 items divided into specific (6 items) and usual (4 items) activities, with responses ranging from 0 (no difficulty) to 10 (unable to do). A composite score can also be obtained. The objectives of administering the PRWE include determining the degree of wrist-related musculoskeletal disability; predicting the prognosis for a patient with wrist injury considering his or her baseline score; and communicating the pain and degree of musculoskeletal disability associated with wrist injuries in a meaningful way to the patients, health care professionals, and insurance companies (PRWE, 2011).

Michigan Hand Outcomes Questionnaire (MHQ)

The MHQ is a self-report outcome assessment used to measure and compare outcomes across different hand conditions. It consists of 37 core questions, takes approximately 15 minutes to complete, and can be self-administered or administered by research personnel (University of Michigan, 2014a). It can be used to assess an individual's general hand function or, if administered several times (e.g., pre- and postoperatively), to assess changes in hand function. It includes demographic information and contains six distinct scales: overall hand function, activities of daily living, pain, work performance, aesthetics, and patient satisfaction with hand function (University of Michigan, 2014a).

Back and Neck

Instruments for assessment of low back pain have been used to evaluate the correlation between individuals' self-report scores and their actual work status (Sivan et al., 2009). In general, the correlation between chronic back pain and work status has been demonstrated to be modest at best. The instruments evaluated include the RMDQ, the ODI, and the Orebro Musculoskeletal Pain Questionnaire. Descriptions of these and several other commonly used instruments for assessment of back and neck pain follow.

Roland-Morris Disability Questionnaire (RMDQ)

The RMDQ is a self-report measure consisting of 24 items. First published in 1983, it has become one of the most widely used outcome measures for assessing the level of disability experienced by a person suffering from low back pain (Chapman et al., 2011; Roland and Morris, 1983). The RMDQ has been shown to yield reliable measurements that are valid for and are sensitive to change over time for groups of patients with such pain (Roland and Morris, 1983). Its questions relate to an individual's perceptions of his or her back pain and associated disability, including items on physical ability/activity (15), sleep/rest (3), psychosocial factors (2), household management (2), eating (1), and pain frequency (1). The questionnaire can be completed in about 5 minutes and without assistance (Stevens et al., 2016). The RMDQ is most sensitive for patients with mild to moderate disability (Davies and Nitz, 2013), with greater levels of disability being indicated by higher numbers on a 24-point scale (Stevens et al., 2016). Internal consistency and test-retest reliability are good (Smeets et al., 2011).

Results from the RMDQ have a moderate to large correlation with those from other self-report disability questionnaires, such as the Quebec Back Pain Disability Scale (Quebec Scale) and the ODI, and this instrument has demonstrated strong qualities with respect to content and construct validity, feasibility, linguistic adaptation, and international use (Calmels et al., 2005). The original 24-item questionnaire has been shortened to create 18- and 23-item versions and has been cross-culturally adapted or translated for use in other countries, although the original is still the most widely used and validated version. The RMDQ is highly regarded among questionnaires designed to evaluate disability caused by low back pain because of its psychometrics and feasibility (Rocchi et al., 2005).

Oswestry Disability Index (ODI)

The ODI, derived from the Oswestry Low Back Pain Questionnaire, is used to quantify disability due to low back pain. Developed in 1990, this 10-point self-report, self-administered outcome questionnaire is one of the outcome measures most commonly used to assess individuals with low back pain (Vianin, 2008). Its psychometric properties have been well established (Vianin, 2008), and it is the most effective questionnaire for assessing persistent severe disability (Davies and Nitz, 2013). It takes 5 minutes to complete and 1 minute to score. Scores reflect the severity of the individual's back pain, ranging from minimal to bedbound. The ODI is an extremely important tool for use by both researchers and disability evaluators to measure individuals' permanent functional disability, and is considered the "gold standard" among low back functional outcome tools (Fairbank and Pynsent, 2000).

The ODI evaluates pain intensity based on 10 topics: intensity of pain, lifting, ability to care for oneself, ability to walk, ability to sit, sexual function, ability to stand, social life, sleep quality, and ability to travel (Fairbank and Pynsent, 2000; Fairbank et al., 1980; Yates and Shastri-Hurst, 2017). Each question is scored on a scale of 0 to 5, from least amount of disability to severe disability (a total score of 0 = no disability, 100 = maximum disability possible) (Fairbank and Pynsent, 2000). Self-report scores on the ODI were found to be correlated with those on the RMDQ and the Quebec Scale. The ODI may be considered more reliable than the RMDQ, but both are considered well validated (Grotle et al., 2005).

Quebec Back Pain Disability Scale (Quebec Scale)

The Quebec Scale is another commonly used self-report outcome measure designed to evaluate functional disability in patients with chronic back pain (Kopec et al., 1995). This 20-item questionnaire is used to assess symptoms and severity of low back pain and the degree to which the individual's pain impacts functional activities. Rocchi and colleagues (2005) critically compared nine self-administered questionnaires designed to evaluate disability caused by low back pain with respect to their psychometric properties (reliability, validity, responsiveness) and practical and technical aspects (number of items, number and kind of domains, scaling of items, scoring, time to complete, validated transitions). The authors determined that the ODI, the Quebec Scale, and the RMDQ appeared to be fully validated from a psychometric standpoint. They found the RMDQ and the ODI to be preferable based on their psychometrics and feasibility (Rocchi et al., 2005). The reliability of the Quebec Scale has been found to compare favorably with that of the ODI (Davidson and Keating, 2002).

Neck Disability Index (NDI)

Published in 1991, the NDI was the first instrument designed to assess self-rated disability in patients with neck pain. This self-administered, self-report instrument consists of 10 questions asking individuals to provide information about how neck pain affects their everyday life activities. The NDI is the most widely used and strongly validated instrument for assessing self-rated disability in patients with neck pain and has been used effectively in both clinical and research settings in the treatment of such pain (Vernon, 2008).

Lower Extremities and Feet

Lower-Extremity Functional Scale (LEFS)

The LEFS comprises 20 questions designed to help determine how a person's lower-limb function affects the ability to perform daily activities, with a focus on disorders of the hip, knee, leg, ankle, and foot (Martin and Irrgang, 2007). Questions address activities that range from walking between rooms to running on uneven ground (Dingemans et al., 2017). The questions are subdivided into four groups, which cover activities that impose increasing physical demands. The lower the score, the greater is the disability. In general, scores have been observed to decrease with age, and men tend to score slightly higher than women; no statistical correlation has been found between socioeconomic status and LEFS scores. People who are unfit for work have significantly lower LEFS scores (Dingemans et al., 2017). The scale has been found to be reliable and valid for assessing functional impairment in a wide array of patient groups with lower-extremity musculoskeletal conditions (Binkley et al., 1999) and for a broad range of disorders and treatments (Dingemans et al., 2017).

Functional Gait Assessment (FGA)

The FGA is an ambulation-based balance test used to assess postural stability during walking tasks, with a focus on 10 facets of gait and balance (Leddy et al., 2011). The performance-based measure consists of a timed walk with a 10-item clinical gait test that includes tasks requiring many postural adjustments (Wrisley et al., 2004). Participants are asked to perform gait activities such as "walk at normal speeds, at fast and slow speeds, with vertical and horizontal head turns, with eyes closed, over obstacles, in tandem, backward, and while ascending and descending stairs" (Wrisley and Kumar, 2010, p. 762). The FGA is scored on a four-level (0–3) ordinal scale. The scores range from 0 to 30 with lower scores indicating greater impairment (0 = severe impairment, 30 = normal ambulation). It takes 5–10 minutes to administer (AbilityLab, 2016). The FGA demonstrates acceptable reliability, internal consistency, and concurrent validity with other balance measures used for patients with vestibular disorders (Wrisley et al., 2004). It has shown reliability and validity for assessing balance in individuals with Parkinson's disease and has been validated for use in individuals not only with vestibular disorders but also with such diagnoses as Parkinson's disease, as well as in community-dwelling older adults (Leddy et al., 2011). It has been shown to have high interrater reliability across the patient populations studied and high concurrent validity in patients with Parkinson's disease and stroke and community-dwelling older adults (Weber et al., 2016).

Foot and Ankle Ability Measure (FAAM)

The FAAM is a self-report measure used to assess the physical performance of individuals with musculoskeletal disorders affecting the lower leg, foot, and ankle (AbilityLab, 2015; Martin and Irrgang, 2007; Martin et al., 2005). It consists of 29 questions divided into two subscales, which are scored separately: an activities of daily living subscale (21 items) and a sports subscale (8 items). A higher score represents a higher level of ability (Martin and Irrgang, 2007; Martin et al., 2005). The 5-point Likert scale responses range from zero (unable to do) to 5 (no difficulty in performing).

Summary

Numerous performance-based measures are available for assessing physical function in specific areas (e.g., range of motion, strength, balance). While such measures provide quantitative information about the areas assessed, they are less useful for predicting whether or how well the individual will be able to perform everyday activities, including work, on a sustained basis. Self-report outcome measures complement the results of performance-based assessments to provide a more complete picture of an individual's overall functional status. An important consideration in determining the utility of these assessment tools is that if an individual has a solitary injury such as a wrist injury, and pain is not a significant element of the presentation, then using PRWE alone may suffice. However, if more than one area of the upper extremity is involved, then a more global assessment of upper-extremity function, such as the DASH or the PROMIS Upper-Extremity Questionnaire, may be in order. In the event the entire body is involved, the WD-FAB physical function scales may be helpful. However, it would be operationally challenging to identify distinct constructs for use depending on the body location affected that would provide the answer as to whether an individual is disabled. Such use also would be counter to the premise of the importance of considering overlapping and synergistic conditions an individual may display in determining disability (see Chapter 4). Thus while this section reviews some of the tools commonly used to assess musculoskeletal function, it would be naïve to conclude that one tool can be used in isolation to determine the degree of an individual's disability; this is the case particularly for tools that address a single joint and that may be validated only for certain diagnoses. It is beyond the scope of this report to provide a comprehensive approach to disability determination based on a combination of these and similar assessment tools while also taking into account pain and other factors that might be relevant.

PHYSICAL FUNCTION AND PAIN

As discussed, assessing an individual's functional abilities is a complex undertaking. For example, a musculoskeletal impairment resulting from an upper-body injury may preclude performing a job that requires heavy lifting, but not a job that involves substantial walking without heavy lifting. However, it is important also to consider the presence of potential modulating factors, such as depression (discussed in Chapters 6 and 7) or pain, that can affect an individual's performance and may preclude his or her ability to perform either job. Pain often impairs function and occurs in many forms, including monophasic events, chronic episodic conditions, and chronic persistent problems (Stewart et al., 2003). Multiple studies focus on the impact of pain as a predictor of work absenteeism and reduced work performance (Bergström et al., 2014; Félin-Germain et al., 2018; Kawai et al., 2017; Kresal et al., 2015; McDonald et al., 2011; Tsuji et al., 2018). Potential side effects of treatment for pain that can affect the ability to work include drowsiness, nausea, vomiting, diarrhea, impaired capacity for concentration, deficits in information processing and memory, and slower psychomotor speed and reaction time (Bet et al., 2013; Kendall et al., 2010; Porreca and Ossipov, 2009). To the extent that use of these medications interfered with an individual's ability to work, they would be considered "interrupters" in the committee's conceptual framework (see Chapter 2).

Although pain is a symptom, not a function, a few common pain assessment instruments are described here because its presence affects so many functional abilities. It is important to note that there are many valid approaches to measuring pain, and the tools presented are by no means exhaustive.

Visual Analog Scale (VAS) for Pain

The VAS for the measurement of pain is a horizontal (more common) or vertical scale depicted as a 10-centimeter line typically labeled "no pain" at one end and "worst imaginable pain" at the other (Hawker et al., 2011; Williamson and Hoggart, 2005).[2] The individual is asked to place a mark along the line to indicate his or her level of pain, typically as experienced within the past 24 hours. The individual's pain intensity is scored by measuring the distance in millimeters from zero (no pain) to the mark. The higher the number, the greater is the individual's pain, up to a maximum of 100 (worst imaginable pain). The VAS for pain has been shown to be reliable and valid for the assessment of pain in an acute setting, such as a

[2]Evidence suggests that the orientation of the line should reflect normal reading direction (horizontal or vertical) for the person being assessed to decrease error (Williamson and Hoggart, 2005).

hospital emergency department (Bijur et al., 2001). Individuals with cognitive limitations or those with fine motor deficits may have trouble completing the assessment (Hawker et al., 2011; Williamson and Hoggart, 2005).

Numeric Rating Scale (NRS) for Pain

The NRS for pain is another self-report measure of pain intensity in adults, typically within the past 24 hours. The most common version is an 11-point scale anchored by 0 ("no pain") at one end and 10 (e.g., "worst pain imaginable") at the other (Hawker et al., 2011; Williamson and Hoggart, 2005). The NRS for pain is valid and reliable. Unlike VAS for pain, it can be administered verbally as well as in writing, and it is simpler for respondents to comprehend and for examiners to score. As measures of pain intensity, the VAS and the NRS are not designed to capture the complexity of individuals' experience of pain (Hawker et al., 2011).

PROMIS Pain Interference Instruments

The PROMIS Pain Interference instruments measure self-reported effects of pain on relevant aspects of one's life, including the extent to which pain hinders engagement with social, cognitive, emotional, physical, and recreational activities (HealthMeasures, 2017). The instruments assess pain interference over the past 7 days. The PROMIS measure yields a single summary score, which is then converted to a T-score (mean of 50, standard deviation of 10). (Meterko et al., 2015). The instrument has been shown to be valid and reliable (Amtmann et al., 2010).

CARDIOVASCULAR ASSESSMENTS

Cardiovascular disease can limit work capacity in multiple ways. First, the heart can be too weak to pump an adequate amount of blood to provide oxygen to exercising muscle. Even when the contractility of the heart appears adequate at rest, such structural impairments as coronary artery or valve disease may limit the response to exercise. Some individuals who are physically capable of exercise may be restricted by their cardiologist because of the risk of a sudden life-threatening arrhythmia or other collapse during exertion, such as can occur with some inherited genetic heart diseases, advanced hypertrophic cardiomyopathy, or critical aortic stenosis, but this would be rare. SSA listing criteria for cardiovascular conditions, such as chronic heart failure or ischemic heart disease, generally include both (1) demonstration of objective abnormality of cardiac structure/cardiac function that could seriously limit patient function, and (2) demonstration of patient functional limitation by exercise testing, although the functional criteria are occasionally met instead by three or more events requiring intervention during a 12-month period (IOM, 2010, pp. 86, 118; SSA, 2008, 4.02, 4.04).

Demonstration of functional cardiac limitation can be supported by patient questionnaire and a medical provider's subjective designation of functional impairment based on integrated clinical assessment. As patient mood, recall, and incentives can influence both patient and provider assessment, these factors are best combined with exercise testing to provide objective functional measurement during an exercise tolerance test. This test is most reliable and least subject to assumptions when it can be performed with simultaneous gas exchange analysis. Annex Table 5-5 provides relevant information for selected cardiac and cardiovascular assessments described below.

Patient Questionnaires

Multiple validated patient questionnaires can be used to assess symptomatic limitation with cardiovascular disease. The two standard questionnaires for heart failure are the Minnesota Living with Heart Failure Questionnaire (MLHFQ) (Rector et al., 2006) and the Kansas City Cardiomyopathy Questionnaire (KCCQ) (Joseph et al., 2013). Both have been extensively validated and shown good reliability, responsiveness, performance across populations, feasibility, and interpretability (Kelkar et al., 2016). These instruments were not developed specifically to address physical function but to quantitate the overall impact of a decrease in heart function on the life of an individual. They encompass multiple domains, including physical, social, and emotional. Both questionnaires are approved by the U.S. Food and Drug Administration as valid for demonstrating the value of medical interventions, either medications or devices.

The MLHFQ contains 21 items for which the scores (maximum of 5 points each, with a higher score indicating more limitations) are summed. Conversely, a higher score indicates fewer limitations in the KCCQ, which includes 23 items and more specific description of limitation in specific activities, such as walking, climbing stairs, housework, and yardwork. Both questionnaires have been administered to many thousands of patients with impaired cardiac function and have been translated extensively into other languages. Both have been demonstrated to predict death or permanent disability with moderate accuracy for the minority of patients who deteriorate soon after testing, but they have not been used specifically to determine intermediate levels of function. General ranges of the aggregate scores correlate with mild, moderate, and severe limitation of activity in heart failure populations but do not correlate reliably with capacity for individuals.

The Seattle Angina Questionnaire was developed using a framework similar to that for the KCCQ for heart failure. Shorter versions of both of these instruments were recently developed and validated, consisting of 7 items for the angina form (Chan et al., 2014) and 12 items for the heart failure form (Spertus and Jones, 2015).

Recognizing the limitations of all patient questionnaires, the two major limitations of these heart failure and angina questionnaires for the assessment of ability to work are (1) the broad range of the questions, which include social and emotional domains as well as physical functioning; and (2) the definition of limitations as being specifically attributable to the diagnosis of heart failure or angina. Many patients with cardiac disease have general symptoms of fatigue and breathlessness regardless of their specific diagnosis. Two other patient questionnaires are keyed directly to the performance of work activities with any cardiac disease: the Specific Activity Scale (Goldman et al., 1981) and the Duke Activity Status Index (Hlatky et al., 1989). Both include questions on specific daily activities that elicit information directly translatable to a measure of work intensity (see discussion of metabolic equivalents [METS] below). However, these instruments have been validated only in small cohorts and have not been widely adopted.

Provider Classification of Cardiac Limitation

The New York Heart Association (NYHA) Functional Classification was proposed in 1902 as a convenient tool for medical providers seeking to describe the symptomatic severity of cardiac disease in their patients. Its wording has been modified for different diagnoses but in its broadest form specifies the nature of activity that is limited and not the limiting symptoms (see Table 5-1). It is used most commonly in populations of adults with heart failure or angina, but is often used as well to describe the degree of limitation with other cardiac disorders, such as adult congenital heart disease and acquired valve disease. The universality and simplicity of the NYHA classification have rendered it common vernacular in routine patient care, triage for major interventions, and eligibility for specific therapies. Because the classification explicitly describes limitation in relation to routine activity, it can serve as a useful initial screen for employability: Class I patients would be expected to have no cardiac limitation on work; Class II patients to have limitation on moderately strenuous work; some Class III patients to be able to do sitting work with occasional walking, and others to be unable to work in any job requiring physical activity; and Class IV patients, with symptoms at rest or with any movement, generally to be unable to perform regular work, except for occasional highly motivated individuals doing cognitive work.

TABLE 5-1
New York Heart Association Functional Classification

Class I. No limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, dyspnea, or anginal pain.
Class II. Comfortable at rest with only slight limitation of physical activity. Ordinary physical activity results in fatigue, palpitation, dyspnea, or anginal pain.
Class III. Comfortable at rest but marked limitation of physical activity. Less than ordinary activity causes fatigue, palpitation, dyspnea, or anginal pain.
Class IV. Inability to carry on any physical activity without discomfort. Symptoms often present even at rest. If any physical activity is undertaken, discomfort is increased.

SOURCES: Adapted from Greenberg and Kahn, 2012, p. 507; NYHA, 1994, pp. 253–256.

The NYHA classification is notoriously subject to interobserver variability by one class, although very rarely by two classes (Raphael et al., 2007). Unless they question patients systematically about specific activities, physicians commonly underestimate the limitations on specific activities that patients report (Albert et al., 2010). There is a strong statistical correlation between scores on the patient heart failure questionnaires and the NYHA classes, such that more severe patient-described limitation tends to correlate with worse classification of limitation by the physician. However, the ranges for each class have wide overlap, and the correlation is not tight; for example, *r* values range from 0.55 to 0.62 in one large study (Joseph et al., 2013). Similarly, the NYHA classification has been aligned with performance on exercise tests, but again, although the correlation is strong, wide overlap in ranges limits the ability to classify an individual patient correctly.

Despite the serious limitations of the NYHA classification, it remains a surprisingly robust predictor of outcomes with heart failure, as measured by rates of hospitalization, heart transplant or mechanical circulatory support, or death. When it indicates severe limitation, as in Class III or IV, objective assessment is likely to confirm substantial limitation on work capacity. When a discrepancy exists between a provider's designation of Class I or II and a patient's description of more severe activity limitation, an exercise test is frequently helpful to provide more objective assessment of such limitation.

Exercise Testing for Measurement of Functional Capacity

A major advantage of exercise testing is the resulting estimates of work capacity that can be compared directly with the levels of work estimated for activities in many common occupations. The most frequently used means of objectively assessing cardiac exercise capacity is exercise testing on a treadmill (Ellestad et al., 1969). This method requires an exercise testing laboratory with exercise and cardiac monitoring equipment, a technician or physician/nurse practitioner/physician's assistant trained in performance of exercise testing, and either the same or a different provider with experience in interpreting the results of the tests. Most hospitals can meet these requirements because this type of testing is commonly used to evaluate patients with chest pain that may be cardiac in origin. Outpatients may have this testing in a nearby hospital or in the outpatient office of a cardiology practice; in some cases, the testing may be carried out by experienced primary care providers in their offices. Although most exercise testing is performed on an upright treadmill, some facilities and some patients may have the option of using a bicycle for greater stability or in some cases for more consistent calibration with actual work performed. In general, peak exercise on a bicycle requires a slightly lower total energy expenditure relative to that on a treadmill.

In rare cases, patients may be deemed to be at excessive risk of a serious event during exercise testing, which would then be supported by written documentation. Ideally, a waiver of exercise testing should be completed by a board-certified cardiologist and include details of the cardiac diagnosis and the nature of the risk. The importance of the safety of exercise testing and exercise for rehabilitation is increasingly appreciated such that few patients should require a waiver of exercise testing because of cardiac risk.

Routine Exercise Testing

Standard exercise testing is generally performed with consistent encouragement to the level of patient exhaustion, unless early termination is required because of abnormal events—such as a drop in blood pressure, chest pain or electrocardiographic evidence of severe ischemia, or life-threatening arrhythmias—suggesting active disease. Such events generally are considered sufficient evidence of cardiac limitation even if the test was stopped early. Most tests otherwise continue to completion until limited by shortness of breath; fatigue; or noncardiac conditions, such as musculoskeletal pain or weakness. Shortness of breath or occasional oxygen desaturation due to pulmonary disease may limit exercise performance, but these symptoms can also be due to cardiac disease, so distinguishing limitations resulting from these conditions is not easy at the time of exercise testing.

The professional administering the test needs to be experienced in recognizing typical signs of disease and risks during the test, as well as in assessing the individual's degree of effort. An individual's intentional restraint of effort can be suspected from a lack of typical heart rate and blood pressure changes or of the appearance of vigorous effort during exercise. More difficult is distinguishing individuals with true cardiac limitations from those who are peripherally deconditioned as a result of prolonged inactivity.

Assessment of functional capacity is influenced by peak blood pressure and heart rate but generally summarized by the total exercise time and highest workload achieved before stopping. The latter is calculated from the speed and elevation of a treadmill (or resistance on a bicycle) and expressed in terms of METS. One metabolic equivalent is the energy required each minute to support normal existence at rest. For example, easy walking is estimated to require approximately 2.5–3 METS of energy expenditure. Inability to perform 5 METS generally is sufficient to meet the current SSA listing for cardiac disability (SSA, 2008). There are tables indicating the estimated level of METS for common activities involved in employment.

Estimation of the actual energy expenditure from the external workload performed during exercise requires several assumptions, however. The first is that the average watts expended to perform an activity per kilogram will be the same for individuals with different levels of muscular fitness and body composition. The second is that the individual is in a steady state where the energy used each minute is the energy required to meet the ongoing needs of the exercising muscle. The third assumption is that the individual will not exercise beyond the level of what would normally be tolerated. Extreme performance by athletes or by individuals trying to override their limitations can be endured for a few minutes during which the workload exceeds the physical capacity and is being performed in a state of energy deficit that is paid back slowly during prolonged fatigue. Individuals working with chronic cardiac disease can become physiologically and psychologically adapted to brief bouts of activity at levels that are beyond the capacity of their heart to sustain. For this reason, the estimation of METS for an individual's workload during a routine exercise test commonly exceeds the individual's capacity to sustain a given level of work.

Exercise Testing with Simultaneous Gas Exchange Analysis

The addition of gas exchange analysis during exercise testing allows work performance to be measured directly in terms of oxygen consumption per kilogram per minute rather than estimated. This approach is referred to as cardiopulmonary exercise testing, often abbreviated CPET or CPX. Measurement of peak oxygen consumption has consistently been shown to be the most objective and reproducible means of assessing exercise capacity in cardiac disease (ATS/ACCP, 2003), and has long been known to relate closely to the peak cardiac output achieved during exercise (Metra et al., 1990). According to the statement of the American Thoracic Society/American College of Chest Physicians on cardiopulmonary exercise testing, such testing "complements other clinical and diagnostic modalities and by directly quantitating work capacity improves the diagnostic accuracy of impairment/disability evaluation … [and] may be particularly helpful when job-related or exertional complaints are disproportionate to the measured … impairments" (ATS/ACCP, 2003, p. 217).

The usual conversion from estimated to measured energy requirements is that 1 MET is equivalent to 3–4 ml/kg/min of oxygen consumption, such that walking would require about 10 ml/kg/min. The threshold of 5 METS for the current SSA listing for cardiac disability for both ischemic heart disease and chronic heart failure would then correspond to about 15–18 ml/kg/min (IOM, 2010, pp. 86, 118). For patients who exercise to exhaustion, this exact measurement often yields a lower estimate of work capacity than that derived from their last attempted level of the routine exercise test. Stoic patients or those accustomed to strenuous activity despite cardiac disease often achieve an exercise workload associated with a higher level of METS than what is actually measured by oxygen consumption. The result is overestimation of the work such an individual could actually perform on a regular basis. On the other hand, the simultaneous monitoring of carbon dioxide production and its comparison with oxygen consumption can be used to measure intensity of effort so that individuals can be urged to maximum effort, and their failure to give that level of effort can be recognized.

A unique advantage of cardiopulmonary exercise testing compared with most tests of functional capacity is that it provides ancillary information about the sustainability of exercise. While gas exchange analysis measures peak oxygen consumption that reflects the highest level of exertion possible, it also provides information on the individual's anaerobic threshold, or the level of work he or she would likely be able to sustain for at least 50 minutes (if not limited by other conditions). This level is generally about 70 percent of the person's peak oxygen consumption. Individuals who stop voluntarily before achieving their anaerobic threshold are considered to have undergone an incomplete and inadequate test of cardiovascular capacity (Wasserman, 1994, p. 122).

Extensive tables exist with which to compare the estimated METS or oxygen consumption required during specific work-related or recreational activities (Arizona State University, 2011; Erb, 1970, Table F-4; Wasserman, 1994, p. 469). For instance, the estimated oxygen consumption while sitting during computer use would be just slightly more than while sitting at rest, while pumping gas at a gas station would require three times as much oxygen consumption as sitting at rest. It may appear simple to align these requirements with the METS or oxygen consumption measured at peak exercise, but such an interpretation would overestimate the capacity to sustain a given work activity for a prolonged period of time. Instead, the cardiac work level that could be sustained under ideal conditions for about 50 minutes is best estimated by the direct measurement of anaerobic threshold during the exercise test, or approximated by the calculation of 70 percent of the peak METS or peak oxygen consumption achieved during peak exercise.

The gas exchange apparatus requires the individual being tested to wear a noseclip and mask or mouthpiece similar to a snorkel, which most people can tolerate with adequate coaching. Gas exchange measurement equipment requires additional training to maintain and calibrate, and performance of the test requires technicians with specific expertise, often exercise physiologists or respiratory therapists. Exercise testing with gas exchange frequently is performed in pulmonary function laboratories in community hospitals and clinics. These facilities may be supervised by cardiologists in large cardiac centers, particularly those centers offering cardiac transplantation, where a peak oxygen consumption of less than 12–14 ml/kg/min (3.5–4.0 METS) is often used as a general indication of cardiac disease that is sufficiently severe to warrant consideration for transplantation.

The 6-minute hall walk has been used as a convenient way to assess exercise capacity without requiring an exercise laboratory. Originally developed to assess heart failure patients in trials of investigational therapies, this test involves encouraging a patient to cover as much floor as possible along a measured corridor during a 6-minute period. The results are most helpful when they are very high or very low, but do not correlate well with more objective exercise testing in the middle ranges, where ability to work would be assessed (Lucas et al., 1999). This method is also highly dependent on individuals' level of effort, which may be reduced during evaluation for disability. The gait speed measured during a 6-meter walk is gaining popularity as an index of frailty in populations with chronic disease (Peel et al., 2012), but is also highly dependent on effort and has not been compared against work requirements.

Summary

Patient questionnaires and NYHA classification by providers offer useful information on limitation to routine daily activities, from dressing to climbing stairs and doing household chores. Current SSA listing criteria for disability require both the demonstration of cardiac structural abnormality on imaging and routine exercise treadmill testing to estimate peak work performance, unless patients have had at least three cardiac events requiring intervention in the preceding year. The addition of metabolic gas exchange analysis to regular treadmill testing can provide more precise measurement of the energy expenditure that can be sustained, which can be compared with the energy requirements documented for specific work activities.

VISUAL ASSESSMENTS[3]

The number of individuals with visual impairment is increasing in the United States, with approximately 11 million estimated in 2015, and by 2050, that number is projected to double (Varma et al., 2016). This is particularly important because vision problems affect productivity, activities of daily living, and quality of life.

Visual functions relevant to work requirements include far and near visual acuity and peripheral vision (DOL, 2018, p. 130). In addition, O*NET lists depth perception, night vision, glare sensitivity, and color discrimination among its visual sensory abilities (DOL, n.d.). With regard to driving, visual acuity, peripheral visual field sensitivity and extent, and color vision appear to be important visual functions for achieving good performance. Good performance refers to vision impairments that do not increase a meaningful risk of failures in safe driving (e.g., collisions, accidents, pedestrian or bicycle collisions) that may affect the driver and/or those in other nearby vehicles. As discussed further below, while it is not necessary that drivers have normal color vision, there are certain situations in which persons with color vision deficiencies (moderate to advanced anomalous trichromats and dichromats) may experience confusion that may lead to increased risk of a driving problem (e.g., inability to distinguish between a flashing red or yellow light, recognizing hazard warnings).

Reductions in visual function can occur when environmental conditions, such as poor weather (rain, snow, fog, sleet), smoke, dim or nonexistent lighting, low contrast, or abrupt changes in lighting (bright to dark or vice versa), compromise or diminish visibility. Most visual functions are degraded under these conditions, although a few (flicker detection, motion sensitivity, object localization) are affected only minimally. Given the effect of environmental conditions on visual function, it is important to keep in mind that most vision tests are designed to evaluate abilities under optimum visibility conditions that may not reflect performance when those conditions are degraded. The use of eyeglasses or contact lenses to correct refractive errors also can be adversely affected by poor environmental conditions. Additionally, there are a number of ocular and neurological conditions that are subject to the greatest amount of impairment when visibility is

[3]This section draws heavily on a paper commissioned by the committee for this study (Wilkinson et al., 2018).

degraded. Annex Table 5-6 lists selected instruments used to measure visual function.

Visual Acuity

Visual acuity, or the finest spatial detail that can be resolved, is the most commonly used index of visual capacity. Acuity predicts the ability to perform a number of important visual discrimination tasks, such as those involved in reading and the identification of critical details at near and far distances. If acuity is reduced, for example, it is not possible to discriminate objects at a distance or the fine differences that distinguish letters from each other.

Visual acuity is measured most commonly with the use of high-contrast letters (dark letters on a white background). The letters are constructed such that their thickness is one-fifth the size of their overall width and height. The minimum size of letters that can be accurately recognized determines the visual acuity measure. Visual acuity is normally specified according to the visual angle subtended by the fine detail (thickness) of the letters, known as the minimum angle of resolution. By convention, a minimum angle of resolution of 1 minute of arc corresponds to a visual acuity of 20/20 in Snellen notation. The Snellen acuity notation is reciprocally related to minimum angle of resolution; that is, 20/40 corresponds to 2 minutes of arc, 20/200 corresponds to 10 minutes of arc, and so forth. A number of physiological factors (e.g., refractive state, pupil size) and environmental factors (e.g., illumination level, contrast) can affect visual acuity measurements. Cognitive status is another important factor that can affect vision testing.

Two types of visual acuity charts are currently in use. The Snellen chart, developed in 1862, is a vertical chart of letters that decrease in size from top to bottom. It contains a different number of letters on each line, and the size changes from one line to the next are variable. This chart can be used to test distance visual acuity at a 10- or 20-foot viewing distance. The Bailey-Lovie chart and later the Early Treatment Diabetic Retinopathy Study visual acuity chart were developed to provide a more standardized method of measuring visual acuity (Kaiser, 2009). These newer charts have an equal number of letters on each line, and the size changes from one line to the next are equal steps on a logarithmic scale, that is, a geometric progression in size. A refined scoring system is also available for these charts. Moreover, the light levels are specified, improving test-retest reliability. These charts, and their projected and electronic equivalents, are thus preferable for accurate determination of visual acuity (Kaiser, 2009; NRC, 2002).

Near visual acuity typically is measured using handheld charts at a distance of 40 centimeters. “If the near vision test chart has the same or similar design features as the letter chart used for distance visual acuity, if other test conditions (luminance, contrast, etc.) are the same, and if the subject is wearing appropriate refractive error correction, then the distance and near visual acuity scores should be equivalent to each other" (NRC, 2002, p. 66).

Peripheral Vision (Visual Fields)

Peripheral vision allows individuals to see objects around them without turning their head. It permits the detection of objects of interest, approaching threats, and the like that are critical for the safe and effective performance of daily activities. Detection of objects in the peripheral visual field is necessary for directing head and eye movements to fixate the objects and thereby inspect them in detail. Also, skills involving visually guided behavior (eye–hand coordination; driving; and other mobility tasks, such as ascending or descending stairs) are heavily dependent on peripheral vision. In addition to the importance of peripheral vision for driving, Marron and Bailey (1982) report that visual field status is one of the most important factors in predicting the mobility skills of individuals with vision loss.

The prevalence of significant loss of visual field due to ocular and/or neurological disorders has been reported by various investigators to be between 1 and 9 percent (Keltner and Johnson, 1980). In a study of 10,000 California drivers, Johnson and Keltner (1983) found that 3–3.5 percent of the population exhibited significant visual field abnormalities. However, the prevalence of visual field loss was much greater in the older population, increasing to 7 percent for individuals aged 60–65 and more than 13 percent for those over 65. Only one-third of individuals with visual field loss demonstrated binocular (both eyes) defects, which were found to be associated with an increased number of driving accidents and convictions relative to those with visual field loss in only one eye. Many subsequent investigations have examined the effect on driving performance of visual field loss produced by ocular and neurological impairment.

Although it is possible to perform visual field testing using manual procedures, such methods are susceptible to high rates of both false-positive and false-negative errors unless performed by specially trained personnel using appropriate equipment. Automated visual field testing is a preferred method of evaluation because it affords a greater degree of standardization and reliability. Automated rapid screening procedures are available for visual field testing, although the cost of such equipment remains rather high.

Depth Perception (Stereopsis)

Stereopsis refers to the process by which the slight differences in the location of images in the two eyes (retinal disparity) result in an impression of depth for objects located at different distances from an observer. It is one of many cues for depth, and is most important for tasks at distances nearer than 4 feet (e.g., threading a needle) where other depth cues are absent. At long distances, stereopsis is only a minor cue for depth, with linear perspective, size, overlapping or relative movement of objects, shadows, and other cues being more important for depth perception. Stereo acuity or stereopsis is measured in terms of the differences in angle between the two eyes and is usually represented in seconds of arc. Under optimal conditions, skilled observers are able to produce thresholds of 3–5 seconds of arc (approximately 10 to 20 times better than for visual acuity). For this reason, stereo acuity is often regarded as a "hyperacuity" function of the visual system. A number of tests (both automated and manual) are available commercially for performing rapid-screening evaluations of stereo acuity.

As noted, stereopsis is just one of a number of cues or sources of depth information that are available to an observer. To the extent that other sources of this information are available, stereopsis probably is not required for performing depth discriminations. Additionally, stereopsis is often quite a slow response function, so it may not be useful for tasks that require rapid visual inspection. Because stereopsis involves the resolution of very fine spatial differences between the two eyes, it is easily disrupted by a wide variety of visual anomalies and is, therefore, an effective screening test for ocular pathology, such as amblyopia. With the exception of such specialized jobs as stereophotogrammetry (e.g., 3D mapping of aerial photographs), there is little information to suggest that stereopsis is a critical visual component of job-related tasks or an important component of activities of daily living and quality of life.

Dark Adaptation and Night Vision

Dark adaptation refers to the increase in sensitivity of the visual system under low illumination or darkness. For most portions of the visual field, the dark adaptation process is described by a two-part function. During the initial 5–10 minutes of dark adaptation, the cone photoreceptor system (used for daylight vision) achieves its maximum sensitivity. This is followed by an additional increase in sensitivity produced by recovery of the rod photoreceptor system used for nighttime vision (30–40 minutes for full dark adaptation). Dark adaptation is an important visual function that may have implications for visual performance in dimly illuminated environments. However, its measurement is difficult for large populations because it requires a large amount of time (about 45 minutes per person), careful control of preadapting test conditions, a completely "light-tight" test facility, precise calibration of equipment, and personnel with training in this type of vision testing. (A review of dark adaptation and all aspects of night vision is available in NRC [1987].)

Dark adaptation and night vision are known to be compromised in a number of eye diseases, in particular those preferentially affecting rod function, such as retinitis pigmentosa. These patients may be severely disabled in darkened surroundings that may be important for job-related activities. However, job accommodations may be possible, and in more advanced disease, these patients are more severely disabled by the loss of more primary visual functions, such as peripheral field extent and visual acuity. As noted above, moreover, adequate testing of night vision capabilities imposes a number of difficult requirements. In addition, there are significant degradations in night vision sensitivity with age, as well as large individual differences. As a consequence, a rapid, simple test of night vision is not available.

Glare Disability and Glare Recovery

Glare refers to the disruption of vision from the presence of a veiling light source in the visual field, such as the oncoming headlights of an approaching vehicle at night. Light from this source is scattered by the optical components of the eye and degrades contrast of the images of other objects. Problems with glare are typically greater under low-luminance viewing conditions. Glare disability is usually measured in terms of the decrement in visual function (e.g., visual acuity) that occurs in the presence of the glare source compared with visual function in its absence. Glare recovery refers to the amount of time needed to reach visual performance equivalent to the level prior to being exposed to the glare source. Both glare disability and glare recovery are important factors related to driving at night. Glare disability becomes greater in older age groups.

Several problems are associated with the use of glare as a screening test for visual performance. First, there are large individual differences in glare disability at all ages, but especially in older age groups. Second, some tests of glare disability focus on the transient adaptation problems associated with the presence of a glare source (glare recovery), whereas others focus primarily on the presence of a steady glare source. Third, no standard methodologies, stimulus conditions, measurement techniques, or test protocols have been established for glare disability. At least two dozen different devices and techniques are currently used for glare testing, each of which yields different results. This variation also makes it difficult to examine the prevalence of glare disability problems in the general population.

Color Discrimination

Color vision, or the ability to perform discriminations on the basis of wavelength differences among stimuli, is a primary means by which objects are discriminated in the environment. Some visual tasks, such as detection and recognition of traffic lights, warning signs, and taillights and their increased brightness produced by braking, may rely on color vision.

Normal color vision is referred to as trichromacy. Color-deficient observers are described as anomalous trichromats, dichromats, or monochromats, depending on the severity of the impairment. Anomalous trichromats are the most similar to normal color observers, except they differ in various degrees with regard to how color mixtures appear. A number of studies related to color vision and job performance have revealed that, with the exception of certain specialized jobs (e.g., diamond grading, quality control of dyes or paints), individuals with mild to moderate anomalous trichromacies are able to perform job-related color discriminations with little or no problem. Vivid colors such as those used for traffic lights and warning signs can readily be distinguished by individuals with mild to moderate color vision deficiencies; such individuals confuse only very desaturated or pastel colors. Approximately 8 percent of the male population has color vision deficiencies of one form or another; women are much less likely to be affected by such deficiencies, which are most commonly due to genes on the X chromosome. In general, deficient color vision does not appear to be a factor that meaningfully impacts quality of life or activities of daily living, including job-related activities.

Summary

Vision problems affect productivity, activities of daily living, and quality of life. Visual acuity, the most commonly used index of visual capacity, is measured with the use of high-contrast letters. In addition, peripheral vision (visual field status) is one of the most important factors in predicting the mobility skills of individuals with vision loss.

HEARING ASSESSMENTS[4]

Approximately 30 million Americans of working age are estimated to have significant hearing loss in both ears, and an additional 18 million are estimated to have hearing loss in one ear (Lin et al., 2011). The proportion of Americans with hearing loss who are of working age and use hearing aids, cochlear implants, or other assistive technologies has not been definitively reported in the literature. Overall hearing aid market penetration data cited by the President's Council of Advisors on Science and Technology in 2016 suggest that only about 15 to 30 percent of the 30 million adults who could benefit from amplification use a hearing aid (Abrams and Kihm,

[4]This section draws heavily on a paper commissioned by the committee for this study (McCreery, 2018).

2015; Chien and Lin, 2012; Kochkin, 2009). This means that a majority of adults with hearing loss are likely to experience hearing-related difficulty in the workplace. Therefore, functional assessment of hearing is critical to ensure participation, safety, and efficiency in the workplace.

Occupations vary considerably in hearing-related performance requirements and environmental conditions that are likely to affect the assessment methods used to measure functional hearing. Occupational requirements for hearing encompass communication with others, including one-on-one and group conversations, as well as use of a telephone or radio; listening in high levels of background noise; and environmental awareness of sound related to changes in working conditions (Montgomery et al., 2011; Punch et al., 1996; Tufts et al., 2009). Many considerations involved in functional hearing assessment are context dependent. As a result, numerous assessment techniques have been developed to quantify functional hearing ability related to occupational requirements in adults. This section provides an overview of the methods used to assess functional hearing for occupational standards across a wide range of contexts. Annex Table 5-7 lists selected instruments used to measure hearing function.

Considerations for Functional Hearing Assessment

Hearing is a complex process involving multiple levels of sensory processing that are often organized hierarchically (Tye-Murray, 2014). Hearing assessments can be categorized by the level of auditory skills required to complete a given task. Erber (1982) describes four categories of auditory skills for assessment: sound awareness, sound discrimination, sound identification, and comprehension. Sound awareness, the most basic level of auditory function, refers to the ability to detect the presence or absence of sound. Sound discrimination is the ability to determine when a sound is different from another sound or changes over time. Sound identification occurs when a listener is able to label or categorize an auditory signal. Comprehension, the highest level of the auditory skills hierarchy, requires the listener to understand and interpret incoming auditory stimuli. Each level of the auditory skills hierarchy is dependent on the skills lower in the hierarchy. Functional assessment of hearing for purposes of occupational fitness draws on tests that assess each of these auditory skill levels.

Auditory–verbal communication with other people in occupational contexts requires that an individual be able to function at all four levels of the auditory skills hierarchy. This interdependence across levels of auditory skill has led to two distinct approaches to functional hearing assessment related to occupational listening environments. The first approach is foundational in that it measures the individual level of hearing acuity and attempts to estimate the impact of environmental factors (such as levels of background noise or reverberation) on higher-order auditory skills using models that predict performance in realistic listening environments (e.g., Soli et al., 2018a). The second approach attempts to measure an individual's auditory performance using tasks that directly involve higher-order auditory skills, such as identification or comprehension. Such approaches may also entail recreating elements of a real-world listening environment in order to measure functional hearing as it would be used for specific employment tasks (e.g., McGregor, 2003).

Foundational approaches to functional hearing assessment are useful at predicting functional hearing abilities at the group level (Soli et al., 2018a), but also may have difficulty predicting performance at the individual level or for highly specific work environments. Because such approaches do not measure higher-order auditory skills directly, they may not be sensitive to auditory problems that have isolated effects on higher-order auditory functions. Direct measurement of higher-order auditory skills, on the other hand, can also be limited by their specific characteristics. For example, a functional assessment of hearing on a radio or telephone may not generalize to other occupational activities, such as face-to-face communication or detection of alarms or auditory signals from equipment. For these reasons, functional hearing assessments related to employment often combine both types of approaches to reflect both hearing abilities that apply more generally to overall functioning in the workplace and those applicable to tasks that may be highly specific to an occupation.

Functional Hearing Assessments

The order in which functional hearing assessment methods are described here is based on the hierarchy of auditory skills discussed above (Erber, 1982), from awareness of sound to comprehension-level tasks.

Pure-Tone Audiometric Screening

Hearing screening with pure-tone stimuli at the frequency range most important for speech understanding is a common method for determining the presence or absence of hearing loss in adults (McBride et al., 1994; Yueh et al., 2003). Pure-tone screening generally involves presenting one or more pure-tone signals in the frequency range for speech (500–4000 Hz) at a single level in each ear. If the listener responds to the sound at that level, he or she passes the screening; if not, the listener is referred for further diagnostic assessment. Pure-tone hearing screening can be accomplished by physicians, nurses, or trained medical assistants in an office setting using handheld or portable devices. Its sensitivity ranges from 0.96 for detecting high-frequency hearing losses to 0.64 for detecting low-frequency hearing losses (McBride et al., 1994). Sensitivity increases as the number of different frequencies in the screening test is increased. The specificity of pure-tone screening ranges from 0.77 to 0.89 across studies (Lichtenstein et al., 1988; Yueh et al., 2003).

The advantages of pure-tone hearing screening are that it requires minimal expertise to administer and interpret, can be conducted in a quiet office setting, and can be administered in just a few minutes. The screening also can be administered to listeners with any language background. The disadvantages of pure-tone hearing screening are that it does not provide information about the type, degree, or configuration of hearing loss and may be contaminated by ambient room noise, particularly at low frequencies.

Otoacoustic Emissions Screening

Otoacoustic emissions (OAEs) are an acoustic response that originates from the outer hair cells in the healthy cochlea within the inner ear. The OAE response is measured using a soft probe placed in the ear canal that presents auditory stimuli to the ear and uses a microphone to measure the OAEs that occur in response. Ears with measured OAEs are considered to have normal cochlear function, whereas absent OAE responses are associated with at least mild cochlear hearing loss. OAE screening has been widely used in hearing screening for newborns (e.g., Gorga et al., 1997) but has also been proposed as an objective hearing screening test for adults for occupational and medical screening purposes (Hotz et al., 1993; Wang et al., 2002). OAEs can be measured in a few minutes per ear by a physician, nurse, or trained medical assistant in a quiet office setting.

In adults, the sensitivity of OAE screening for hearing loss ranges from 0.91 to 0.98, with specificity ranging from 0.62 to 0.86 (Wang et al., 2002). The advantages of OAE screening for adults are that it is an objective measure that requires no response from the listener and can be administered by a trained medical assistant in a quiet office setting. The disadvantages are that the absence of OAE responses does not provide information about the type or severity of hearing loss, OAEs do not predict functional hearing ability in real-world environments, and the screening can be affected by middle-ear dysfunction or contaminated by ambient room noise.

Assessment of Localization

The ability to determine the location of sounds in a listening environment can be an important functional hearing ability for specific occupations. Additionally, protective equipment may reduce environmental awareness of sound, including the ability to determine its location. Localization tasks spatialize sound in a test environment using multiple sound sources or speakers. Listeners must respond by indicating the direction of sounds, typically along the azimuthal plane (Abel et al., 2009; Przewozny, 2016). Localization tests must be administered by an audiologist or scientist with expertise in spatial hearing in a sound-treated audiometric test booth or anechoic chamber. Localization tasks can be administered using loudspeakers or under headphones using head-related transfer functions (Wenzel et al., 1993).

The test-retest reliability of localization tasks is reported to be 0.8 (Häusler et al., 1983), with localization differences ranging from 1 to 3 degrees. Localization tasks have advantages that include high ecological validity for spatialized sound in real-world conditions and assessment of the ability to locate sound in space. However, localization tasks have not been widely adopted because of technical challenges in their implementation, the requirement for specialized equipment and spaces, and the lack of commercially available tests or software with which to administer the tasks.

Diagnostic Pure-Tone Audiometry

Pure-tone audiometric threshold assessment is currently the gold standard for functional hearing assessment (ASHA, 2005). Pure-tone audiometry varies the level of pure-tone signals at multiple frequencies to determine the threshold of hearing for the frequency range of speech. Signals can be presented through air and bone conduction to help determine the specific site of lesion for hearing loss. Unlike pure-tone screening, which results only in a pass or refer result, diagnostic pure-tone audiometry makes it possible to determine the type (conductive, sensorineural, or mixed), degree (mild, moderate, severe, or profound), and configuration (flat or sloping) of hearing loss.

Diagnostic pure-tone audiometry is typically conducted by a licensed audiologist in a sound-treated audiometric test booth (ASHA, 2005). Its test-retest reliability is 0.92, with differences of 3–5 decibels (dB) between repeated tests (Roeser et al., 2000). Pure-tone audiometry has advantages that include the specification of hearing sensitivity at multiple frequencies important for communication. In addition, the degree of hearing loss determined from the audiogram has been shown to predict a range of outcomes in real-world listening environments (e.g., Woods et al., 2013). However, the disadvantages of this method are also notable, including the fact that audiometric thresholds do not measure auditory skills above the level of detection, and are unlikely to reflect listening performance in background noise or with reverberation without additional interpretation. Moreover, because pure-tone audiometry is most often conducted by an audiologist in a sound-treated booth, it may be impractical in some situations, limiting access for specific populations.

Extended Speech Intelligibility Index Methods

The speech intelligibility index (SII) (ANSI S3.5-1997) is a standardized method for predicting the proportion of acoustic information in a speech signal that is audible to a listener in a given situation. The standard approach to estimating SII speech audibility involves a talker who is facing the listener at 1 meter away, speaking with normal vocal effort and in a quiet listening environment. Although this standard approach has been useful for estimating the effects of hearing loss and noise on a listener's access to speech information, recent research has focused on extending the SII estimates into real-world listening environments that workers may encounter on the job (Soli et al., 2018a,b). The extended SII (ESII) methods resulting from this work have been validated to predict speech recognition for adult listeners with normal hearing and those with hearing loss across a range of plausible acoustic environments derived from occupational studies of noise (Soli et al., 2018a). These methods extend the information gathered from the audiogram by incorporating information about ear canal acoustics, background noise and reverberation, and talker-level effects.

The test-retest reliability of ESII methods ranges from 0.78 to 0.97, depending on the listening situation being simulated. The 95 percent confidence interval of the speech recognition prediction obtained with these methods is approximately 7 to 14 percent (Soli et al., 2018a). Because these methods are based on the audiogram, their advantages and disadvantages are similar to those of pure-tone audiometry. An additional disadvantage is that the software for calculating the ESII was not available at the time this report was written.

Speech Recognition in Noise Testing

Speech recognition in noise is an important assessment for determining a listener's ability to identify speech and communication in real-world environments with background noise. Listeners repeat back words or point to pictures that are presented in a background of noise (Giguère et al., 2008; Laroche et al., 2003). A wide range of different stimuli sets, including nonwords, monosyllabic words, and sentences, are available as recorded materials, including tests in multiple languages other than English. The background noise can consist of steady-state noise (designed to simulate noise from heating, ventilation, and air conditioning systems) or multitalker babble (designed to simulate noise from other people talking). The testing is most often completed under headphones or through speakers by an audiologist in a sound-treated audiometric test booth.

The test-retest reliability for recorded speech recognition materials in noise is above 0.85 (Laroche et al., 2003). The testing generally takes about

15 minutes or less in a clinical setting to provide enough trials for reliable results. The advantages of speech recognition in noise tests are that they have strong face validity because, unlike pure-tone audiometry, they require the listener to perform a task critical to real-world listening and use stimuli likely to be encountered in everyday life. Their disadvantages include that performance on the tests is dependent on the language proficiency of the listener and the tester. Additionally, because the testing is typically conducted by an audiologist in a sound-treated audiometric test booth, some individuals may lack access to this type of assessment.

Live-Voice Speech Testing

Because audiometric test equipment and facilities are not always available, physicians and other medical providers have used live-voice speech testing to screen for hearing loss in medical office settings. The person administering the test often stands behind or out of view of the listener and says words or phrases at an average or soft speech level. The listener must accurately repeat the words or phrases back to pass the screening. One study found sensitivity of 100 percent for detecting hearing loss and specificity of 84 percent for whispered speech in an office setting (MacPhee et al., 1988). Live-voice speech testing can be administered in only a few minutes in the native language of the listener. The advantages of live-voice speech testing are its ease of administration and the lack of requirements for expensive equipment or spaces. This method also has many disadvantages. Because the live voice cannot be calibrated, there is likely to be substantial variability among and within talkers across tests that threatens the method's reliability. Live-voice speech testing is also influenced by the dialect of the talker and ambient room noise. There is no evidence at this time to suggest that this method can predict communication or hearing in occupational environments with background noise and reverberation.

Internet and Telephone Hearing Screening

Efforts to increase access to functional hearing assessments for individuals not in close geographic proximity to hearing health care providers has led to the development of Internet and telephone methods (Laplante-Lévesque et al., 2015; Smits et al., 2004, 2006; Watson et al., 2012). In most cases, such tests present speech to a listener either via Internet connection to a computer or over a terrestrial phone line. The listener registers a response either by typing a response into the computer or speaking. Results indicate a strong correlation with results of pure-tone diagnostic tests (0.73) and sensitivity for significant hearing loss of 0.80, with specificity of 0.83 (Watson et al., 2012). The advantage of Internet and telephone hearing screening is in providing access to a hearing screening for anyone with an Internet connection or telephone. Telephone-based hearing tests have the disadvantage of not currently being validated for cellular telephones, although this weakness is mitigated by the fact that many cellular phones have data connections that allow for an Internet-based application of the test. Additionally, as with other screening methods, the results of Internet and telephone hearing screening do not provide direct information about occupational function in real-world listening environments, only whether hearing is normal or abnormal. Results may also be affected by the quality of the transducers if the listener is using a computer.

Hearing Questionnaires

Questionnaires have been developed for assessing functional hearing in occupational settings. Published data are available for two questionnaires for adults in the workplace: the Hearing Handicap Inventory for Adults (HHIA) (Newman et al., 1990) and the Speech, Spatial, and Qualities of Hearing Questionnaire (SSQ) (Gatehouse and Noble, 2004). The HHIA has 25 items and can be self-administered. It has been shown to be related to both pure-tone audiometric data and speech recognition abilities (Newman et al., 1990). Its internal reliability based on Cronbach's alpha is 0.93, and its test-rest reliability is reported as 0.97 (Newman et al., 1991). The SSQ has multiple versions. One of these, the SSQ5, is a five-item version that has been used to assess hearing loss in individuals of employment age (Demeester et al., 2012). The SSQ has reported test-retest reliability of 0.83 for interview-based administration and 0.73 for self-administration (Singh and Pichora-Fuller, 2010). The advantages of both the HHIA and SSQ questionnaires are that they can be used to assess a person's self-perceived hearing difficulty in ecologically relevant employment contexts. Their disadvantages are that they may not always be available in the language of the listener and may not assess a wide range of work-related hearing challenges that are relevant to specific occupations.

Effects of Environmental Conditions on Hearing

The environmental conditions of the work environment have significant effects on individuals' hearing for the purposes of communication. Of these conditions, the presence of noise or reverberation has the most significant negative effects. Noise is ubiquitous in most work environments, but the range and type of background noise vary considerably. Proximity to moving mechanical parts and heavy vibration affect verbal communication primarily because both are associated with an increased noise intensity level in the workplace. The presence of noise has negative effects not only on the ability to communicate but also on the ability to store and process auditory information that is heard.

The ideal listening environment for verbal communication at work is a quiet one with minimal background noise. Noise intensity levels of 45 decibels (acoustic) (dBA) or less provide optimal signal-to-noise ratios for face-to-face and small-group communication. Moderate noise levels in the work environment (45–70 dBA) are likely to have a negative impact on one-on-one and group communication, particularly as the level of the background noise increases. Loud working environments (70–85 dBA) are likely to have a significant negative impact on communication, including one-on-one communication, group meetings, and communication by radio or telephone (see Soli et al., 2018b, for examples). Communication in very loud environments (>85 dBA) is difficult across all verbal communication contexts. The level of the noise in such environments is likely to require hearing protection devices, preventing any verbal communication or awareness of alarms or auditory signals. In addition to the effects of noise on verbal communication, limited environmental awareness of noise is a significant safety risk. Picard and colleagues (2008) report that up to 12 percent of workplace injuries in a large study of more than 40,000 claims could be accounted for by the interaction of hearing-related factors and background noise.

Effects of Hearing Loss on Work-Related Hearing Function

The presence of hearing loss can present particular challenges for communication in the workplace (Jennings and Shaw, 2008). Even with workplace accommodations, hearing loss can reduce participation in employment by 10 to 15 percent (Hogan et al., 2009).

The degree of hearing loss identified from a pure-tone audiogram can serve as a reasonable predictor of difficulty with workplace communication. Although individual performance varies considerably based on the degree of hearing loss, use of hearing technology, and occupational hearing requirements, problems with communication in both quiet and noisy environments increase as the degree of hearing loss increases (Kramer et al., 2006). A mild degree of hearing loss (25–45 decibels hearing level [dB HL]) is likely to have a minimal effect on one-on-one communication or communication by radio or telephone for adults, with performance near 100 percent accuracy. Such individuals will not have difficulty communicating in meetings or even in the presence of background noise. A moderate degree of hearing loss (50–65 dB HL) is often associated with a decrease in the accuracy of speech recognition of 20–30 percent (Dubno et al., 1984; Humes, 2002) and increased effort for one-on-one communication and communication by radio or telephone, along with significant difficulty communicating in the presence of background noise. A severe degree of hearing loss (70–90 dB HL) is associated with a limited ability to understand words or sentences for communication purposes. Even in quiet environments or with visual cues, recognition is often less than 50 percent for such individuals (Dubno et al., 1984; Grant et al., 1998). A profound degree of hearing loss (greater than 90 dB HL) is associated with limited ability for verbal communication in either quiet environments or those with background noise. People with mild to severe hearing loss can improve communication by using hearing aids, while those with severe or profound hearing loss can improve their auditory access through cochlear implants.

Predictions of workplace performance can be enhanced by using estimates from the ESII (Soli et al., 2018a,b) to predict speech recognition in specific work environments. Other functional hearing assessments may serve as indicators for occupational challenges related to hearing. Poor understanding of speech, even with the use of hearing aids or cochlear implants, may be an indication that an individual listener will have difficulty performing hearing-related occupational functions. Clinical assessments of speech recognition are conducted under controlled conditions, which means that these scores represent an optimistic estimate of real-world communication performance (Giguère et al., 2008). High scores on questionnaires, such as the HHIA (Newman et al., 1990), can indicate that an individual is likely to experience communication difficulty in real-world environments, including the workplace. While the presence of hearing loss increases challenges related to occupational function, the use of hearing aids, cochlear implants, or hearing assistance technology should lower HHIA scores over time and can help increase workforce participation for people with hearing loss.

Summary

Functional hearing assessments related to employment often combine two approaches to reflect both hearing abilities that apply more generally to overall functioning in the workplace and those applicable to tasks that may be highly specific to an occupation. Currently, pure-tone audiometric threshold assessment is the gold standard for functional hearing assessment. The assessment can serve as a reasonable predictor of difficulty with workplace communication, because it makes it possible to determine the type (conductive, sensorineural, or mixed), degree (mild, moderate, severe, or profound), and configuration (flat or sloping) of hearing loss. Speech recognition in noise testing is an important assessment for determining a listener's ability to identify speech and communication in real-world environments with background noise. Hearing questionnaires such as the HHIA and the SSQ can be used to assess a person's self-perceived hearing difficulty in ecologically relevant employment contexts; however, they may not assess a wide range of work-related hearing challenges that are relevant to specific occupations.

SPEECH AND LANGUAGE ASSESSMENTS[5]

Communication disabilities can profoundly influence a person's life, impacting his or her ability to work, interact, and engage with others. The American Speech-Language-Hearing Association (ASHA) describes a communication disorder as "impairment in the ability to receive, send, process, and comprehend concepts or verbal, nonverbal and graphic symbol systems" (ASHA, 1993). In the context of functional abilities relevant to work requirements, an individual's comprehension of language produced by the communication partner, processing of that language, expression of his or her own ideas, and pragmatic interactions in dialogue appropriate to the work context are all relevant to assessing such disorders for purposes of disability determination. The ORS Collection Manual definition of speaking encompasses skills necessary for "expressing or exchanging ideas by means of the spoken word to impart oral information to clients or the public and to convey detailed verbal instructions to other workers accurately, loudly, or quickly" (DOL, 2018, p. 125). Based on this definition, specific aspects most relevant to assessing a person's functional communication in an occupational setting include receptive-expressive language (i.e., to impart information or exchange ideas) and speech production (i.e., accuracy, loudness, and speed). A person's ability to participate in certain work settings may include other aspects of functional communication (e.g., alternative expressive modalities, nonverbal interactions, written language, social communication). This section focuses on speech and language assessment for the purposes of determining and describing functional communication skills, with particular attention to employment-related interactions. Selected instruments for assessing speech and language function are listed in Annex Tables 5-8 and 5-9, respectively. This section provides an overview of functional communication, followed by a review of assessment instruments for the physical and mental components of such communication.

Understanding Functional Communication

Functional communication comprises both speech (i.e., verbal speech production) and language (i.e., comprehension and/or production of phonology, morphology, syntax, semantics, pragmatics) mechanisms (ASHA,

---

[5]This section draws heavily on a paper commissioned by the committee for this study (Ball, 2018).

n.d.). Given the physical planning and execution of oral-facial movements to produce speech, speech is considered a physical process for purposes of this discussion; similarly, language involves cognitive processing and formulation and therefore is considered a mental process.

There are many different definitions for functional communication. One definition of functional communication as the "ability to communicate basic physical needs and emotional states" has been used by funding agencies attempting to limit resource outlays (Elman and Bernstein-Ellis, 1995). In some cases, this practice has resulted in discontinuing treatment when an individual simply achieves the ability to express very basic needs, but has not yet advanced to be capable of more complex occupation-specific communication (Elman and Bernstein-Ellis, 1995). In such cases, much of the information available in assessment and treatment records does not directly address the individual's functional abilities for work, yet clearly these individuals are likely to have a restricted employment outlook.

The ORS Collection Manual (DOL, 2018) implies that functional communication involves effective communication skills in natural environments (e.g., with respect to setting [indoors, outside, noise, light]) with typical communication partners (e.g., one-on-one, groups). Effective communication yields the intended result or response and, particularly in many employment settings, is efficient in getting the message across (Hustad, 1999). Components of functional communication in an employment situation include (1) the physical and cognitive characteristics of the people interacting; (2) the content of messages and how they are represented and conveyed (e.g., face-to-face, by phone); (3) conditions in which messages are transmitted and understood (e.g., immediate or delayed timing, noise or quiet, bright or dim light, office or warehouse); (4) societal relationship (e.g., familiar or unfamiliar, boss or employee, educator or student); and (5) interaction purposes (e.g., distributing information, expressing attitudes, social relations) (Blackstone et al., 2007; Light, 1988).

Myriad factors contribute to functional communication, including personal (individual) skills, environmental conditions, and the skills of communication partners. At the level of the individual, physical factors specific to communication (articulation accuracy, speaking rate, voice quality, loudness, fluency, effort, and fatigue) affect the intelligibility and comprehensibility of speech production. Mental factors specific to communication at the individual level (receptive, expressive, pragmatic language skills) affect message comprehension. And because communication is a dynamic, transactional process that involves at least two people, characteristics of communication partners (e.g., personal factors, sensory skills, motivation) and other external factors (e.g., environmental conditions, adaptations) also play a critical role in functional communication. These factors, along with the anticipated course and severity of the communication impairment, whether physical or mental in nature, will impact assessment results and need to be considered in determining current and longitudinal employment expectations.

At present, insufficient research is available regarding assessment of functional communication; therefore, limited standardized measures exist for this purpose. The descriptions of assessment instruments that follow were derived from the few such instruments that have been published and the literature. These instruments, which reflect standard clinical practices in the field of speech-language pathology, are often supported by only a few studies. Speech-language assessments are performed by speech-language pathologists (SLPs), who have expertise in assessment, differential diagnosis, and treatment of communication disorders. These assessments may be completed for many purposes (e.g., detect and describe a problem, establish diagnostic options, establish a diagnosis, specify severity, determine implications and functional impacts) (Duffy, 2013). They may include observation of interactions, standardized and/or criterion-referenced tools, instrumentation (e.g., endoscopy, aerodynamics, ultrasound), prior test results, and history (ASHA, 2016). A key consideration in assessment is to examine the individual's ability to successfully access test items. Attempts to adapt various testing materials for persons with physical disabilities may introduce errors in measurement, thus decreasing validity and reliability. Although Ratcliff (1994) and White and colleagues (2010) studied children, the research provides legitimate considerations for adults. That is, changing the way an individual accesses a test instrument may impose additional cognitive and physical challenges that can impact performance and lead to an inaccurate estimation of actual ability.

Clearly, a major factor in functional communication is the severity of an individual's communication impairment. Communication impairments, whether physical or mental in nature, can range in severity from mild to profound. Some assessment instruments with standard protocols include a description of severity, while for others, severity is determined by percentage of occurrence, by levels of impact on daily activities, and in some cases by the individual's need for additional supports. Often, severity is rated by the SLP evaluator, the individual, and communication partners using a scale ranging from 0 (normal, minimal disturbance) to 6 (extreme/aphonic). Severity is discussed in greater detail later in this section.

Assessment of the Physical Components of Functional Communication

Any component of speech impairment may result in impaired functional ability to communicate verbally in the context of work. The focus in this section is on those physical aspects of functional communication most relevant to adults communicating in a work setting: speech (i.e., sound accuracy), voice (i.e., quality, loudness), and fluency (i.e., smooth flow).

Speech

Common measures of speech sound production focus on body structure and function aspects of speech impairments that are used to establish a diagnosis and identify targets for intervention (WHO, 2002). At present, the gold standard for speech assessment involves auditory-perceptual analysis of productions by an SLP experienced with these impairments and the associated assessment procedures (Duffy, 2013). Although this assessment approach is dependent on the SLP's perceptual skills, evidence supports its effectiveness in differential diagnosis and severity scoring, with high intrarater reliability (Mumby et al., 2007). Annex Table 5-8 lists published tests that are effective for evaluating the accuracy of speech sound productions in various contexts, including

- Frenchay Dysarthria Assessment (Second Edition) (FDA-2) (Enderby, 1983; Enderby and Palmer, 2008),
- Dysarthria Examination Battery (Drummond, 1993),
- Quick Assessment for Dysarthria (Tanner and Culbertson, 1999),
- Apraxia of Speech Rating Scale (Strand et al., 2014), and
- Apraxia Battery for Adults (Second Edition) (Dabul, 2000).

Although the results of this auditory-perceptual assessment contribute to understanding an individual's speech impairment, the extent of the impairment's impact on functional communication and work participation can be judged based on measures of intelligibility (i.e., decontextualized words, sentences) (Ishikawa et al., 2017; Kent et al., 1989; McAuliffe et al., 2010; Patel et al., 2014; Stelzle et al., 2013; Yorkston et al., 1992); efficiency (i.e., rate of intelligible/comprehensible utterance production) (Przysiezny and Przysiezny, 2015; Ross and Wertz, 2003; Yorkston and Beukelman, 1981); and comprehensibility and listener comprehesion (i.e., words, sentences, discourse in context) (Hustad, 2008). Together, these features of speech contribute to communication effectiveness, including in a work context.

Intelligibility and efficiency Sentence intelligibility assessment tasks are typically concerned with quantifying the severity of a speech impairment or level of performance (Yorkston et al., 1992). Clinically defined as the extent to which a listener understands a person's speech, intelligibility is a critical measure of the severity of a communication disorder (Yorkston et al., 2010). Assessment methods used to estimate levels of intelligibility include direct amplitude estimation, Likert-type scales, and categorical estimation

(McAuliffe et al., 2010). Evidence suggests that estimation procedures yield less accurate and more variable results among speakers relative to word-by-word transcription measures of intelligibility, which yield more consistent results across speakers (Hustad, 2006). Intelligibility has a ceiling and floor effect; thus it is limited in sensitivity to a narrow range of severity (Yorkston and Beukelman, 1981). A broader index of speech performance is obtained by combining measures of speaking rate and intelligibility, which helps illuminate the complex interaction necessary for functional communication (Yorkston and Beukelman, 1981; Yorkston et al., 1992).

A measure of sentence production will provide a decontextualized assessment of speech performance. Published measures of intelligibility for literate English-language speakers (see Annex Table 5-8) include the Assessment of Intelligibility in Dysarthric Speech (Yorkston and Beukelman, 1981) and its computerized version, the Speech Intelligibility Test (Yorkston et al., 2007). Both versions quantify word and sentence intelligibility (percent of intelligible words) and speaking rate (words per minute), and provide a measure of communication efficiency (ratio of intelligible words per minute). Intelligibility is calculated by dividing the number of correctly identified words by the number of possible words (Yorkston et al., 1996, 2007). The reliability of this instrument has been established for interrater agreement (score dispersion, variance), with reliability coefficients of 0.93 to 0.99 for intelligibility and 0.99 for rate of intelligible speech (Yorkston et al., 2007).

Comprehensibility and listener comprehension Different from intelligibility, comprehensibility is the extent to which a listener understands the speech produced when provided with all available additional information that may add to understanding, or within context (Duffy, 2013). Although measures of intelligibility illustrate the severity of communication impairment, the direct relationship of intelligibility with functional communication remains unclear, as it has not shown to be a significant predictor of the effectiveness of communication (Donovan et al., 2008). The addition of comprehensibility measures to intelligibility measures can provide a more thorough understanding of the "information-bearing capability" of a person's speech (Hustad, 2008).

Comprehensibility assessment yields a valid measurement of the functional participation impacts of a speech impairment (Duffy, 2013). One method for assessing comprehensibility is to provide two transcriptions: first, the individual is recorded saying a list of prearranged utterances, and a transcriptionist transcribes this recording for intelligibility and speaking rate efficiency (i.e., intelligible words per minute); second, the transcriptionist is provided with a contextual statement for each spoken utterance, and then transcribes for comprehensibility and speaking rate efficiency (i.e., comprehensible words per minute).

Similarly, listener comprehension tasks require listeners to decode speech signals and subsequently process the linguistic information (Marslen-Wilson, 1989). Listener comprehension tasks gauge a communication partner's ability to interpret the meaning of messages produced without regard for speech accuracy (e.g., regardless of inaccuracy, dysfluency, or reduced loudness). Listener comprehension is calculated by the listener's ability to answer questions about or summarize message content (Hustad and Beukelman, 2002). As yet, there is no known standardized assessment protocol for measuring comprehension or comprehensibility, nor is there a definitive metric for determining the threshold of comprehensibility for functional interactions. One strategy for obtaining this information is for the evaluator to devise a set of content-based questions about a passage or set of utterances and ask a listener unfamiliar with the material to respond to the questions after listening to the speaker relate them (Hustad and Beukelman, 2002). Comprehension is quantified by dividing the number of correct responses to content questions by the possible number correct (e.g., (# correct responses/# content questions) × 100 = % comprehension).

Communication effectiveness The Communication Effectiveness Index, developed to assess functional communication among individuals with aphasia, can be used to evaluate progress toward the recovery of communication ability (Lomas et al., 1989). This tool was subsequently adapted for use by individuals with amyotrophic lateral sclerosis (ALS) and found to have high internal test reliability ($r = 0.97$) and strong correlation ($r^2 = 0.80$, $p < 0.01$) between ratings obtained from individuals with ALS and their communication partners (Ball et al., 2004). The adapted index, the Communication Effectiveness Survey, was recently found to have construct validity and statistical significance (Donovan et al., 2008). It also identified differences between individuals with speech disorders and with no speech disorder (Donovan et al., 2008). Annex Table 5-8 includes other available measures that can be used to assess communication effectiveness.

Voice

As with assessment of speech, elements of a clinical voice assessment generally focus on body structure and functional aspects of voice. Less helpful for assessment of overall performance are analyses of discrete subsystems (pitch/frequency, loudness/intensity, quality/sound wave complexity, duration/respiratory-phonatory control, and muscle tension). One purpose of the SLP evaluation is to assess voice production and determine the impact of impairments on daily interactions (ASHA, 2004).

The variability of assessment procedures has led to a recent attempt to standardize protocols for acoustic and instrumental assessments of vocal function (ASHA DIV3, n.d.; Patel et al., 2018; Roy et al., 2013). Instruments are now available for measuring the severity of voice impairment (Awan et al., 2009). The Cepstral Spectral Index of Dysphonia is a multifactorial assessment of voice severity in connected speech. This index discriminates normal from impaired voice, and its results correlate with a visual analog scale rating of overall voice impairment severity on the Consensus Auditory-Perceptual Evaluation of Voice (Awan et al., 2010, 2016; Patel et al., 2018; Peterson et al., 2013; Zraick et al., 2011). As previously discussed, most commonly used clinical assessment procedures involve auditory-perceptual evaluation of voice and may include categorical ratings (mild, moderate, severe), equal appearing/visual analog scales, or direct magnitude estimation (see Table 5-2).

A recent review of patient-report measures identified a group of measures addressing voice impairments from a broad perspective (i.e., some targeting specific disorders, such as spasmodic dysphonia) with strong psychometric characteristics. The integration of these measures into voice assessment provides supportive evidence of the consequences of voice disorders directly from the affected individual (Francis et al., 2017). Annex Table 5-8 contains selected measures addressing general voice impairments.

Fluency

Fluency in speech refers to continuity, smoothness, rate, and effort. Stuttering is a disorder in the rhythm of speech in which the individual knows what he or she wants to say but is unable to say it because of involuntary, repetitive prolongation or cessation of sound (Andrews et al., 1983). The disorder may include repetitions (sounds, syllables, words, phrases), word avoidance, sound prolongations, blocks (difficulty starting a sound), and interjections of unnecessary sounds in speech (Craig et al., 1996). Many assessment instruments focus on discrete aspects of speech (e.g., sound repetitions, number of prolongations), although they also often involve a greater focus on activity and participation interactions relative to assessments of other aspects of speech. Fluency assessment involves auditory-perceptual analysis of productions and is completed by an SLP experienced with stuttering. Less helpful for assessment of overall performance are analyses of discrete characteristics (percentage of stuttered syllables, stutter duration, and muscle tension). One purpose of the SLP evaluation is to assess fluency of speech and determine the impact of impairments on daily interactions (ASHA, 2004). Individuals who stutter are at a higher risk of experiencing difficulty obtaining and maintaining

TABLE 5-2
Standard Voice Assessment Procedures, Speech Tasks, and Anticipated Results

	Procedure	Patient Task(s)	Assessment Results	ICF Level
Instrumental	Laryngeal imaging (endoscope)	Inhale-exhale cycles Produce [ʔiʔiʔiʔi] [i:]-sniff or [i:]-quick inhale Sustained [i:]	Vocal fold edges, mobility, vertical level, closure duration Supraglottic activity Regularity, amplitude Mucosal wave movement Glottal closure, phase symmetry	BSF
	Acoustic analysis	Standard reading passage Vowel glide [a] Sustained vowel [a]	Vocal sound level, frequency, range, noise (cepstral peak prominence)	BSF ACT
	Aerodynamic analysis	Produce [pi:pi:pi:pi]	Glottal airflow rate, pressure Mean frequency, vocal loudness	BSF ACT
Noninstrumental	Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)	Sustained vowel [a, i:] Sentence production Spontaneous speech	Severity, roughness, breathiness, strain, pitch, loudness	ACT
	Voice Outcome Survey	Patient report	Emotional, functional, physical aspects of voice	PAR
	Voice-related quality of life	Patient report	Physical, socioemotional aspects of voice	PAR

NOTES: [ʔiʔiʔiʔi] refers to production of a glottal stop–vowel combination; [pi:pi:pi:pi] refers to production of a consonant–vowel combination; [a] and [i:] refer to vowel productions. ACT = activity; BSF = body structure and function; ICF = *International Classification of Functioning, Disability and Health*; PAR = participation.
SOURCE: Assessment procedures adapted and summarized from Patel et al., 2018.

employment (Bloodstein and Bernstein Ratner, 2008; Craig and Calver, 1991; Craig et al., 2009; Klein and Hood, 2004).

Patient-report measures of stuttering with strong psychometric characteristics have been identified (Craig et al., 2009). One example is the Overall Assessment of the Speaker's Experience of Stuttering, designed to estimate the impact of stuttering on quality of life (Constantino et al., 2016; Siew et al., 2017) (see Annex Table 5-8). The integration of these measures into an assessment provides supportive evidence of the impacts of voice disorders directly from the affected individual (Francis et al., 2017).

The intelligibility and efficiency of speech sound production, and thus functional communication, are affected by effort and fatigue from the perspective of both listener and speaker. For example, although an individual may produce completely intelligible speech via a tracheoesophageal prosthesis following a total laryngectomy (Iverson-Thoburn and Hayden, 2000; Singer and Blom, 1980), the differences between the speech produced in this manner and natural speech require increased listener effort (Nagle and Eadie, 2012). A relationship has been identified between attention allocation (i.e., the amount of effort a listener expends in a conversation) and intelligibility, with the highest levels of attentional need or focus being associated with intelligibility in the range of 75–80 percent (Beukelman et al., 2011). Perceived listening effort (i.e., the amount of effort, attention, or concentration required to understand a speech sample) has a strong negative correlation with intelligibility, and therefore may be useful as an outcome measure (Nagle and Eadie, 2018). Although data are sparse on this question, this measure may serve to supplement objective speech measures of accuracy, intelligibility, and efficiency by illustrating the level of effort involved to sustain performance for functional interactions within a given workplace. Perceived listening effort can be quantified by asking a listener to indicate the effort required on a 10-centimeter vertical visual analog scale marked at the endpoints (0 = very little effort, 10 = extreme effort) after responding to transcription and comprehension questions.

Fatigue may also be experienced by speakers with speech impairments. This is particularly the case for individuals with degenerative disease (e.g., ALS, multiple sclerosis, myasthenia gravis), who may experience rapid declines in speech subsystem performance when in adverse speaking conditions or with repeated or lengthy interactions. In addition to stress testing completed by an SLP, a patient-reported measure of fatigue associated with communication may be included in the assessment process. A strategy similar to that used to measure listener effort can be used for this purpose, with the speaker being asked to indicate the effort required on a 10-centimeter vertical visual analog scale marked at the endpoints (0 = very little effort, 10 = extreme effort) after completing specific speech production tasks.

The Levels of Speech Usage instrument is used for assessment of speech requirements to meet daily communication needs, focusing on variables necessary for functional communication and shifting needs across daily activities (Baylor et al., 2008). With this tool, individuals report the frequency, type, amount, and perceived importance of daily speaking situations (Anderson et al., 2016; Baylor et al., 2008). The results provide categorical ratings of the individual demands of speech (i.e., undemanding, intermittent, routine, extensive, or extraordinary speech usage) (Baylor et al., 2008). Working for pay, time spent talking at work, and education levels are strongly associated with speech usage; in one study, the majority of participants who worked for pay indicated that speech was very to extremely important to their work (Anderson et al., 2016; Gray et al., 2012). Annex Table 5-8 includes available measures that can be used to assess the level of effort and/or fatigue associated with speech production.

Mental Components of Verbal Communication

The mental components of verbal communication addressed here include language skills. Any aspect of language impairment may result in impaired functional ability to communicate verbally within the context of work. The focus of this section is on those aspects most relevant to adults communicating in a work setting: receptive, expressive, and pragmatic language.

Receptive, Expressive, and Pragmatic Language

Language is considered a system "of signs or symbols used according to prescribed rules to convey meaning" (Kent, 1998). Language impairment is generally categorized as a set of specific disabilities impacting a person's ability to receive, understand, process, produce, and respond to language. Two presentations of language impairment are identified among adults, based primarily on the type of impairment. The first is associated with intellectual disability and/or developmental delay and involves chronically impaired acquisition of language (i.e., receptive, expressive, pragmatic). An example of this type of disorder is that associated with Down syndrome, the most common genetic cause of intellectual disability (Martin et al., 2009). Another example of a developmental disability in which communication is at the core of the impairment is autism spectrum disorder (ASD). In many cases, individuals with this type of language impairment have not fully achieved language proficiency, and, combined with other risks, the impairment adversely impacts them academically and vocationally, with long-term social and economic costs (Williams, 1970). The second type of language impairment is an acquired impairment of language abilities resulting from damage to portions of the brain responsible for language. For example, aphasia often occurs as a result of brain injury from stroke or trauma but may develop from tumors or progressive neurological disease (NIDCD, 2015). Individuals with this type of impairment have achieved efficient natural language proficiency, often for a number of years, prior to the onset of their loss of language skill.

Although considerable individual variability exists among all causes of language impairment, specific features of language are encompassed by receptive, expressive, and pragmatic skills. In the remainder of this section, three impairments—Down syndrome, ASD, and aphasia—are used to highlight aspects of language impairment and potential influences on employment.

A variety of instruments for assessing language skills among adults exist; some target specific etiologies (i.e., stroke, brain injury, ASD), while others focus on aspects of language skills (i.e., receptive, expressive, pragmatic; see Annex Table 5-9). The focus of many assessment protocols is on identifying the severity of discrete impairments or skills consistent with body structure and function and activity components of the *International Classification of Functioning, Disability and Health* (Raghavendra et al., 2007; Simeonsson et al., 2012; WHO, 2002). With any assessment, measured severity does "not necessarily relate to measures of life participation … it is not only those with mild aphasia who return to work" (Hinkley, 2002, p. 544); therefore, additional assessment is required to capture potential employment-related communication abilities. In the context of this report, language assessment provides information to assist in determining the ability to express or exchange "ideas by means of the spoken word to impart oral information to clients or the public and to convey detailed verbal instructions to other workers accurately, loudly, or quickly" (DOL, 2018, p. 125).

Receptive language refers to the ability to understand information. It involves understanding spoken words, sentences, and meaning expressed by others within employment (and other) settings. Expressive language refers to being able to put thoughts into words and sentences in a way that makes sense and is grammatically accurate. Expressive language involves speaking words and sentences in a meaningful way to express ideas rapidly and accurately within employment (and other) settings. Pragmatic language refers to the ability to integrate context-dependent aspects of language into interactions relevant during employment-related communication. Discourse—commonly described as any unit of language larger than a single sentence used for a communicative purpose—involves a combination of all three of these language skills and is a crucial component of functional language in an employment setting (Pritchard et al., 2018).

Language Assessment

Formulating a battery of standard language assessments for use in determining an individual's ability to work involves combining standardized assessment procedures, narrative discourse sampling and analyses, and self- (and communication partner) report measures to examine the impact of language impairments on daily interactions. Annex Table 5-9 summarizes commonly used language protocols in each of these categories (i.e., standard procedures, discourse sampling, patient report). Noteworthy in this table is the number of tests available and the variety of language aspects assessed by each, with the largest focus being on acquired language impairments (i.e., stroke, brain injury). Assessments commonly used to identify discrete language impairments (e.g., body structure and function, activity) include the Boston Diagnostic Aphasia Examination (Third Edition) with the Boston Naming Test (Second Edition) (Goodglass et al., 2000); the Western Aphasia Battery (Kertesz, 2007); the Peabody Picture Vocabulary Test (Fifth Edition) (Dunn, 2018); and the Expressive Vocabulary Test (Third Edition) (Williams, 2018). The protocols range from a broad assessment of language functioning to assessment of specific aspects of language (e.g., picture recall and naming, vocabulary comprehension and expression). Functional language assessment integrates information obtained from these discrete tests with discourse analyses, informant description, and patient self-reports.

Discourse analyses Everyday discourse, such as communication used in descriptions (e.g., describe a work project in detail), recounts (e.g., detail events from a previous meeting), and procedural discourse (e.g., relay instructions to a colleague), entails key interactions expected of many employees. Assessments of discourse require extensive resources. Individuals being assessed typically produce discourse based on an external cue (e.g., topic, image, interview). Their discourse productions are then audio recorded and subsequently manually transcribed. Finally, the transcription is processed using a variety of methods to identify language components (e.g., meaning, topic coherence, reference chains, verb structure and complexity) found to meet acceptability, reliability, and validity standards for clinical assessment (Pritchard et al., 2018). Although discourse analysis is of paramount importance (Wallace et al., 2014), these procedures are rarely performed as components of a clinical assessment because of the prohibitive time, expertise, and training they require.

Informant descriptions and patient self-reports Patient-reported outcomes (e.g., voice-related quality of life measures, communication effectiveness) are important indices of successful communication (Eadie, 2003) and complement other measurements of language production (Nagle and Eadie, 2012). With these measures, individuals with language impairments complete a structured self-evaluation and report the effects of their communication disability on daily interactions to illuminate functional impacts.

Recently, a number of self-report as well as informant-report measures have been validated for clinical use. Many evaluate communication effectiveness in day-to-day situations, communication quality of life, and overall communication-related function (see Annex Table 5-9). Three frequently used measures address functional communication:

- The ASHA Functional Assessment of Communication Skills for Adults, Revised Edition (Frattali et al., 1995, 2017) measures the functional communication of adults with speech, language, and cognitive-communication impairments. The items on this assessment encompass social communication, communication of basic needs, reading/writing/number concepts, and daily planning.
- The ASHA Quality of Communication Life Scale (QCL) (Paul et al., 2004) measures the extent to which a person's communication allows participation in life situations. The ASHA QCL captures information about the impact of a communication disorder on communication, interactions, participation (e.g., in social, leisure, work, and education activities), and overall quality of life. It is intended to provide information about the psychosocial, vocational, and educational effects of having a communication impairment. The ASHA QCL is a valid measure of communication-related quality of life for use with adults with neurogenic communication disorders (i.e., aphasia, cognitive communication disorders, and dysarthria).
- The Communication Participation Item Bank is a self-report instrument that can be used to evaluate the extent to which communication disorders impede communicative participation. It is increasingly validated for use with various etiologies (e.g., Parkinson's disease, spasmodic dysphonia, stroke).

Considerations for Functional Communication

The extent to which communication is functional in a work setting is dependent on (1) the accuracy and effectiveness of an individual's speech-language communication skills and (2) the communication environment and partner skills. The communication environment links to language disability type; for example, individuals with social communication impairments, such as those with ASD, may be successful with work activities that do not require extensive social interactions (e.g., computer programming, account management). Figure 5-1 illustrates communication functionality based on the effectiveness of the individual's skills and the impact of the work environment on situational communication. Each aspect varies, based on the needs of the specific communication act (i.e., face-to-face interactions, small group/conference [business meeting], large group [educator], telephone, intercom, public address system). As is illustrated, an individual with poor communication skills in an employment environment that is also poor for communication (e.g., noise, distance from speakers, rapid response requirement) may have dysfunctional communication. However, an individual with the same poor communication skills may benefit from work-related communication supports sufficiently to become a functional communicator for employment. And a person with good individual communication skills in an adverse work environment may yet require work-related environmental supports for functional communication. The highlighted central areas indicate crossover zones where "fair or good" communication and work environment conditions converge. Assessment will identify the extent to which functional communication is achieved when work environment or individualized communication supports are provided. Employment success will be maximized when the individual is communicating (with or without communication and work environment supports) in one of the three "Functional" quadrants. Therefore, personal communication and work environment adaptations focus on achieving functional communication in that employment setting.

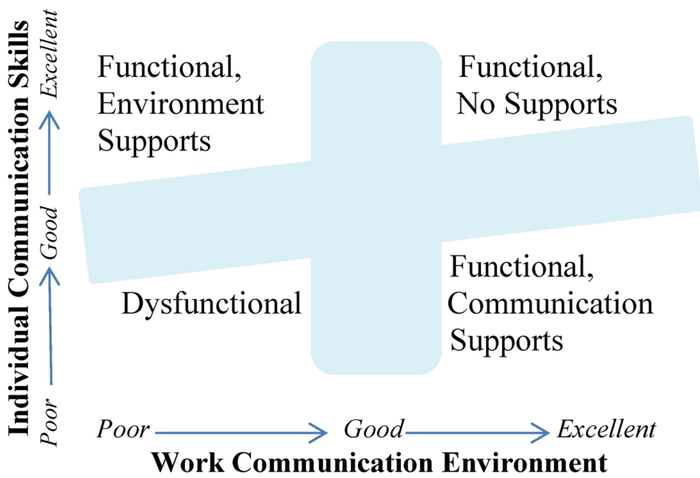

FIGURE 5-1 Relationship between communication skills and environment and impact on functionality.

Summary

In the context of functional abilities relevant to work requirements, an individual's comprehension of language produced by the communication partner, processing of that language, expression of his or her own ideas, and pragmatic interactions in dialogue appropriate to the work context are all relevant to assessing speech and language disorders for purposes of disability determination. Any aspect of speech and/or language impairment may result in impaired functional ability to communicate verbally in the context of work. Physical aspects of functional communication most relevant to adults communicating in a work setting include speech, voice, and fluency, while mental components of verbal communication include receptive, expressive, and pragmatic language.

FINDINGS AND CONCLUSIONS

Findings

5-1. Self-report and performance-based measures provide different perspectives on physical functional ability.

5-2. Numerous self-report measures of physical function have been developed, tested, and verified. Self-report measures are widely used, often require fewer clinical resources to administer, and provide the individual's perspective on his or her physical functional ability.

5-3. Numerous validated performance-based measures are available for assessing physical function in specific areas (e.g., range of motion, strength, balance). Such measures provide quantitative information about the areas assessed.

5-4. Self-report and performance-based measures of physical function may be limited by a number of factors, including an individual's underlying physical condition and cognitive status; the experience of pain, depression, or anxiety; and respondent bias or the person's level of effort.

5-5. Physical performance-based measures can add important information to that obtained with self-report questionnaires, and use of the two together can improve the prediction of work ability and allow comparison of results for consistency.

5-6. Functional capacity evaluations (FCEs) can provide information on an individual's ability to work.

5-7. There are multiple FCE instruments with varying degrees of reliability and validity. No single FCE instrument has proven superior for determining an individual's functional ability. The reliability and validity of FCEs can reflect a variety of confounders, including assessors' training; nonstandard testing environments; and examinees' effort, cooperation, and interest in returning to work. Assessors' estimate of the examinee's level of effort can enhance the accuracy of test results.

5-8. Metabolic exercise testing provides evidence of both maximal capacity and capacity for sustained work for at least 50 minutes.

5-9. Metabolic exercise testing provides information on functional capacity in units of work energy that are available for many physical activities of employment.

5-10. Performance-based measures have been developed and validated to measure visual, hearing, and communication ability and can be used to inform determinations of work ability.

5-11. Visual field status is one of the most important factors in predicting the mobility skills of individuals with vision loss.

5-12. The degree of hearing loss identified from a pure-tone audiogram can serve as a reasonable predictor of difficulty with workplace communication.

5-13. The extent to which communication is functional in a work setting is dependent on the accuracy and effectiveness of an individual's speech-language communication skills, as well as the communication environment and partner skills.

Conclusions

5-1. Given the complexity of measuring physical function and the multidimensional nature of work participation, no single instrument has yet been demonstrated to provide a comprehensive assessment of an individual's physical functional abilities relevant to work.

5-2. Self-report and performance-based measures of physical function can provide useful information for disability determination.

5-3. While individual performance-based measures may be of limited value in determining whole-body work ability, several tests can be of value when combined.

5-4. Self-report and performance-based measures of physical function provide complementary information, and together can be used to assess an individual's overall functional status, providing a more complete picture of whether or how well the individual will be able to perform everyday activities, including work, on a sustained basis than can be obtained with either type of measure alone.

5-5. Although FCE is not sufficient to predict successful performance of a particular occupation, FCE assessment can be useful to highlight strengths and weaknesses and may help focus intervention to improve ability to work in a specific setting.

5-6. When the ability to work is limited by cardiac function, metabolic exercise testing can help identify specific occupations that should be compatible with the residual cardiac functional capacity.

REFERENCES

Aasheim, T., and V. Finsen. 2014. The DASH and the QuickDASH instruments: Normative values in the general population in Norway. *Journal of Hand Surgery (European Volume)* 39(2):140–144.

Abel, S. M., S. Boyne, and H. Roesler-Mulroney. 2009. *Sound localization with an Army helmet worn in combination with an in-ear advanced communications system*. Ottawa, Ontario: Defense Research and Development Canada. http://www.dtic.mil/docs/citations/ADA525193 (accessed April 4, 2019).

AbilityLab (Shirley Ryan AbilityLab). 2013a. *Lower Extremity Functional Scale*. https://www.sralab.org/rehabilitation-measures/lower-extremity-functional-scale (accessed April 4, 2019).

AbilityLab. 2013b. *Numeric Pain Rating Scale*. https://www.sralab.org/rehabilitation-measures/numeric-pain-rating-scale (accessed April 8, 2019).

AbilityLab. 2015. *Foot and Ankle Ability Measures*. https://www.sralab.org/rehabilitation-measures/foot-and-ankle-ability-measures (accessed April 4, 2019).

AbilityLab. 2016. *Functional Gait Assessment*. https://www.sralab.org/rehabilitation-measures/functional-gait-assessment (accessed April 4, 2019).

Abrams, H. B., and J. Kihm. 2015. An introduction to MarkeTrak IX: A new baseline for the hearing aid market. *Hearing Review* 22(6):16.

Ackelman, B. H., and U. Lindgren. 2002. Validity and reliability of a modified version of the Neck Disability Index. *Journal of Rehabilitation Medicine* 34(6):284–287.

Albert, N., K. Trochelman, J. Li, and S. Lin. 2010. Signs and symptoms of heart failure: Are you asking the right questions? *American Journal of Critical Care* 19(5):443–452.

Amtmann, D., K. F. Cook, M. P. Jensen, W. H. Chen, S. Choi, D. Revicki, D. Cella, N. Rothrock, F. Keefe, L. Callahan, and J. S. Lai. 2010. Development of a PROMIS item bank to measure pain interference. *Pain* 150(1):173–182.

Anderson, L., C. R. Baylor, T. L. Eadie, and K. M. Yorkston. 2016. Describing speech usage in daily activities in typical adults. *Journal of Voice* 30(1):42–52. doi: 10.1016/j.jvoice.2015.02.001.

Andrews, G., A. Craig, A. Feyer, S. Hoddinott, P. Howie, and M. Neilson. 1983. Stuttering: A review of research findings and theories circa 1982. *Journal of Speech and Hearing Disorders* 48(3):226–246.

Ansuategui Echeita, J., B. J. van Holland, D. P. Gross, J. Kool, P. Oesch, M. A. Trippolini, and M. F. Reneman. 2018. Association between social factors and performance during functional capacity evaluations: A systematic review. *Disability and Rehabilitation*. Epub ahead of print. Pp. 1–11. doi: 10.1080/09638288.2018.1448120.

Arcara, G., and V. Bambini. 2016. A test for the assessment of pragmatic abilities and cognitive substrates (APACS): Normative data and psychometric properties. *Frontiers in Psychology* 7:70. doi: 10.3389/fpsyg.2016.00070.

Arffa, R. E., P. Krishna, J. Gartner-Schmidt, and C. A. Rosen. 2012. Normative values for the Voice Handicap Index-10. *Journal of Voice* 26(4):462–465.

Arizona State University. 2011. *Compendium of physical activities: Occupation*. https://sites.google.com/site/compendiumofphysicalactivities/Activity-Categories/occupation (accessed April 4, 2019).

ASHA (American Speech-Language-Hearing Association). 1993. *Definitions of communication disorders and variations*. https://www.asha.org/policy/rp1993-00208 (accessed April 4, 2019). doi: 10.1044/policy.RP1993-00208.

ASHA. 2004. *Preferred practice patterns for the profession of speech-language pathology*. Rockville, MD: ASHA. doi: 10.1044/policy.PP2004-00191.

ASHA. 2005. *Guidelines for manual pure-tone threshold audiometry*. https://www.asha.org/policy/gl2005-00014.htm (accessed April 4, 2019).

ASHA. 2016. *Scope of practice in speech-language pathology.* https://www.asha.org/policy/sp2016-00343 (accessed April 4, 2019).

ASHA. n.d. *Spoken language disorders*. https://www.asha.org/PRPSpecificTopic.aspx?folderid=8589935327&section=Overview (accessed April 4, 2019).

ASHA DIV3. n.d. *Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) ASHA Special Interest Division 3, voice and voice disorders*. https://www.asha.org/uploaded-Files/members/divs/D3CAPEVprocedures.pdf (accessed April 4, 2019).

ATS/ACCP (American Thoracic Society/American College of Chest Physicians). 2003. Statement on cardiopulmonary exercise testing. *American Journal of Respiratory and Critical Care Medicine* 167(2):211–238.

Awan, S. N., N. Roy, and C. Dromey. 2009. Estimating dysphonia severity in continuous speech: Application of a multi-parameter spectral/cepstral model. *Clinical Linguistics and Phonetics* 23(11):825–841. doi: 10.3109/02699200903242988.

Awan, S. N., N. Roy, M. E. Jetté, G. S. Meltzner, and R. E. Hillman. 2010. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics & Phonetics* 24(9):742–758.

Awan, S. N., N. Roy, D. Zhang, and S. M. Cohen. 2016. Validation of the Cepstral Spectral Index of Dysphonia (CSID) as a screening tool for voice disorders: Development of clinical cutoff scores. *Journal of Voice* 30(2):130–144.

Bailey, I. L., and J. E. Lovie. 1976. New design principles for visual acuity letter charts. *American Journal of Optometry and Physiological Optics* 53(11):740–745.

Ball, L. J. 2018. *Functional communication*. Paper commissioned by the Committee on Functional Assessment for Adults with Disabilities, National Academies of Sciences, Engineering, and Medicine, Washington, DC.

Ball, L. J., D. R. Beukelman, and G. L. Pattee. 2004. Communication effectiveness of individuals with amyotrophic lateral sclerosis. *Journal of Communication Disorders* 37(3):197–215. doi: 10.1016/j.jcomdis.2003.09.002.

Bayles, K. A., D. R. Boone, C. K. Tomoeda, T. J. Slauson, and A. W. Kaszniak. 1989. Differentiating Alzheimer's patients from the normal elderly and stroke patients with aphasia. *Journal of Speech and Hearing Disorders* 54(1):74–87.

Bayles, K. A., C. K. Tomoeda, and R. I. Dharmaperwira-Prins. 1993. *ABCD: Arizona Battery for Communication Disorders of Dementia*. Tuscon, AZ: Canyonlands Publishing.

Baylor, C. R., K. M. Yorkston, T. L. Eadie, R. Miller, and D. Amtmann. 2008. Levels of speech usage: A self-report scale for describing how people use speech. *Journal of Medical Speech-Language Pathology* 16(4):191–198. doi: 10.1109/TMI.2012.2196707.

Baylor, C., K. M. Yorkston, T. L. Eadie, J. Kim, H. Chung, and D. Amtmann. 2013. The Communicative Participation Item Bank (CPIB): Item bank calibration and development of a disorder-generic short form. *Journal of Speech, Language, and Hearing Research* 56(4):1190–1208.

Bean, J. F., D. D. Ölveczky, D. K. Kiely, S. I. LaRose, and A. M. Jette. 2011. Performance-based versus patient-reported physical function: What are the underlying predictors? *Physical Therapy* 91(12):1804–1811. doi: 10.2522/ptj.20100417.

Beaton, D. E., A. M. Davis, P. Hudak, and S. McConnell. 2001. The DASH (Disabilities of the Arm, Shoulder and Hand) outcome measure: What do we know about it now? *The British Journal of Hand Therapy* 6(4):109–118.

Beauchemin, G., J. C. MacDermid, K. Bourduas, M. F. Poirier, C. Gaudelli, and D. M. Rouleau. 2015. Translation and validation of the PREE (Patient Rated Elbow Evaluation) to a French version. *Orthopaedics & Traumatology: Surgery & Research* 101(4):405–409.

Bergström, G., J. Hagberg, H. Busch, I. Jensen, and C. Björklund. 2014. Prediction of sickness absenteeism, disability pension and sickness presenteeism among employees with back pain. *Journal of Occupational Rehabilitation* 24(2):278–286.

Bet, P. M., J. G. Hugtenburg, B. W. Penninx, and W. J. Hoogendijk. 2013. Side effects of antidepressants during long-term use in a naturalistic setting. *European Neuropsychopharmacology* 23(11):1443–1451.

Beukelman, D. R., J. Childes, T. Carrell, T. Funk, L. J. Ball, and G. L. Pattee. 2011. Perceived attention allocation of listeners who transcribe the speech of speakers with amyotrophic lateral sclerosis. *Speech Communication* 53(6):801–806. doi: 10.1016/j.specom.2010.12.005.

Bieniek, S., and M. Bethge. 2014. The reliability of WorkWell Systems Functional Capacity Evaluation: A systematic review. *BMC Musculoskeletal Disorders* 15:106. doi: 10.1186/1471-2474-15-106.

Bijur, P. E., W. Silver, and E. J. Gallagher. 2001. Reliability of the visual analog scale for measurement of acute pain. *Academic Emergency Medicine* 8(12):1153–1157.

Bilberg, A., T. Bremell, and K. Mannerkorpi. 2012. Disability of the Arm, Shoulder and Hand questionnaire in Swedish patients with rheumatoid arthritis: A validity study. *Journal of Rehabilitation Medicine* 44(1):7–11.

Bilek, L., D. Venema, G. Willett, and E. Lyden. 2008. Use of the Human Activity Profile for estimating fitness in persons with arthritis. *Arthritis and Rheumatism* 59(5):659–664.

Binkley, J. M., P. W. Stratford, S. A. Lott, and D. L. Riddle. 1999. The Lower Extremity Functional Scale (LEFS): Scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. *Physical Therapy* 79(4):371–383.

Blackstone, S., M. Williams, and D. Wilkins. 2007. Key principles underlying research and practice in AAC. *Augmentative and Alternative Communication* 23(3):191–203. doi: 10.1080/07434610701553684.

Bloodstein, O., and N. Bernstein Ratner. 2008. *A handbook of stuttering*, 6th ed. New York: Thomson Delmar Learning.

Boonstra, A. M., H. R. Schiphorst Preuper, M. F. Reneman, J. B. Posthumus, and R. E. Stewart. 2008. Reliability and validity of the visual analogue scale for disability in patients with chronic musculoskeletal pain. *International Journal of Rehabilitation Research* 31(2):165–169. doi: 10.1097/MRR.0b013e3282fc0f93.

Brutten, G., and M. Vanryckeghem. 2003. *Behavior Assessment Battery: A multi-dimensional and evidence-based approach to diagnostic and therapeutic decision making for children who stutter*. Belgium: Stichting Integratie Gehandicapten and Acco Publishers.

Brutten, G. J., and M. Vanryckeghem. 2007. *Behavior Assessment Battery for school-age children who stutter.* San Diego, CA: Plural Publishing.

Calmels, P., F. Béthoux, A. Condemine, and I. Fayolle-Minon. 2005. Low back pain disability assessment tools. *Annales de Réadaptation et de Médecine Physique* 48(6):288–297.

Carding, P. N., J. A. Wilson, K. MacKenzie, and I. J. Deary. 2009. Measuring voice outcomes: State of the science review. *The Journal of Laryngology & Otology* 123(8):823–829.

Chan, P. S., P. G. Jones, S. A. Arnold, and J. A. Spertus. 2014. Development and validation of a short version of the Seattle Angina Questionnaire. *Circulation: Cardiovascular Quality Outcomes* 7(5):640–647.

Chansirinukor, W., C. G. Maher, J. Latimer, and J. Hush. 2005. Comparison of the Functional Rating Index and the 18-item Roland-Morris Disability Questionnaire: Responsiveness and reliability. *Spine* 30(1):141–145.

Chapman, J. R., D. C. Norvel, J. T. Hermsmeyer, R. J. Bransford, J. DeVine, M. J. McGirt, and M. J. Lee. 2011. Evaluating common outcomes for measuring treatment success for chronic low back pain. *Spine* 36(21 Suppl.):S54–S68.

Chen, A. H., F. N. N. Norazman, and N. H. Buari. 2012. Comparison of visual acuity estimates using three different letter charts under two ambient room illuminations. *Indian Journal of Ophthalmology* 60(2):101–104.

Chen, J. J. 2007. Functional capacity evaluation & disability. *Iowa Orthopaedic Journal* 27:121–127.

Chien, W., and F. R. Lin. 2012. Prevalence of hearing aid use among older adults in the United States. *Archives of Internal Medicine* 172(3):292–293.

Chung, K. C., M. S. Pillsbury, M. R. Walters, and R. A. Hayward. 1998. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *Journal of Hand Surgery* 23(4):575–587.

Cleland, J. A., J. D. Childs, and J. M. Whitman. 2008. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Archives of Physical Medicine and Rehabilitation* 89(1):69–74.

Constantino, C. D., P. Leslie, R. W. Quesal, and J. S. Yaruss. 2016. A preliminary investigation of daily variability of stuttering in adults. *Journal of Communication Disorders* 60(2016):39–50. doi: 10.1016/j.jcomdis.2016.02.001.

Craig, A., and P. Calver. 1991. Following up on treated stutterers: Studies of perceptions of fluency and job status. *Journal of Speech and Hearing Research* 34(2):279–284.

Craig, A., K. Hancock, E. Chang, C. McCready, A. Shepley, and A. McCaul. 1996. A controlled clinical trial for stuttering in persons aged 9 to 14 years. *Journal of Speech and Hearing Research* 39(4):808–826.

Craig, A., E. Blumgart, and Y. Tran. 2009. The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders* 34(2):61–71. doi: 10.1016/j.jfludis.2009.05.002.

Curb, J. D., C. D. Ceria-Ulep, B. L. Rodriguez, J. Grove, J. Guralnik, B. J. Willcox, T. A. Donlon, K. H. Masaki, and R. Chen. 2006. Performance-based measures of physical function for high-function populations. *Journal of the American Geriatrics Society* 54(5):737–742.

Dabul, B. 2000. *Apraxia battery for adults: Examiner's manual*. Austin, TX: PRO-ED.

Davidson, M., and N. de Morton. 2007. A systematic review of the Human Activity Profile. *Clinical Rehabilitation* 21(2):151–162.

Davidson, M., and J. L. Keating. 2002. A comparison of five low back disability questionnaires: Reliability and responsiveness. *Physical Therapy* 82(1):8–24.

Davies, C. C., and A. J. Nitz. 2013. Psychometric properties of the Roland-Morris Disability Questionnaire compared to the Oswestry Disability Index: A systematic review. *Physical Therapy Reviews* 14(6):399–408. doi: 10.1179/108331909X12540993898134.

De Baets, S., P. Calders, N. Schalley, K. Vermeulen, S. Vertriest, L. Van Peteghem, M. Coussens, F. Malfait, G. Vanderstraeten, G. Van Hove, and D. Van de Velde. 2017. Updating the evidence on functional capacity evaluation methods: A systematic review. *Journal of Occupational Rehabilitation* 28(3):418–428. doi: 10.1007/s10926-017-9734-x.

Demeester, K., V. Topsakal, J. J. Hendrickx, E. Fransen, L. Van Laer, G. Van Camp, P. Van de Heyning, and A. Van Wieringen. 2012. Hearing disability measured by the Speech, Spatial, and Qualities of Hearing Scale in clinically normal-hearing and hearing-impaired middle-aged persons, and disability screening by means of a reduced SSQ (the SSQ5). *Ear and Hearing* 33(5):615–616. doi: 10.1097/AUD.0b013e31824e0ba7.

Dingemans, S. A., S. C. Kleipool, M. A. M. Mulders, J. Winkelhagen, N. W. L. Schep, J. C. Goslings, and T. Schepers. 2017. Normative data for the Lower Extremity Functional Scale (LEFS). *Acta Orthopaedica* 88(4):422–426. doi: 10.1080/17453674.2017.1309886.

Dixon, D., M. Johnston, and M. McQueen. 2008. The Disabilities of the Arm, Shoulder and Hand Questionnaire (DASH) can measure the impairment, activity limitations and participation restriction constructs from the *International Classification of Functioning, Disability and Health* (ICF). *BMC Musculoskeletal Disorders* 9:114. doi: 10.1186/1471-2474-9-114.

DOL (U.S. Department of Labor). 2018. *Occupational Requirements Survey (ORS): Collection manual*, 4th ed. https://www.bls.gov/ors/about/information-for-survey-participants/pdf/occupational-requirements-survey-collection-manual-082018.pdf (accessed April 4, 2019).

DOL. n.d. *Abilities*. https://www.onetonline.org/find/descriptor/browse/Abilities (accessed April 4, 2019).

Donovan, N. J., D. L. Kendall, M. E. Young, and J. C. Rosenbek. 2008. The Communicative Effectiveness Survey: Preliminary evidence of construct validity. *American Journal of Speech-Language Pathology* 17(4):335–347. doi: 10.1044/1058-0360(2008/07-0010).

Douglas, J. M., C. A. O'Flaherty, and P. C. Snow. 2000. Measuring perception of communicative ability: The development and evaluation of the La Trobe communication questionnaire. *Aphasiology* 14(3):251–268.

Drummond, S. S. 1993. *Dysarthria Examination Battery*. Tuscon, AZ: Communication Skill Builders.

Dubno, J. R., D. D. Dirks, and D. E. Morgan. 1984. Effects of age and mild hearing loss on speech recognition in noise. *The Journal of the Acoustical Society of America* 76(1):87–96.

Duffy, J. R. 2013. *Motor speech disorders: Substrates, differential diagnosis, and management*, 3rd ed. St. Louis, MO: Elsevier Mosby.

Dunn, D. 2018. *Peabody Picture Vocabulary Test*. San Antonio, TX: Pearson. https://www.pearsonclinical.com/language/products/100001984/peabody-picture-vocabulary-test-fifth-edition.html (accessed April 4, 2019).

Eadie, T. L. 2003. The ICF: A proposed framework for comprehensive rehabilitation of individuals who use alaryngeal speech. *American Journal of Speech-Language Pathology* 12(2):189–197.

Ellestad, M. H., W. Allen, M. C. K. Wan, and G. L. Kemp. 1969. Maximal treadmill stress testing for cardiovascular evaluation. *Circulation* 39(4):517–522.

Elman, R. J., and E. Bernstein-Ellis. 1995. What is functional? *American Journal of Speech-Language Pathology* 4(4):115–117.

Enderby, P. M. 1983. *Frenchay Dysarthria Assessment*. Austin, TX: PRO-ED.

Enderby, P. M., and R. Palmer. 2008. *FDA-2: Frenchay Dysarthria Assessment: Examiner's manual*. Austin, TX: PRO-ED.

Erb, B. D. 1970. *Physician's handbook for evaluation of cardiovascular and physical fitness* (Table F-4 Estimated VO2 for various activities). Nashville, TN: Tennessee Heart Association Physical Exercise Committee.

Erber, N. P. 1982. *Auditory training*. Washington, DC: Alex Graham Bell Association for Deaf and Hard of Hearing.

Fairbank, J. C., and P. B. Pynsent. 2000. The Oswestry Disability Index. *Spine* 25(22):2940–2952.

Fairbank, J. C., J. Couper, J. B. Davies, and J. P. O'Brien. 1980. The Oswestry Low Back Pain Disability Questionnaire. *Physiotherapy* 66(8):271–273.

Félin-Germain, A., I. Denis, S. Turcotte, R. Fleet, P. Archambault, C. E. Dionne, and G. Foldes-Busque. 2018. Work absenteeism and presenteeism loss in patients with non-cardiac chest pain. *Journal of Occupational and Environmental Medicine* 60(9):781–786.

Fore, L., Y. Perez, R. Neblett, S. Asih, T. G. Mayer, and R. J. Gatchel. 2015. Improved functional capacity evaluation performance predicts successful return to work one year after completing a functional restoration rehabilitation program. *PM&R: The Journal of Injury, Function, and Rehabilitation* 7(4):365–375. doi: 10.1016/j.pmrj.2014.09.013.

Francis, D., J. Daniero, K. Hovis, N. Sathe, B. Jacobson, D. Penson, I. D. Feurer, and M. L. McPheeters. 2017. Voice-related patient-reported outcome measures: A systematic review of instrument development and validation. *Journal of Speech, Language, and Hearing Research* 60(1):62–88. doi: 10.1044/2016_JSLHR-S-16-0022.

Frattali, C., C. Thompson, A. Holland, M. Ferketic, and C. Wohl. 1995. *Functional Assessment of Communication Skills for Adults (ASHA FACS).* Rockville, MD: ASHA.

Frattali, C. M., C. K. Thompson, A. L. Holland, C. B. Wohl, C. J. Wenck, S. C. Slater, and D. Paul. 2017. *American Speech-Language-Hearing Association Functional Assessment of Communication Skills for adults (ASHA FACS).* Rockville, MD: ASHA.

Gatehouse, S., and W. Noble. 2004. The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology* 43(2):85–99.

Genovese, E., and J. S. Galper (Eds.). 2009. *Guide to the evaluation of functional ability: How to request, interpret, and apply functional capacity evaluations*. Chicago, IL: American Medical Association.

Giguère, C., C. Laroche, S. D. Soli, and V. Vaillancourt. 2008. Functionally-based screening criteria for hearing-critical jobs based on the Hearing in Noise Test. *International Journal of Audiology* 47(6):319–328.

Gliklich, R. E., R. M. Glovsky, and W. W. Montgomery. 1999. Validation of a voice outcome survey for unilateral vocal cord paralysis. *Otolaryngology—Head and Neck Surgery* 120(2):153–158.

Goldman, L., B. Hashimoto, E. F. Cook, and A. Loscalzo. 1981. Comparative reproducibility and validity of systems for assessing cardiovascular functional class, advantages of a new specific activity scale. *Circulation* 64(6):1227–1234.

Goldman, N., D. A. Glei, L. Rosero-Bixby, S.-T. Chiou, and M. Weinstein. 2014. Performance-based measures of physical function as mortality predictors: Incremental value beyond self-reports. *Demographic Research* 30(7):227–252.

Goodglass, H., E. Kaplan, and B. Barresi. 2000. *Boston Diagnostic Aphasia Examination*, 3rd ed. Austin, TX: PRO-ED.

Gorga, M. P., S. T. Neely, B. Ohlrich, B. Hoover, J. Redner, and J. O. Peters. 1997. From laboratory to clinic: A large scale study of distortion product otoacoustic emissions in ears with normal hearing and ears with hearing loss. *Ear and Hearing* 18(6):440–455.

Gouttebarge, V., H. Wind, P. P. Kuijer, and M. H. Frings-Dresen. 2004. Reliability and validity of functional capacity evaluation methods: A systematic review with reference to Blankenship system, Ergos work simulator, Ergo-Kit and Isernhagen work system. *International Archives of Occupational and Environmental Health* 77(8):527–537.

Grant, K. W., B. E. Walden, and P. F. Seitz. 1998. Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *The Journal of the Acoustical Society of America* 103(5):2677–2690.

Gray, C., C. Baylor, T. Eadie, D. Kendall, and K. Yorkston. 2012. The levels of speech usage rating scale: Comparison of client self-ratings with speech pathologist ratings. *International Journal of Language and Communication Disorders* 47(3):333–344. doi: 10.1111/j.1460-6984.2011.00112.x.

Greenberg, B., and A. M. Kahn. 2012. Clinical assessment of heart failure. In *Braunwald's Heart Disease: A textbook of cardiovascular medicine,* 9th ed., edited by R. O. Bonow, D. L. Mann, D. P. Zipes, and P. Libby. Philadelphia, PA: Elsevier Saunders. Pp. 505–516.

Grobbel, J., J. Dietzsch, C. A. Johnson, R. Vonthein, K. Stingl, R. G. Weleber, and U. Schiefer. 2016. Normal values for the full visual field, corrected for age-and reaction time, using semiautomated kinetic testing on the Octopus 900 perimeter. *Translational Vision Science & Technology* 5(2):5.

Gross, D. P. 2006. Are functional capacity evaluations affected by the patient's pain? *Current Pain and Headache Reports* 10(2):107–113.

Gross, D. P., and M. C. Battié. 2004. The prognostic value of functional capacity evaluation in patients with chronic low back pain: Part 2: Sustained recovery. *Spine* 29(8):920–924.

Gross, D. P., and M. C. Battié. 2005. Functional capacity evaluation performance does not predict sustained return to work in claimants with chronic back pain. *Journal of Occupational Rehabilitation* 15(3):285–294.

Gross, D. P., M. C. Battié, and J. D. Cassidy. 2004. The prognostic value of functional capacity evaluation in patients with chronic low back pain: Part 1: Timely return to work. *Spine* 29(8):914–919.

Grotle, M., J. I. Brox, and N. K. Vøllestad. 2005. Functional status and disability questionnaires: What do they assess? A systematic review of back-specific outcome questionnaires. *Spine* 30(1):130–140.

Häusler, R., S. Colburn, and E. Marr. 1983. Sound localization in subjects with impaired hearing: Spatial-discrimination and interaural-discrimination tests. *Acta Oto-Laryngologica* 96(Suppl. 400):1–62.

Hawker, G. A., S. Mian, T. Kendzerska, and M. French. 2011. Measures of adult pain. *Arthritis Care & Research* 63(Suppl. 11):S240–S252.

HealthMeasures. 2017. *Pain interference: A brief guide to the PROMIS pain interference instruments*. http://www.healthmeasures.net/images/PROMIS/manuals/PROMIS_Pain_Interference_Scoring_Manual.pdf (accessed April 8, 2019).

Helm-Estabrooks, N. 2017. *Cognitive-Linguistic Quick Test-Plus*. https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Cognition-&-Neuro/Non-Verbal-Ability/Cognitive-Linguistic-Quick-Test-Plus/p/100000459.html (accessed April 12, 2019).

HHS (U.S. Department of Health and Human Services). 2018. *Patient-Reported Outcomes Measurement Information System: Program snapshot*. https://commonfund.nih.gov/promis/index (accessed April 4, 2019).

Hicks, G. E., and T. J. Manal. 2009. Psychometric properties of commonly used low back disability questionnaires: Are they useful for older adults with low back pain? *Pain Medicine* 10(1):85–94, doi: 10.1111/j.1526-4637.2008.00548.x.

Hilari, K., S. Byng, D. L. Lamping, and S. C. Smith. 2003. Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of acceptability, reliability, and validity. *Stroke* 34(8):1944–1950.

Hinckley, J. J. 2002. Vocational and social outcomes of adults with chronic aphasia. *Journal of Communication Disorders* 35(6):543–560.

Hirano, M. 1981. *Clinical Examination of Voice*. New York: Springer-Verlag.

Hlatky, M. A., R. E. Boineau, M. B. Higginbotham, K. L. Lee, D. B. Mark, R. M. Califf, F. R. Cobb, and D. B. Pryor. 1989. A brief self-administered questionnaire to determine functional capacity (the Duke Activity Status Index). *American Journal of Cardiology* 64(10):651–654.

Hogan, A., K. O'Loughlin, A. Davis, and H. Kendig. 2009. Hearing loss and paid employment: Australian population survey findings. *International Journal of Audiology* 48(3):117–122.

Hogikyan, N. D., and G. Sethuraman. 1999. Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice* 13(4):557–569.

Holland, A. L., C. Frattali, and D. Fromm. 1999. *Communication Activities of Daily Living: CADL-2*. Austin, TX: PRO-ED.

Hotz, M. A., R. Probst, F. P. Harris, and R. Hauser. 1993. Monitoring the effects of noise exposure using transiently evoked otoacoustic emissions. *Acta Oto-Laryngologica* 113(4):478–482.

Hula, W. D., P. J. Doyle, C. A. Stone, S. N. Austermann Hula, S. Kellough, J. L. Wambaugh, K. B. Ross, J. G. Schumacher, and A. St. Jacque. 2015. The Aphasia Communication Outcome Measure (ACOM): Dimensionality, item bank calibration, and initial validation. *Journal of Speech, Language, and Hearing Research* 58(3):906–919.

Humes, L. E. 2002. Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *The Journal of the Acoustical Society of America* 112(3):1112–1132.

Hustad, K. C. 1999. Optimizing communication effectiveness: Bringing it together. In *Management of motor speech disorders in children and adults*, edited by K. M. Yorkston, D. R. Beukelman, E. A. Strand, and M. Hakel. Austin, TX: PRO-ED.

Hustad, K. C. 2006. Estimating the intelligibility of speakers with dysarthria. *Folia Phoniatrica et Logopaedica* 58(3):217–228.

Hustad, K. C. 2008. The relationship between listener comprehension and intelligibility scores for speakers with dysarthria. *Journal of Speech, Language, and Hearing Research* 51(3):562–573. doi: 10.1044/1092-4388(2008/040).

Hustad, K. C., and D. Beukelman. 2002. Listener comprehension of severely dysarthric speech: Effects of linguistic cues and stimulus cohesion. *Journal of Speech, Language, and Hearing Research* 45(3):545–558. doi: 10.1044/1092-4388(2002/043).

IOM (Institute of Medicine). 2010. *Cardiovascular disability: Updating the Social Security listings*. Washington, DC: The National Academies Press.

Ishikawa, K., S. Boyce, L. Kelchner, M. G. Powell, H. Schieve, A. de Alarcon, and S. Khosla. 2017. The effect of background noise on intelligibility of dysphonic speech. *Journal of Speech, Language, and Hearing Research* 60(7):1919–1929. doi: 10.1044/2017_JSLHR-S-16-0012.

Iverson-Thoburn, S., and D. Hayden. 2000. Alaryngeal speech utilization: A survey. *Journal of Medical Speech–Language Pathology* 8(2):85–99.

IWH (Institute for Work & Health). n.d. *The DASH outcome measure: About the DASH*. http://www.dash.iwh.on.ca/about-dash (accessed April 4, 2019).

Jahn, W. T., L. N. Cupon, and J. H. Steinbaugh. 2004. Functional and work capacity evaluation issues. *Journal of Chiropractic Medicine* 3(1):1–5. doi: 10.1016/S0899-3467(07)60059-7.

Jennings, M. B., and L. Shaw. 2008. Impact of hearing loss in the workplace: Raising questions about partnerships with professionals. *Work* 30(3):289–295.

Jester, A., A. Harth, G. Wind, G. Germann, and M. Sauerbier. 2005. Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire: Determining functional activity profiles in patients with upper extremity disorders. *Journal of Hand Surgery* 30(1):23–28.

John, M., F. Angst, G. Pap, A. Junge, and A. F. Mannion. 2007. Cross-cultural adaptation, reliability and validity of the Patient Rated Elbow Evaluation (PREE) for German-speaking patients. *Clinical and Experimental Rheumatology* 25(2):195–205.

Johnson, C. A., and J. L. Keltner. 1983. The incidence of visual field loss in 20,000 eyes and its relationship to driving performance. *Archives of Ophthalmology* 101(3):371–376.

Joseph, S. M., E. Novak, S. Z. Arnold, P. G. Jones, H. Khattak, A. E. Platts, V. Davila-Roman, D. L. Mann, and J. A. Spertus. 2013. Comparable performance of the Kansas City Cardiomyopathy Questionnaire in heart failure patients with preserved and reduced ejection fraction. *Circulation: Heart Failure* 6(6):1139–1146. doi: 10.1161/CIRCHEARTFAILURE.113.000359.

Kaiser, P. K. 2009. Prospective evaluation of visual acuity assessment: A comparison of Snellen versus ETDRS charts in clinical practice (an AOS thesis). *Transactions of the American Ophthalmological Society* 107:311–324.

Kawai, K., A. T. Kawai, P. Wollan, and B. P. Yawn. 2017. Adverse impacts of chronic pain on health-related quality of life, work productivity, depression and anxiety in a community-based study. *Family Practice* 34(6):656–661.

Kay, J. M., R. Lesser, and M. Coltheart. 1992. *PALPA: Psycholinguistic assessment of language performance in aphasia*. Hove, England: Lawrence Erlbaum Associates.

Kelkar, A. A., J. Spertus, P. Peter, R. F. Pierson, R. J. Cody, I. L. Pina, A. Hernandez, and J. Butler. 2016. Utility of patient-reported outcome instruments in heart failure. *Journal of the American College of Cardiology* 4(3):116–125.

Keltner, J. L., and C. A. Johnson. 1980. Mass visual field screening in a driving population. *Ophthalmology* 87(8):785–792.

Kendall, S. E., P. Sjøgren, C. A. de Mattos Pimenta, J. Højsted, and G. P. Kurita. 2010. The cognitive effects of opioids in chronic non-cancer pain. *Pain* 150(2):225–230.

Kennedy, C. A., D. E. Beaton, P. Smith, D. Van Eerd, K. Tang, T. Inrig, S. Hogg-Johnson, D. Linton, and R. Couban. 2013. Measurement properties of the QuickDASH (Disabilities of the Arm, Shoulder and Hand) outcome measure and cross-cultural adaptations of the QuickDASH: A systematic review. *Quality of Life Research* 22(9):2509–2547. doi: 10.1007/s11136-013-0362-4.

Kent, R. D. 1998. Normal aspects of articulation. In *Articulation and phonological disorders*, 4th ed., edited by J. E. Bernthal and N. Bankson. Needham Heights, MA: Allyn and Bacon. Pp. 1–62.

Kent, R. D., G. Weismer, J. F. Kent, and J. C. Rosenbek. 1989. Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders* 54(4):482–499.

Kertesz, A. 1982. *Western Aphasia Battery test manual*. San Antonio, TX: The Psychological Corporation.

Kertesz, A. 2007. *Western Aphasia Battery-revised*. San Antonio, TX: Pearson.

Klein, J., and S. Hood. 2004. The impact of stuttering on employment opportunities and job performance. *Journal of Fluency Disorders* 29(4):255–273.

Ko, S., and S. Chae. 2017. Correlations between the SF-36, the Oswestry-Disability Index and Rolland-Morris Disability Questionnaire in patients undergoing lumbar decompression according to types of spine origin pain. *Clinical Spine Surgery* 30(6):E804–E808.

Kochkin, S. 2009. MarkeTrak VIII: 25-year trends in the hearing health market. *Hearing Review* 16(11):12–31.

Kopec, J. A., J. M. Esdaile, M. Abrahamowicz, L. Abenhaim, S. Wood-Dauphinee, D. L. Lamping, and J. I. Williams. 1995. The Quebec Back Pain Disability Scale: Measurement properties. *Spine* 20(3):341–352.

Kramer, S. E., T. S. Kapteyn, and T. Houtgast. 2006. Occupational performance: Comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work: Desempeño laboral: Comparación de empleados con audición normal o alterada usando el Listado Amsterdam para Audición y Trabajo. *International Journal of Audiology* 45(9):503–512.

Kresal, F., V. Roblek, A. Jerman, and M. Meško. 2015. Lower back pain and absenteeism among professional public transport drivers. *International Journal of Occupational Safety and Ergonomics* 21(2):166–172.

Krupp, L. B., N. G. LaRocca, J. Muir-Nash, and A. D. Steinberg. 1989. The fatigue severity scale: Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology* 46(10):1121–1123.

Kuijer, P. P. F. M., V. Gouttebarge, S. Brouwer, M. F. Reneman, and M. H. W. Frings-Dresen. 2012. Are performance-based measures predictive of work participation in patients with musculoskeletal disorders? A systematic review. *International Archives of Occupational and Environmental Health* 85(2):109–123.

Laplante-Lévesque, A., K. J. Brännström, E. Ingo, G. Andersson, and T. Lunner. 2015. Stages of change in adults who have failed an online hearing screening. *Ear and Hearing* 36(1):92–101.

Laroche, C., S. Soli, C. Giguère, J. Lagacé, V. Vaillancourt, and M. Fortin. 2003. An approach to the development of hearing standards for hearing-critical jobs. *Noise and Health* 6(21):17–37.

Leddy, A. L., B. E. Crowner, and G. M. Earhart. 2011. Functional Gait Assessment and Balance Evaluation System Test: Reliability, validity, sensitivity, and specificity for identifying individuals with Parkinson Disease who fall. *Physical Therapy* 91(1):102–113.

Lichtenstein, M. J., F. H. Bess, and S. A. Logan. 1988. Validation of screening tools for identifying hearing-impaired elderly in primary care. *JAMA* 259(19):2875–2878.

Light, J. 1988. Interaction involving individuals using augmentative and alternative communication systems: State of the art and future directions. *Augmentative and Alternative Communication* 4(2):66–82. doi: 10.1080/07434618812331274657.

Lin, F. R., J. K. Niparko, and L. Ferrucci. 2011. Hearing loss prevalence in the United States. *Archives of Internal Medicine* 171(20):1851–1853.

Llewellyn-Thomas, H., H. J. Sutherland, R. Tibshirani, A. Ciampi, J. E. Till, and N. F. Boyd. 1984. Describing health states: Methodologic issues in obtaining values for health states. *Medical Care* 22(6):543–552.

Lomas, J., L. Pickard, S. Bester, H. Elbard, A. Finlayson, and C. Zoghaib. 1989. The Communicative Effectiveness Index: Development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders* 54(1):113–124. doi: 10.1044/jshd.5401.113.

Lucas, C., L. W. Stevenson, J. Johnson, H. Hartley, M. A. Hamilton, J. Walder, V. Lem, and E. Eagen-Bengsten. 1999. The 6-min walk and peak oxygen consumption in advanced heart failure. *American Heart Journal* 138(4 Pt. 1):618–624.

MacDermid, J. C. 2001. Outcome evaluation in patients with elbow pathology: Issues in instrument development and evaluation. *Journal of Hand Therapy* 14(2):105–114.

MacDermid, J. C. 2010. *The Patient-Rated Elbow Evaluation (PREE) user manual.* https://srs-mcmaster.ca/wp-content/uploads/2015/05/English-PREE-User-Manual_June-2010.pdf (accessed April 4, 2019).

MacDermid, J. C., T. Turgeon, R. S. Richards, M. Beadle, and J. H. Roth. 1998. Patient rating of wrist pain and disability: A reliable and valid measurement tool. *Journal of Orthopaedic Trauma* 12(8):577–586.

MacPhee, G. J., J. A. Crowther, and C. H. McAlpine. 1988. A simple screening test for hearing impairment in elderly patients. *Age and Ageing* 17(5):347–351.

Marron, J. A., and I. L. Bailey. 1982. Visual factors and orientation and mobility performance. *American Journal of Optometry and Physiological Optics* 59(5):413–426.

Marslen-Wilson, W. 1989. Access and integration: Projecting sound onto meaning. In *Lexical representation and processing*, edited by W. Marslen-Wilson. Cambridge, MA: MIT Press. Pp. 3–24.

Martin, G. E., J. Klusek, B. Estigarribia, and J. E. Roberts. 2009. Language characteristics of individuals with Down syndrome. *Topics in Language Disorders* 29(2):112–132. doi: 10.1097/TLD.0b013e3181a71fe1.

Martin, R. T., and J. J. Irrgang. 2007. A survey of self-reported outcome instruments for the foot and ankle. *The Journal of Orthopaedic and Sports Physical Therapy* 37(2):72–84.

Martin, R. T., J. J. Irrgang, R. G. Burdett, S. F. Conti, and J. M. Van Swearingen. 2005. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). *Foot & Ankle International* 26(11):968–983. doi: 10.1177/107110070502601113.

Mathiowetz, V., S. Rogers, M. Dowe-Keval, and C. Rennels. 1986. The Purdue Pegboard Norms for 14-19 year olds. *American Journal of Occupational Therapy* 40(3):174–179.

McAuliffe, M. J., S. Carpenter, and C. Moran. 2010. Speech intelligibility and perceptions of communication effectiveness by speakers with dysarthria following traumatic brain injury and their communication partners. *Brain Injury* 24(12):1408–1415. doi: 10.3109/02699052.2010.511590.

McBride, W. S., C. D. Mulrow, C. Aguilar, and M. R. Tuley. 1994. Methods for screening for hearing loss in older adults. *American Journal of the Medical Sciences* 307(1):40–42.

McCarthy, M. J. H., M. P. Grevitt, P. Silcocks, and G. Hobbs. 2007. The reliability of the Vernon and Mior Neck Disability Index, and its validity compared with the Short Form-36 health survey questionnaire. *European Spine Journal* 16(12):2111–2117.

McCreery, R. 2018. *Functional assessment of hearing*. Paper commissioned by the Committee on Functional Assessment for Adults with Disabilities, National Academies of Sciences, Engineering, and Medicine, Washington, DC.

McDonald, M., M. da Costa DiBonaventura, and S. Ullman. 2011. Musculoskeletal pain in the workforce: The effects of back, arthritis, and fibromyalgia pain on quality of life and work productivity. *Journal of Occupational and Environmental Medicine* 53(7):765–770.

McGregor, A. 2003. Fitness standards in airline staff. *Occupational Medicine* 53(1):5–9.

Mehta, S. P., A. Fulton, C. Quach, M. Thistle, C. Toledo, and N. A. Evans. 2016. Measurement properties of the Lower Extremity Functional Scale: A systematic review. *Journal of Orthopaedic & Sports Physical Therapy* 46(3):200–216.

Meterko, M., E. E. Marfeo, C. M. McDonough, A. M. Jette, P. Ni, K. Bogusz, E. K. Rasch, D. E. Brandt, and L. Chan. 2015. The Work Disability Functional Assessment Battery (WD-FAB): Feasibility and psychometric properties. *Archives of Physical Medicine and Rehabilitation* 96(6):1028–1035. doi: 10.1016/j.apmr.2014.11.025.

Meterko, M., M. Marino, P. Ni, E. Marfeo, C. M. McDonough, A. Jette, K. Peterik, E. Rasch, D. E. Brandt, and L. Chan. 2018. Psychometric evaluation of the improved Work-Disability Functional Assessment Battery. *Archives of Physical Medicine and Rehabilitation*. In press. Corrected proof available online December 19, 2018. doi: 10.1016/j.apmr.2018.09.125.

Metra, M., R. Raddino, L. D. Cas, and O. Visioli. 1990. Assessment of peak oxygen consumption, lactate, and ventilator thresholds and correlation with resting and exercise hemodynamic data in chronic congestive heart failure. *American Journal of Cardiology* 65(16):1127–1133.

Montgomery, S., S. D. Soli, L. S. Meyers, L. Alvaraz, K. A. Hunley, S. L. Fuller, and P. Duenas. 2011. *Hearing standard for selection of entry-level correctional officers*. Sacramento, CA: State of California Corrections Standards Authority, California Department of Corrections and Rehabilitation.

Mumby, K., A. Bowen, and A. Hesketh. 2007. Apraxia of speech: How reliable are speech and language therapists' diagnoses? *Clinical Rehabilitation* 21(8):760–767. doi: 10.1177/0269215507077285.

Nagle, K. F., and T. L. Eadie. 2012. Listener effort for highly intelligible tracheoesophageal speech. *Journal of Communication Disorders* 45(3):235–245. doi: 10.1016/j.jcomdis.2012.01.001.

Nagle, K. F., and T. L. Eadie. 2018. Perceived listener effort as an outcome measure for disordered speech. *Journal of Communication Disorders* 73(May–June):34–49. doi: 10.1016/j.jcomdis.2018.03.003.

Newman, C. W., B. E. Weinstein, G. P. Jacobson, and G. A. Hug. 1990. The Hearing Handicap Inventory for Adults: Psychometric adequacy and audiometric correlates. *Ear and Hearing* 11(6):430–433.

Newman, C. W., B. E. Weinstein, G. P. Jacobson, and G. A. Hug. 1991. Test-retest reliability of the Hearing Handicap Inventory for Adults. *Ear and Hearing* 12(5):355–357.

NIDCD (National Institute on Deafness and Other Communication Disorders). 2015. *NIDCD Fact Sheet|Voice, Speech, and Language: Aphasia*. Publication no. 97–4257. https://www.nidcd.nih.gov/sites/default/files/Documents/health/voice/Aphasia.pdf (accessed April 4, 2019).

Nielsen, L. M., H. Kirkegaard, L. G. Østergaard, K. Bovbjerg, K. Breinholt, and T. Maribo. 2016. Comparison of self-reported and performance-based measures of functional ability in elderly patients in an emergency department: Implications for selection of clinical outcome measures. *BMC Geriatrics* 16:199. doi: 10.1186/s12877-016-0376-1.

NINDS (National Institute of Neurological Disorders and Stroke). 2011. *NIH Stroke Scale*. Bethesda, MD: National Institute of Neurological Disorders and Stroke, U.S. Department of Health and Human Services.

NRC (National Research Council). 1987. *Night vision: Current research and future directions, symposium proceedings*. Washington, DC: National Academy Press.

NRC. 2002. *Visual impairments: Determining eligibility for Social Security benefits*. Washington, DC: National Academy Press.

NYHA (Criteria Committee of the New York Heart Association). 1994. *Nomenclature and criteria for diagnosis of diseases of the heart and great vessels*, 9th edition. Boston, MA: Little, Brown & Co.

Oesch, P., K. Meyer, B. Jansen, P. Mowinckel, S. Bachmann, and K. B. Hagen. 2012. What is the role of "nonorganic somatic components" in functional capacity evaluations in patients with chronic nonspecific low back pain undergoing fitness for work evaluation? *Spine* 37(4):243–250.

Overbeek, C. L., S. P. Nota, P. Jayakumar, M. G. Hageman, and D. Ring. 2015. The PROMIS physical function correlates with the QuickDASH in patients with upper extremity illness. *Clinical Orthopaedics and Related Research* 473(1):311–317. doi: 10.1007/s11999-014-3840-2.

Packham, T., and J. C. MacDermid. 2013. Measurement properties of the Patient-Rated Wrist and Hand Evaluation: Rasch analysis of responses from a traumatic hand injury population. *Journal of Hand Therapy* 26(3):216–223. doi: 10.1016/j.jht.2012.12.006.

Patel, R., N. Usher, H. Kember, S. Russell, and J. Laures-Gore. 2014. The influence of speaker and listener variables on intelligibility of dysarthric speech. *Journal of Communication Disorders* 51(2014):13–18. doi: 10.1016/j.jcomdis.2014.06.006.

Patel, R. R., S. N. Awan, J. Barkmeier-Kraemer, M. Courey, D. Deliyski, T. Eadie, D. Paul, J. G. Švec, and R. Hillman. 2018. Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech–Language Pathology* 27(3):887–905. doi: 10.1044/2018_AJSLP-17-0009.

Paul, D. R., C. Frattali, A. L. Holland, C. K. Thompson, C. J. Caperton, and S. C. Slater. 2004. *Quality of Communication Life Scale: Manual*. Rockville, MD: American Speech-Language-Hearing Association.

Peel, N. M., S. S. Kuys, and K. Klein. 2012. Gait speed as a measure in geriatric assessment in clinical settings: A systematic review. *Journal of Gerontology* 68(1):39–46.

Pepin, V., J. L. Alexander, and W. T. Phillips. 2004. Physical function assessment in cardiac rehabilitation: Self-report, proxy-report and performance-based measures. *Journal of Cardiopulmonary Rehabilitation and Prevention* 24(5):287–295.

Peterson, E. A., N. Roy, S. N. Awan, R. M. Merrill, R. Banks, and K. Tanner. 2013. Toward validation of the Cepstral Spectral Index of Dysphonia (CSID) as an objective treatment outcomes measure. *Journal of Voice* 27(4):401–410.

Picard, M., S. A. Girard, M. Simard, R. Larocque, T. Leroux, and F. Turcotte. 2008. Association of work-related accidents with noise exposure in the workplace and noise-induced hearing loss based on the experience of some 240,000 person-years of observation. *Accident Analysis & Prevention* 40(5):1644–1652.

Poburka, B. J. 1999. A new stroboscopy rating form. *Journal of Voice* 13(3):403–413.

Porch, B. E. 2001. *The Porch Index of Communicative Ability*, 3rd ed. Palo Alto, CA: Consulting Psychologists Press.

Porreca, F., and M. H. Ossipov. 2009. Nausea and vomiting side effects with opioid analgesics during treatment of chronic pain: Mechanisms, implications, and management options. *Pain Medicine* 10(4):654–662.

Pritchard, M., K. Hilari, N. Cocks, and L. Dipper. 2018. Psychometric properties of discourse measures in aphasia: Acceptability, reliability, and validity. *International Journal of Language & Communication Disorders* 53(6):1078–1093. doi: 10.1111/1460-6984.12420.

Prutting, C. A., and D. M. Kirchner. 1987. A clinical appraisal of the pragmatic aspects of language. *Journal of Speech and Hearing Disorders* 52(2):105–119.

PRWE (Patient-Rated Wrist Evaluation). 2011. *The Patient-Rated Wrist Evaluation (PRWE) user manual, June 2011*. https://srs-mcmaster.ca/wp-content/uploads/2015/05/English-PRWE-User-Manual.pdf (accessed April 4, 2019).

Przewozny, T. 2016. Possibilities of spatial hearing testing in occupational medicine. *International Journal of Occupational Medicine and Environmental Health* 29(4):527–538. doi: 10.13075/ijomeh.1896.00537.

Przysiezny, P. E., and L. T. S. Przysiezny. 2015. Work-related voice disorder. *Brazilian Journal of Otorhinolaryngology* 81(2):202–211. doi: 10.1016/j.bjorl.2014.03.003.

Punch, J. L., D. O. Robinson, and D. F. Katt. 1996. Development of a hearing performance standard for law enforcement officers. *Journal of the American Academy of Audiology* 7(2):113–119.

Raghavendra, P., J. Bornman, M. Granlund, and E. Björck-Åkesson. 2007. The World Health Organization's *International Classification of Functioning, Disability and Health*: Implications for clinical and research practice in the field of augmentative and alternative communication. *Augmentative and Alternative Communication* 23(4):349–361. doi: 10.1080/07434610701650928.

Raphael, C., C. Briscoe, J. Davies, Z. I. Whinnett, C. Manistry, R. Sutton, K. J. Mayet, and D. P. Francis. 2007. Limitations of the New York Heart Association functional classification system and self-reported walking distances in chronic heart failure. *Heart* 93(4):476–482.

Ratcliff, A. 1994. Comparison of relative demands implicated in direct selection and scanning: Considerations from normal children. *Augmentative and Alternative Communication* 10(2):67–74.

Rector, T. S., I. S. Anand, and J. N. Cohn. 2006. Relationships between clinical assessments and patient perceptions of the effects of heart failure on their quality of life. *Journal of Cardiac Failure* 12(2):87–92.

Reiman, M. P., and R. C. Manske. 2011. The assessment of function: How is it measured? A clinical perspective. *Journal of Manual & Manipulative Therapy* 19(2):91–99.

Riley, G. D., and K. Bakker. 2009. *Stuttering Severity Instrument: SSI-4*. Austin, TX: PRO-ED.

Rocchi, M. B., D. Sisti, P. Benedetti, M. Valentini, S. Bellagamba, and A. Federici. 2005. Critical comparison of nine different self-administered questionnaires for the evaluation of disability caused by low back pain. *Europa Medicophysica* 41(4):275–281.

Roeser, R. J., K. A. Buckley, and G. S. Stickney. 2000. Pure tone tests. In *Audiology diagnosis*, edited by R. J. Roeser, M. Valente, and H. Hosford-Dunn. New York: Thieme. Pp. 227–251.

Roland, M., and J. Fairbank. 2000. The Roland–Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine* 25(24):3115–3124.

Roland, M., and R. Morris. 1983. A study of the natural history of back pain. *Spine* 8(2):141–144. doi: 10.1097/00007632-198303000-00004.

Rosen, C. A., A. S. Lee, J. Osborne, T. Zullo, and T. Murry. 2004. Development and validation of the Voice Handicap Index-10. *The Laryngoscope* 114(9):1549–1556.

Ross, K. B., and R. T. Wertz. 2003. Discriminative validity of selected measures for differentiating normal from aphasic performance. *American Journal of Speech-Language Pathology* 12(3):312–319. doi: 10.1044/1058-0360(2003/077).

Roy, N., J. Barkmeier-Kraemer, T. Eadie, M. P. Sivasankar, D. Mehta, D. Paul, and R. Hillman. 2013. Evidence based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology* 22(2):212–226. doi: 10.1044/1058-0360(2012/12-0014).

Rustenburg, G., P. P. Kuijer, and M. H. Frings-Dresen. 2004. The concurrent validity of the ERGOS Work Simulator and the Ergo-Kit with respect to maximum lifting capacity. *Journal of Occupational Rehabilitation* 14(2):107–118.

Sager, M., N. Dunham, A. Schwantes, L. Mecum, K. Halverson, and D. Harlowe. 1992. Measurement of activities of daily living in hospitalized elderly: A comparison of self-report and performance-based methods. *Journal of the American Geriatrics Society* 40(5):457–462.

Shauver, M. J., and K. C. Chung. 2013. The Michigan Hand Outcomes Questionnaire (MHQ) after 15 years of field trial. *Plastic and Reconstructive Surgery* 131(5):779e.

Siew, C. S. Q., K. M. Pelczarski, J. S. Yaruss, and M. S. Vitevitch. 2017. Using the OASES-A to illustrate how network analysis can be applied to understand the experience of stuttering. *Journal of Communication Disorders* 65(January–February):1–9. doi: 10.1016/j.jcomdis.2016.11.001.

Simeonsson, R. J., E. Björck-Åkessön, and D. J. Lollar. 2012. Communication, disability, and the ICF-CY. *Augmentative and Alternative Communication* 28(1):3–10. doi: 10.3109/07434618.2011.653829.

Simmons-Mackie, N., A. Kagan, J. C. Victor, A. Carling-Rowland, A. Mok, J. S. Hoch, M. Hujibregts, and D. L. Streiner. 2014. The assessment for living with aphasia: Reliability and construct validity. *International Journal of Speech-Language Pathology* 16(1):82–94.

Singer, M., and E. Blom. 1980. An endoscopic technique for restoration of voice after laryngectomy. *Annals of Otology, Rhinology & Laryngology* 89(6 Pt. 1):529–533.

Singh, G., and M. K. Pichora-Fuller. 2010. Older adults' performance on the Speech, Spatial, and Qualities of Hearing Scale (SSQ): Test-retest reliability and a comparison of interview and self-administration methods. *International Journal of Audiology* 49(10):733–740.

Sivan, M., B. Sell, and P. Sell. 2009. A comparison of functional assessment instruments and work status in chronic back pain. *European Journal of Physical and Rehabilitation Medicine* 45(1):31–36.

Smeets, R., A. Köke, C.-W. Lin, M. Ferreira, and C. Demoulin. 2011. Measures of function in low back pain/disorders: Low Back Pain Rating Scale (LBPRS), Oswestry Disability Index (ODI), Progressive Isoinertial Lifting Evaluation (PILE), Quebec Back Pain Disability Scale (QBPDS), and Roland-Morris Disability Questionnaire (RDQ). *Arthritis Care and Research* 63(Suppl. 11):S158–S173. doi: 10.1002/acr.20542.

Smits, C., T. S. Kapteyn, and T. Houtgast. 2004. Development and validation of an automatic speech-in-noise screening test by telephone. *International Journal of Audiology* 43(1):15–28.

Smits, C., P. Merkus, and T. Houtgast. 2006. How we do it: The Dutch functional hearing–screening tests by telephone and Internet. *Clinical Otolaryngology* 31(5):436–440.

Soer, R., C. P. Van der Schans, J. W. Groothoff, J. H. Geertzen, and M. F. Reneman. 2008. Towards consensus in operational definitions in functional capacity evaluation: A Delphi Survey. *Journal of Occupational Rehabilitation* 18(4):389–400.

Soli, S. D., A. Amano-Kusumoto, O. Clavier, J. Wilbur, K. Casto, D. Freed, C. Laroche, V. Vaillancourt, C. Giguère, W. A. Dreschler, and K. S. Rhebergen. 2018a. Evidence-based occupational hearing screening II: Validation of a screening methodology using measures of functional hearing ability. *International Journal of Audiology* 57(5):323–334.

Soli, S. D., C. Giguère, C. Laroche, V. Vaillancourt, W. A. Dreschler, K. S. Rhebergen, K. Harkins, M. Ruckstuhl, P. Ramulu, and L. S. Meyers. 2018b. Evidence-based occupational hearing screening I: Modeling the effects of real-world noise environments on the likelihood of effective speech communication. *Ear and Hearing* 39(3):436–448.

Speksnijder, C. M., T. Koppenaal, J. A. Knottnerus, M. Spigt, J. B. Staal, and C. B. Terwee. 2016. Measurement properties of the Quebec Back Pain Disability Scale in patients with nonspecific low back pain: Systematic review. *Physical Therapy* 96(11):1816–1831.

Spertus, J. A., and P. G. Jones. 2015. Development and validation of a short version of the Kansas City Cardiomyopathy Questionnaire. *Circulation: Cardiovascular Quality Outcomes* 8(5):469–476.

Spertus, J. A., J. A. Winder, T. A. Dewhurst, R. A. Deyo, J. Prodzinski, M. McDonnell, and S. D. Fihn. 1995. Development and evaluation of the Seattle Angina Questionnaire: A new functional status measure for coronary artery disease. *Journal of the American College of Cardiology* 25(2):333–341.

SSA (U.S. Social Security Administration). 2008. *Listing of impairments—Adult listings (Part A)*. 4.00 Cardiovascular system. https://www.ssa.gov/disability/professionals/bluebook/4.00-Cardiovascular-Adult.htm (accessed April 4, 2019).

SSA. 2018. *OIS project frequently asked questions*. https://www.ssa.gov/disabilityresearch/ois_project_faqs.html (accessed April 3, 2019).

Stelzle, F., C. Knipfer, M. Schuster, T. Bocklet, E. Nöth, W. Adler, L. Schempf, P. Vieler, M. Riemann, F. W. Neukam, and E. Nkenke. 2013. Factors influencing relative speech intelligibility in patients with oral squamous cell carcinoma: A prospective study using automatic, computer-based speech analysis. *International Journal of Oral and Maxillofacial Surgery* 42(11):1377–1384. doi: 10.1016/j.ijom.2013.05.021.

Stevens, M. L., C. C.-W. Lin, and C. G. Maher. 2016. The Roland Morris Disability Questionnaire. *Journal of Physiotherapy* 62(2):116. doi: 10.1016/j.jphys.2015.10.003.

Stewart, W. F., J. A. Ricci, E. Chee, D. Morganstein, and R. Lipton. 2003. Lost productive time and cost due to common pain conditions in the US workforce. *JAMA* 290(18):2443–2454.

Strand, E. A., J. R. Duffy, H. M. Clark, and K. Josephs. 2014. The Apraxia of Speech Rating Scale: A tool for diagnosis and description of apraxia of speech. *Journal of Communication Disorders* 51(September–October):43–50. doi: 10.1016/j.jcomdis.2014.06.008.

Swinburn, K., G. Porter, and D. Howard. 2004. *CAT: Comprehensive Aphasia Test*. Hove, England: Psychology Press.

Tanner, D., and W. Culbertson. 1999. *Quick Assessment for Dysphagia*. Oceanside, CA: Academic Communication Associates.

Tarlov, A. R., J. E. Ware, S. Greenfield, E. C. Nelson, E. Perrin, and M. Zubkoff. 1989. The Medical Outcomes Study: An application of methods for monitoring the results of medical care. *JAMA* 262(7):925–930.

Tiffin, J., and E. J. Asher. 1948. The Purdue Pegboard: Norms and studies of reliability and validity. *Journal of Applied Psychology* 32(3):234–247.

Touré-Tillery, M., and A. Fishbach. 2014. How to measure motivation: A guide for the experimental social psychologist. *Social and Personality Psychology Compass* 8(7):328–341.

Tsuji, T., K. Matsudaira, H. Sato, J. Vietri, and D. H. Jaffe. 2018. Association between presenteeism and health-related quality of life among Japanese adults with chronic lower back pain: A retrospective observational study. *BMJ Open* 8(6):e021160.

Tufts, J. B., K. A. Vasil, and S. Briggs. 2009. Auditory fitness for duty: A review. *Journal of the American Academy of Audiology* 20(9):539–557.

Tye-Murray, N. 2014. *Foundations of aural rehabilitation: Children, adults, and their family members*. Stamford, CT: Nelson Education.

University of Michigan. 2014a. *MHQ Michigan Hand Outcomes Questionnaire: The MHQ*. http://mhq.lab.medicine.umich.edu/mhq (accessed April 4, 2019).

University of Michigan. 2014b. *MHQ Michigan Hand Outcomes Questionnaire: Translations*. http://mhq.lab.medicine.umich.edu/translations (accessed April 4, 2019).

Varma, R., T. S. Vajaranant, B. Burkemper, S. Wu, M. Torres, C. Hsu, F. Choudhury, and R. McKean-Cowdin. 2016. Visual impairment and blindness in adults in the United States: Demographic and geographic variations from 2015 to 2050. *JAMA Ophthalmology* 134(7):802–809.

Vernon, H. 2008. The Neck Disability Index: State-of-the-art, 1991–2008. *Journal of Manipulative and Physiological Therapeutics* 31(7):491–502.

Vianin, M. 2008. Psychometric properties and clinical usefulness of the Oswestry Disability Index. *Journal of Chiropractic Medicine* 7(4):161–163. doi: 10.1016/j.jcm.2008.07.001.

Vincent, J. I., J. C. MacDermid, G. J. King, and R. Grewal. 2015. Rasch analysis of the Patient Rated Elbow Evaluation questionnaire. *Health and Quality of Life Outcomes* 13:84. doi: 10.1186/s12955-015-0275-8.

Vowles, K. E., R. T. Gross, and J. T. Sorrell. 2004. Predicting work status following interdisciplinary treatment for chronic pain. *European Journal of Pain* 8(4):351–358.

Walker, M. L., A. G. Austin, G. M. Banke, S. R. Foxx, L. Gaetano, L. A. Gardner, J. McElhiney, K. Morris, and L. Penn. 2007. Reference group data for the Functional Gait Assessment. *Physical Therapy* 87(11):1468–1477.

Wallace, S., L. Worrall, T. Rose, and G. Le Dorze. 2014. Measuring outcomes in aphasia research: A review of current practice and an agenda for standardisation. *Aphasiology* 28(11):1364–1384. doi: 10.1080/02687038.2014.930262.

Wang, Y. F., S. S. Wang, C. C. Tai, L. C. Lin, and A. S. Shiao. 2002. Hearing screening with portable screening pure-tone audiometer and distortion-product otoacoustic emissions. *Zhonghua Yi Xue Za Zhi [Chinese Medical Journal]* 65(6):285–292.

Wasserman, K., J. E. Hansen, D. Y. Sue, B. J. Whipp, and R. Casaburi 1994. *Principles of exercise testing and interpretation*, 2nd ed. Philadelphia, PA: Lea & Febiger.

Watson, C. S., G. R. Kidd, J. D. Miller, C. Smits, and L. E. Humes. 2012. Telephone screening tests for functionally impaired hearing: Current use in seven countries and development of a U.S. version. *Journal of the American Academy of Audiology* 23(10):757–767.

Weber, C., M. Schwieterman, K. Fier, J. Berni, N. Swartz, R. S. Phillips, and J. C. Reneker. 2016. Reliability and validity of the Functional Gait Assessment: A systematic review. *Physical & Occupational Therapy in Geriatrics* 34(1):88–103. doi: 10.3109/02703181.2015.1128509.

Wenzel, E. M., M. Arruda, D. J. Kistler, and F. L. Wightman. 1993. Localization using non-individualized head-related transfer functions. *The Journal of the Acoustical Society of America* 94(1):111–123.

Westaway, M. D., P. W. Stratford, and J. M. Binkley. 1998. The Patient Specific Functional Scale: Validation of its use in persons with neck dysfunction. *Journal of Orthopaedic & Sports Physical Therapy* 27(5):331–338. doi: 10.2519/jospt.1998.27.5.331.

White, A. R., E. Carney, and J. Reichle. 2010. Group-item and directed scanning: Examining preschoolers' accuracy and efficiency in two augmentative communication symbol selection methods. *American Journal of Speech-Language Pathology* 19(4):311–320.

Whitehouse, A. J. O., and D. V. M. Bishop. 2009. *Communication Checklist—Adult*. London: The Psychological Corporation.

WHO (World Health Organization). 2001. *International classification of functioning, disability and health*. Geneva, Switzerland: WHO.

WHO. 2002. *Towards a common language for functioning, disability and health: ICF*. Geneva, Switzerland: WHO.

Wilkinson, M., G. Dagnelie, and C. A. Johnson. 2018. *Visual functions to test for disability determinations*. Paper commissioned by the Committee on Functional Assessment for Adults with Disabilities, National Academies of Sciences, Engineering, and Medicine, Washington, DC.

Williams, F. 1970. *Language and poverty: Perspectives on a theme*. New York: Academic Press.

Williams, K. 2018. *Expressive Vocabulary Test-3*. San Antonio, TX: Pearson.

Williamson, A., and B. Hoggart. 2005. Pain: A review of three commonly used pain rating scales. *Journal of Clinical Nursing* 14(7):798–804.

Wind, H., V. Gouttebarge, P. P. F. Kuijer, and M. H. Frings-Dresen. 2005. Assessment of functional capacity of the musculoskeletal system in the context of work, daily living, and sport: A systematic review. *Journal of Occupational Rehabilitation* 15(2):253–272.

Wong, J. Y., B. K. Fung, M. M. Chu, and R. K. Chan. 2007. The use of Disabilities of the Arm, Shoulder, and Hand Questionnaire in rehabilitation after acute traumatic hand injuries. *Journal of Hand Therapy* 20(1):49–55.

Woods, W. S., S. Kalluri, S. Pentony, and N. Nooraei. 2013. Predicting the effect of hearing loss and audibility on amplified speech reception in a multi-talker listening scenario. *The Journal of the Acoustical Society of America* 133(6):4268–4278.

Wrisley, D. M., and N. A. Kumar. 2010. Functional Gait Assessment: Concurrent, discriminative, and predictive validity in community-dwelling older adults. *Physical Therapy* 90(5):761–773. doi: 10.2522/ptj.20090069.

Wrisley, D. M., G. F. Marchetti, D. K. Kuharsky, and S. L. Whitney. 2004. Reliability, internal consistency, and validity of data obtained with the Functional Gait Assessment. *Physical Therapy* 84(10):906–918.

Yaruss, J. S. 2010. Assessing quality of life in stuttering treatment outcomes research. *Journal of Fluency Disorders* 35(3):190–202.

Yaruss, J. S., and R. W. Quesal. 2006. Overall Assessment of the Speaker's Experience of Stuttering (OASES): Documenting multiple outcomes in stuttering treatment. *Journal of Fluency Disorders* 31(2):90–115.

Yates, M., and N. Shastri-Hurst. 2017. The Oswestry Disability Index. *Occupational Medicine* 67(3):241–242.

Yorkston, K. M., and D. R. Beukelman. 1981. Communication efficiency of dysarthric speakers as measured by sentence intelligibility and speaking rate. *Journal of Speech and Hearing Disorders* 46(3):296–301.

Yorkston, K. M., D. R. Beukelman, and C. Traynor. 1984. *Assessment of Intelligibility of Dysarthric Speech*. Austin, TX: PRO-ED.

Yorkston, K., P. Dowden, and D. Beukelman. 1992. Intelligibility measurement as a tool in the clinical management of dysarthric speakers. In *Intelligibility in speech disorders: Theory, measurement and management*, edited by R. D. Kent. Amsterdam, The Netherlands: John Benjamin Publishing Company. Pp. 265–285.

Yorkston, K. M., D. Beukelman, and M. Hakel. 1996. *Speech Intelligibility Test for windows*. Lincoln, NE: Madonna Rehabilitation Hospital.

Yorkston, K. M., D. R. Beukelman, M. E. Hakel, and M. Dorsey. 2007. *Speech Intelligibility Test (SIT)*. Lincoln, NE: Madonna Rehabilitation Hospital. https://www.madonna.org/research/software (accessed April 8, 2019).

Yorkston, K. M., D. R. Beukelman, E. A. Strand, and M. Hakel. 2010. *Management of motor speech disorders in children and adults*, 3rd ed. Austin, TX: PRO-ED.

Yueh, B., N. Shapiro, C. H. MacLean, and P. G. Shekelle. 2003. Screening and management of adult hearing loss in primary care: Scientific review. *JAMA* 289(15):1976–1985.

Zraick, R. I., G. B. Kempster, N. P. Connor, S. Thibeault, B. K. Klaben, Z. Bursac, C. R. Thrush, and L. E. Glaze. 2011. Establishing validity of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). *American Journal of Speech-Language Pathology* 20(1):14–22. doi: 10.1044/1058-0360(2010/09-0105).

ANNEX TABLE 5-1
Data Elements on Physical Demands in the Occupational Requirements Survey (ORS)

Category	Data Element	Definition
Sitting Versus Standing/Walking	Sitting	• Remains in a seated position; includes active sitting (e.g., cycling) • Lying down • Option of choosing between sitting and standing (DOL, 2018, p. 92)
	Standing/walking	• Whenever workers are not sitting or lying down, includes time spent stooping, crawling, kneeling, crouching, or climbing (DOL, 2018, p. 92)
	Sitting/standing at will	• Flexibility to choose between sitting and standing throughout the day • No assigned time during the day to sit or stand • No external factors determine whether workers must sit or stand (DOL, 2018, p. 95)
Lifting/Carrying	Lifting	• "Raising or lowering an object from one level to another"; can include an upward pulling motion (DOL, 2018, p. 97)
	Carrying	• "Transporting an object, usually by holding it in the hands or arms, or wearing it on the body, usually around the waist or upper torso" (DOL, 2018, p. 97)
Pushing/Pulling (separated based on the part of the body used: hands/arms, feet/legs [DOL, 2018, p. 105])	Pushing	• "Exerting force upon an object so that the object moves away from the origin of the force" (DOL, 2018, p. 103)
	Pulling	• "Exerting force upon an object so that the object moves toward the origin of the force" (DOL, 2018, p. 103)
Reaching	Reaching	• "Extending the hand(s) and arm(s) in any direction, requiring the straightening and extending of the arm(s) and elbow(s) and the engagement of the shoulder(s)" (DOL, 2018, p. 107)
	Overhead reaching	• "Extending the arm(s) with the hand higher than the head" *and* — the elbow is bent and "the angle at the shoulder is about 90 degrees or more" *or* — the elbow is extended and "the angle at the shoulder is about 120 degrees or more" (DOL, 2018, p. 107)
	At/below the shoulder reaching	• Reaching that "does not meet the threshold for overhead" (DOL, 2018, p. 107)

continued

ANNEX TABLE 5-1
Continued

Category	Data Element	Definition
Manipulation	Gross manipulation	• "Seizing, holding, grasping, turning, or otherwise working with the hand(s)" • "Fingers are involved only to the extent that they are an extension of the hand to hold or operate an object or tool, such as a hammer" (DOL, 2018, p. 109)
	Fine manipulation	• "Touching, picking, pinching, or otherwise working primarily with fingers rather than with the whole hand or arm" (DOL, 2018, p. 109)
	Foot/leg controls	• "The use of one or both feet or legs to move controls on machinery or equipment. Controls include, but are not limited to, pedals, buttons, levers, and cranks" (DOL, 2018, p. 109)
Keyboarding		• "Entering text or data into a computer or other machine by means of a keyboard, using a repetitive motion requiring the use of the whole hand" • Refers to use of a traditional keyboard, i.e., "a panel of keys used as the primary input device on a computer, typographic machine or 10-key numeric keypad" (DOL, 2018, p. 112)
Stooping, Crouching, Kneeling, Crawling	Stooping	• "Bending the body forward and down while bending the spine at the waist 45 degrees or more either over something below waist level or down towards an object on or near the ground" • "Must be performed while standing" (DOL, 2018, p. 114)
	Crouching	• "Bending the body downward and forward by bending the legs and spine" (DOL, 2018, p. 114)
	Kneeling	• "Bending the legs at the knees to come to rest on the knee or knees" (DOL, 2018, p. 114)
	Crawling	• "Moving about on hands and knees or hands and feet" (DOL, 2018, p. 114)

ANNEX TABLE 5-1
Continued

Category	Data Element	Definition
Climbing	Climbing	• "The act of ascending or descending stairs, ramps, ladders, ropes or scaffolding and similar structures using feet, legs, hands, and/or arms" (DOL, 2018, p. 119)
	Climbing ramps or stairs	• Ascending or descending "ramps or stairs primarily using feet and legs ... arms and hands for balance only, as in holding a stair railing" (DOL, 2018, p. 119)
	Climbing ladders, ropes, or scaffolds	• Ascending or descending "ladders, scaffolding, ropes, or poles, using feet/legs and hands/arms" • Typically involves "both upper body and lower body" • Includes climbing "something that requires the use of both the upper and lower body to climb" (DOL, 2018, p. 119)
Vision	Near visual acuity	• "Clarity of vision at approximately 20 inches or less, as when working with small objects or reading small print" • Includes use of "a computer in support of a critical job function, regardless of distance" (DOL, 2018, p. 130)
	Far visual acuity	• "Clarity of vision at a distance of 20 feet or more, involving the ability to distinguish features of a person or objects at a distance" (DOL, 2018, p. 130)
	Peripheral vision	• "What is seen above, below, to the left or right by the eye while staring straight ahead" (DOL, 2018, p. 130)
Speaking and Hearing Requirements	Speaking	• "Expressing or exchanging ideas by means of the spoken word to impart oral information to clients or the public and to convey detailed verbal instructions to other workers accurately, loudly, or quickly" (DOL, 2018, p. 125)
	Hearing requirements	• "Ability to hear, understand, and distinguish speech and/or other sounds" (DOL, 2018, p. 125) • Includes ability to hear speech in person, both one on one and in group or conference; speech through a telephone; speech through other remote communications devices such as radios, walkie-talkies, intercoms, and public address systems; other sounds, such as machinery alarms and equipment sounds (DOL, 2018, pp. 127–128)

SOURCE: DOL, 2018.

ANNEX TABLE 5-2
Physical Functional Abilities Relevant to Work Requirements

Relevant Functional Domains (identified by the committee)	Task	Occupational Requirements Survey		
Lower Extremities and Feet, Back and Neck	Walk	Standing/walking		
	Stand			
Back and Neck	Sit	Sitting		
		Sitting/standing at will		
Lower Extremities and Feet, Upper Extremities, and Hands and Fingers	Climb	Climbing ramps or stairs		
		Climbing ladders, ropes, scaffolds		
Back and Neck	Stoop (bend down and forward at waist)	Stooping		
Lower Extremities and Feet	Kneel (bend legs to rest on knees)	Kneeling		
Lower Extremities and Feet, Back and Neck	Crouch (bend legs and back down and forward)	Crouching		
Lower Extremities and Feet, Upper Extremities and Hands and Fingers	Crawl (move on hands and knees)	Crawling		
Lower Extremities and Feet		Manipulation	Foot/leg controls	
Upper Extremities and Hands and Fingers	Handle large objects	Manipulation	Gross manipulation	
Upper Extremities and Hands and Fingers	Write, type, or handle small objects	Manipulation	Fine manipulation	
		Keyboarding		

	Occupational Information Network (O*NET)		Disability Report-Adult (SSA-3368-BK)	Work History Report (SSA-3369-BK)	Function Report-Adult (SSA-3373-BK)	Physical Residual Functional Capacity Form (SSA-4734-BK)
	Gross body coordination	Trunk strength	Walk	Walk	Walking	Addressed in terms of frequency or number of hours under "exertional limitations"
			Stand	Stand	Standing	
	Trunk strength		Sit	Sit	Sitting	
	Gross body coordination		Climb	Climb	Stair climbing	Climbing ramps or stairs
						Climbing ladders, ropes, scaffolds
	Extent flexibility		Stoop	Stoop	Bending	Stooping
	Extent flexibility		Kneel	Kneel	Kneeling	Kneeling
	Extent flexibility		Crouch	Crouch	Squatting	Crouching
	Gross body coordination		Crawl	Crawl		Crawling
	Gross body equilibrium					Balancing (U.S. Department of Transportation [DOT] element excluded from Occupational Information System [OIS] [SSA, 2018])
	Manual dexterity		Handle large objects	Handle large objects	Using hands	Handling
	Finger dexterity		Write, type, or handle small objects	Write, type, or handle small objects	Using hands	Fingering
	Finger dexterity	Wrist-finger speed				Feeling (skin receptors)

continued

ANNEX TABLE 5-2
Continued

Relevant Functional Domains (identified by the committee)	Task	Occupational Requirements Survey
Upper Extremities and Hands and Fingers	Reach	Overhead reaching
		At/below shoulder reaching
Upper Extremities and Hands and Fingers, Back and Neck Lower Extremities and Feet (for carry)	Lift	Lifting/carrying
	Carry	
Upper Extremities and Hands and Fingers, Back and Neck, Lower Extremities and Feet	Push	Pushing/pulling (exerting force, measure of strength)
	Pull	
Vision	See	Near visual acuity
		Far visual acuity
		Peripheral vision
Communication[a]	Hear	Hearing • In person speech (one on one, group/conference) • Telephone • Other remote speech (e.g., radio, intercom, etc.) • Other sounds (e.g., alarms and other nonverbal sounds)
	Speak	Speaking

[a]Communication: Here refers to anatomical impairments that affect hearing and speech. Certain brain/cognitive disorders (e.g., expressive/receptive aphasias) also can affect communication.

	Occupational Information Network (O*NET)			Disability Report-Adult (SSA-3368-BK)	Work History Report (SSA-3369-BK)	Function Report-Adult (SSA-3373-BK)	Physical Residual Functional Capacity Form (SSA-4734-BK)
	Dynamic flexibility	Extent flexibility	Multilimb coordination	Reach	Reach	Reaching	Reaching all directions
	Static strength		Multilimb coordination	Lifting and carrying	Lifting and carrying	Lifting	Exertional limitations
	Static strength		Multilimb coordination				Exertional limitations
	Near vision					Seeing	Near acuity
	Far vision						Far acuity
	Peripheral vision						Field of vision (visual fields)
	Depth perception						Depth perception
							Accommodation
	Night vision						
	Glare sensitivity						
	Visual color discrimination						Color vision (DOT element excluded from OIS [SSA, 2018])
	Auditory attention					Hearing	Hearing
	Hearing sensitivity						
	Sound localization						
	Speech recognition						
	Speech clarity					Talking	Speaking

continued

ANNEX TABLE 5-2
Continued

Relevant Functional Domains (identified by the committee)	Task	Occupational Requirements Survey	
Strength and Endurance	Push	Pushing/pulling	
	Pull		
	Lift	Lifting/carrying	
	Carry		
	Stand and/or walk	Standing/walking	
	Sit	Sitting	
		Sitting/standing at will	
Stamina			

Occupational Information Network (O*NET)	Disability Report-Adult (SSA-3368-BK)	Work History Report (SSA-3369-BK)	Function Report-Adult (SSA-3373-BK)	Physical Residual Functional Capacity Form (SSA-4734-BK)
Static strength				Exertional limitations
Dynamic strength	Lifting and carrying	Lifting and carrying	Lifting	
Dynamic strength				
Dynamic flexibility				
Trunk strength				
Stamina				

ANNEX TABLE 5-3
Selected General Assessments for Physical Function

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How Administered	
Full-Body Assessments	Functional capacity evaluation (FCE)	Chen, 2007; Fore, 2015; Genovese and Galper, 2009; Jahn et al., 2004; Kuijer et al., 2012; Soer et al., 2008	Physical therapist, occupational therapist, exercise physiologist	Provide instructions to patients and observe task completion	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
	Usually at least 4–6 hours and often administered over 2 days	Often reliable; limited validity.	There are approximately 10 commonly utilized FCE instruments: Blankenship, ERGOS Work Simulator, Ergo-Kit variation, WorkWell Systems (formerly known as Isernhagen Work Systems), Hanoun Medical, Physical Work Performance Evaluation (Ergoscience), WEST-EPIC, Key, ERGOS, and ARCON	Injured workers, patients participating in a rehabilitation program, individuals applying for a strenuous or skill-specific occupation (law enforcement, military, professional sports)	There are no quantifiable measures of employment or activities of daily living prior to a disability or whether a disability is task specific for each job position, defining and measuring the essential demands.

continued

ANNEX TABLE 5-3
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How Administered	
Full-Body Assessments (*continued*)	Work Disability Functional Assessment Battery (WD-FAB) Physical Function	Meterko et al., 2015	No specialized training	Self-administered	
		Meterko et al., 2018			

NOTE: ICC = intraclass correlation coefficient.

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
10 minutes or less	Score reliability criterion (≥0.85); median convergent correlation = 0.55; median discriminant correlation = 0.18.	No	Individuals with self-reported physical disabilities	Strong evidence of construct validity. Data suggest that the WD-FAB could be used to assess physical and behavioral health function related to work disability. Future work will focus on item replenishment and refinement to increase the overall performance of these scales.
	ICCs = 0.69–0.77 (general sample); 0.66–0.86 (disability sample); standard errors of mean for all scales indicated good discrimination; minimum detectable change 90 values = 3.41–10.55.			Although initially developed for use within the U.S. Social Security Administration, the WD-FAB 2.0 could be used for assessment and measurement of work-related physical and mental health function in other contexts as well.

ANNEX TABLE 5-4
Selected Musculoskeletal Assessments

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Upper Extremities and Hands and Fingers	Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH) and QuickDASH	Aasheim and Finsen, 2014	No specialized qualifications	Self-administered	
		Dixon et al., 2008			
		Jester et al., 2005			
		Beaton et al., 2001			
		Kennedy et al., 2013			
		Bilberg et al., 2012			

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
	Minutes	Spearman's correlation coefficient for mean DASH and QuickDASH = 0.965.	No	Individuals with upper-extremity injuries, pain, and/or status postsurgery	DASH is available in 27 languages.
		Valid and frequently used measure of upper-limb function.			
		Significantly higher DASH scores found for older workers, females, and manual workers.			
		Test-retest reliability (ICC = 0.96) exceeded guidelines. Responsiveness of DASH was comparable to or better than that of the joint-specific measures in the whole group and in each region. Evidence for validity, test-retest reliability, and responsiveness of DASH, and validity and responsiveness in both proximal and distal disorders, confirming its usefulness across the whole extremity.			
		QuickDASH English version performs well, with strong positive evidence for reliability and validity.			Information about the measurement properties of the cross-cultural adaptation versions is still lacking.
		Confirmed satisfactory concurrent, convergent, and face validity. Test-retest reliability coefficient 0.99 (95% CI: 0.98–0.99) confirmed.			

continued

ANNEX TABLE 5-4
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Upper Extremities and Hands and Fingers (*continued*)	Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function—Upper Extremity	Overbeek et al., 2015	No specialized qualifications	Self-administered; paper or computer based	
	Patient-Rated Elbow Evaluation (PREE)	Vincent et al., 2015	No specialized qualifications	Self-administered	
		MacDermid, 2001			
		John et al., 2007			
		Beauchemin et al., 2015			

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
	Minutes	Substantial qualitative and quantitative evidence has been gathered that supports the validity of PROMIS measures (http://www.healthmeasures.net/explore-measurement-systems/promis/measure-development-research/validation [accessed April 4, 2019]).	No	Individuals with upper-extremity pain, injury, or status postsurgery	Based on item response theory; questionnaires may be paper based or administered using computer adaptive testing; translated into several languages.
	Minutes	The three subscales appear to be robust.	No	Individuals with elbow pain, injury, surgery	Low patient burden.
		DASH ICC = 0.93; test-retest reliability (2–7 days); pain items ICC = 0.76 to 0.87; function items ICC = 0.60 to 0.88.			
		Test-retest reliability (3–4 days); pain items ICC = 0.56 to 0.76; function items ICC = 0.48 to 0.83.			
		ICC for reliability = 0.89 (95% CI: 0.79–0.94) for the PREE-French; for construct validity, excellent correlation between PREE-Fr and QuickDASH (0.89–0.96); good to excellent correlation between the PREE and Medical Expenditure Panel Survey (MEPS) = (0.70–0.95).			.

continued

ANNEX TABLE 5-4
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Upper Extremities and Hands and Fingers (*continued*)	Patient-Rated Wrist Evaluation (PRWE)	Packham and MacDermid, 2013	No specialized qualifications	Self-administered	
		MacDermid et al., 1998			
	Michigan Hand Outcomes Questionnaire (MHQ)	Shauver and Chung, 2013; University of Michigan, 2014b	No specialized qualifications	Self-administered	
		Chung et al., 1998			

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
	Minutes	α = 0.96; reliability = 0.95.	No	Individuals with hand injury	
		Test-retest reliability was excellent (2–7 days ICC ≥0.90, long-term [1 year] ICC = 0.91); pain subscale and excellent short-term and long-term reliability (ICC = 0.90, 0.91, respectively); function subscale demonstrated excellent short-term reliability (ICC 0.61–0.88).			Psychometric testing on individuals with distal radius and scaphoid fractures.
					PRWE has been translated into Chinese, Dutch, English, German, Hindi, and Japanese.
	15 minutes	ICC range = 0.85–0.96; test-retest reliability = 0.85–0.95; α = 0.84–0.93 for each domain.	Use requires completion of a license	Individuals with hand pain, injury, or status postsurgery	Chinese, Dutch, Farsi, French, German, Japanese, Korean, Polish, Portuguese, Spanish, Turkish.
		Test-retest reliability = 0.81 (aesthetics)–0.97 (activities of daily living [ADLs]); α = 0.86 (pain) and 0.97 (ADLs); correlation between scales gave evidence of construct validity.			Short time to complete; reliable and valid instrument for measuring hand outcomes; can be used in clinic setting with minimal burden to patients; has undergone rigorous psychometric testing.

continued

ANNEX TABLE 5-4
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Back and Neck	Roland-Morris Disability Questionnaire (RMDQ)	Chansirinukor et al., 2005		Self-administered	
		Stevens et al., 2016	No specialized qualifications	Self-administered	
		Smeets et al., 2011			
		Stevens et al., 2016	No training is required to administer or score the questionnaire	Face-to-face, electronically, or over the phone	
		Rocchi et al., 2005		Self-administered	
		Calmels et al., 2005		Self-administered	
		Davies and Nitz, 2013		Self-administered	
	Owestry Disability Index (ODI)	Vianin, 2008	No specialized qualifications	Self-administered	
		Fairbank and Pynsent, 2000			
		Roland and Fairbank, 2000			
		Davidson and Keating, 2002			

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
				Low patient burden.
5 minutes		No		RMDQ is cross-culturally adapted or translated for use in other countries.
	α = 0.84–0.96; test-retest reliability = 0.83–0.91.			
	Correlation with Quebec Scale: *r* = 0.6; α = 0.84–0.96; Test-retest reliability = 0.83–0.91.			
				Low patient burden.
				Strong content and construct validity, feasibility, linguistic adaptation and international use; low patient burden.
				Most sensitive for patients with mild to moderate disability; low patient burden.
5 minutes	α = 0.71–0.87.	No	Individuals with back pain, injury, status postsurgery	Responsiveness is reported to be high.
	α = 0.71–0.87; test-retest reliability *r* = 0.83–0.99.			
	α = 0.71–0.87.			
	ICC = 0.84–0.94.			

continued

ANNEX TABLE 5-4
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Back and Neck (*continued*)	Quebec Back Pain Disability Questionnaire (Quebec Scale)	Kopec et al., 1995		Self-administered	
		Hicks and Manal, 2009			
		Speksnijder et al., 2016			
	Neck Disability Index (NDI)	Vernon, 2008	No specialized training	Self-administered	
		Ackelman and Lindgren, 2002		Self-administered	
		Cleland et al., 2008		Self-administered	
		McCarthy et al., 2007		Self-administered	

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
	Test-retest reliability = 0.92; α = 0.96.			Evaluates functional disability in patients with chronic back pain.
	ICC = 0.94 (95% CI: 0.90, 0.97).			
				There is limited to moderate evidence for good reliability, validity, and responsiveness of the Quebec Scale for different language versions.
Minutes	Spearmans correlation coefficient revealed a significant association.	No		Translated in 22 languages.
	Levels of sensitivity, test-retest reliability, and validity were acceptable.			
	Test-retest = fair to moderate.			Low patient burden.
	α = 0.86, 95% CI: 0.82–0.89; test-retest reliability = high; ICC = 0.93, 95% CI: 0.86–0.97.			

continued

ANNEX TABLE 5-4
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Lower Extremities and Feet	Lower-Extremity Functional Scale (LEFS)	Mehta et al., 2016	No specialized qualifications	Self-administered	
		Binkley et al., 1999			
		AbilityLab, 2013a			
	Functional Gait Assessment (FGA)	Weber et al., 2016	Trained/ experienced individual, such as a physician or physical therapist	Instructions and demonstration—walking test	
		Wrisley and Kumar, 2010; Wrisley et al., 2004			
		Walker et al., 2007			
		Wrisley and Kumar, 2010			
		Leddy et al., 2011			
	Foot and Ankle Ability Measure (FAAM)	AbilityLab, 2015	No specialized training	Self-administered	
		Martin and Irrgang, 2007			

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
		Test-retest reliability = excellent.	No	Individuals with lower-extremity pain, injury, status postsurgery	
		Test-retest reliability = 0.94.			
	5 minutes				
	10 minutes	Interrater reliability high.	No	Individuals at risk of falls	
		For individuals with vestibular dysfunction; interrater reliability: r = 0.86; intrarater reliability: r = 0.74, α = 0.79; Spearman rank correlation coefficients of the FGA scores with balance measurements = 0.11–0.67.			
		For community-dwelling adults, interrater reliability: r = 0.93.			
		Appears to predict falls in community-dwelling older adults.			
		Test-retest reliability = 0.91; interrater reliability >0.93.			
	10 minutes		No		
		Activities of Daily Living subscale ICC = 0.87, α = 0.96; sports subscale ICC = 0.89, α = 0.98.			

continued

ANNEX TABLE 5-4
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Pain Assessments	Visual Analog Scale (VAS) for Pain	Bijur et al., 2001	Minimal training	Self-administered	
	Numeric Rating Scale (NRS) for Pain	AbilityLab, 2013b	No specialized training	Self-administered	
	Patient-Reported Outcomes Measurement Information System (PROMIS) Pain Interference Instruments	HealthMeasures, 2017		Self-administered	

NOTE: α = Cronbach's alpha; CI = confidence interval; ICC = intraclass correlation coefficient.

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
Minutes	Good reliability but "questionable" validity as a measure of disability for individuals with chronic musculoskeletal pain (Boonstra et al., 2008). VAS has been shown to be reliable and valid in an acute setting, such as a hospital emergency department (Bijuer et al., 2001).	No		Measurement of pain can be via a horizontal or vertical scale from 0 to 10 (0 = no pain; 10 = severe pain). Individuals with cognitive limitations or those with fine motor deficits may have trouble completing the assessment (Hawker et al., 2011; Williamson and Hoggart, 2005).
Less than 3 minutes	Instrument has been shown to be valid and reliable (Hawker et al., 2011; Williamson and Hoggart, 2005).	No	Adults	Can be administered verbally as well as in writing.
	Instrument has been shown to be valid and reliable (Amtmann et al., 2010).	No		Measure self-reported effects of pain on relevant aspects of one's life (HealthMeasures, 2017).

ANNEX TABLE 5-5
Selected Cardiovascular Assessments

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Cardio-vascular System	Minnesota Living with Heart Failure Questionnaire	Rector et al., 2006	None except calculation of score	Self-administered	
	Kansas City Cardiomyopathy Questionnaire	Joseph et al., 2013	None except calculation of score	Self-administered	
	Seattle Questionnaire for Angina	Spertus et al., 1995	None except calculation of score	Self-administered	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
	5–8 minutes	Good reliability, responsiveness, performance across populations; feasibility; and interpretability (Kelkar et al., 2016).	Yes		Measures the overall impact of a decrease in heart function on the life of an individual; 21 items for which the responses are summed to a 105-point score, where a higher score indicates more limitations.
	5–8 minutes	Good reliability, responsiveness, performance across populations; feasibility; and interpretability (Kelkar et al., 2016).	Yes		Measures the overall impact of a decrease in heart function on the life of an individual; 23 items and more specific description of limitation on specific activities such as walking, climbing stairs, housework, and yardwork. A difference of 5 points between two time points has been validated as a meaningful change in quality of life.
	5 minutes	Reliability in small-sized populations tested.	Yes		Developed using a framework similar to that for the Kansas City Questionnaire for cardiomyopathy/heart failure; recently adapted and validated as shorter versions consisting of 7 items (Chan et al., 2014).

continued

ANNEX TABLE 5-5
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Cardio-vascular System (*continued*)	New York Heart Association Classification	Albert et al., 2010; Joseph et al., 2013; NYHA, 1994; Raphael et al., 2007	Physician assessment during routine clinical care		
	Exercise tolerance test with treadmill or bicycle exercise (occasionally with hand ergometer) with monitoring of electrocardiogram and blood pressure	Ellestad et al., 1969	Licensed exercise physiologist, nurse, or physician	In an exercise testing laboratory with available equipment for monitoring and resuscitation	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
		Reproducibility across clinicians is limited, in part because it changes in response to fluctuations in clinical status from day to day.	No	Populations of adult patients with heart failure or angina; applicable to and is often used to describe the degree of limitation with other cardiac disorders, such as adult congenital heart disease and acquired valve disease	Robust predictor of outcomes with heart failure, as measured by rates of hospitalization, heart transplant or mechanical circulatory support, or death.
	30–45 minutes	Good responsiveness, reproducibility, and correlation with physical capacity.	No	Populations of adults with multiple etiologies of heart disease	Objective measurements of exercise level allow estimates of peak capacity for specific work and recreational activities. Markedly influenced by level of patient effort.

continued

ANNEX TABLE 5-5
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Cardio-vascular System (*continued*)	Cardiopulmonary exercise tolerance test with gas exchange measurement during treadmill or bicycle exercise (occasionally with hand ergometer)	ATS/ACCP, 2003	Licensed exercise physiologist or other licensed provider; additional expertise needed beyond that for other exercise testing	Equipment for usual exercise testing with additional equipment needed for metabolic equivalents	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
	40–60 minutes	Most responsive and reproducible measurement of integrated heart and lung reserve and high correlation with physical capacity.	No	Populations of adults with multiple etiologies of heart disease	Level of effort can be objectively assessed from measured gas exchange. Provides the most accurate measurement of peak physical work capacity in the exact units that have been measured for specific work and recreational activities. Estimation can be made of the lower level of activity that could likely be sustained for 50 minutes.

ANNEX TABLE 5-6
Selected Visual Assessments

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered
Visual Acuity	Snellen Chart	Chen et al., 2012; Kaiser, 2009	Ophthalmologist, optometrist	Chart, read one eye at a time, at a distance of 10 or 20 feet
	Bailey-Lovie Chart	Bailey and Lovie, 1976; Kaiser, 2009; NRC, 2002	Ophthalmologist, optometrist	Chart, read one eye at a time
	Early Treatment Diabetic Retinopathy Study visual activity chart	Bailey and Lovie, 1976; Chen et al., 2012; Kaiser, 2009; NRC, 2002	Ophthalmologist, optometrist	Chart, read one eye at a time
Peripheral Vision	Kinetic perimetry	Grobbel et al., 2016	Perimetrist	
	Semiautomated kinetic perimetry	Grobbel et al., 2016	Perimetrist	Allows computer-controlled standardized presentation for any chosen Goldmann stimulus size-intensity combination

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
Quick and easy to administer				Disadvantage: each line has a variable letter size and variable letters per line; not all presented characters are equally legible.
Quick and easy to administer	Improved test-retest reliability.		Used for purposes of determining eligibility for driving and legal blindness	Each row has the same number of letters with between-letter spacing equal to 1 letter width; letter size follows a geometric progression.
Prolonged testing time compared with Snellen Chart	Improved test-retest reliability.		Used for purposes of determining eligibility for driving and legal blindness	Equal number of characters per row; equal logarithmic decrement between successive rows; use of character types with relatively uniform legibility.
Less tedious and time-consuming compared with conventional automated static perimetry			Used for qualification for driver's license, assessment of disability, qualification for special support programs for the visually impaired	Advantages: sensitive for detecting peripheral visual field defects; efficient for detection and monitoring progression of steep-edged field defects; kinetic examination results correlate with activities of daily living.
	Greater consistency, reliability, standardization, analysis and interpretation of visual field results.		Individuals with visual field loss	Advantage: can assess the reaction time for a subject for each visual field session.

ANNEX TABLE 5-7
Selected Hearing Assessments

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered
Hearing	Pure-tone audiometric screening	McBride et al., 1994; Yueh et al., 2003	Trained medical assistant, nurse, or physician	In a quiet office setting using a handheld or portable device
	Otoacoustic emissions (OAEs)	Wang et al., 2002	Trained medical assistant, nurse, or physician	In a quiet office setting using portable equipment
	Hearing localization tasks	Abel et al., 2009; Przewozny, 2016	Audiologist or scientist with expertise in spatial hearing	Using multiple sound sources in an acoustically treated environment
	Diagnostic pure-tone audiometry	ASHA, 2005; Roeser et al., 2000	Audiologist	Using specialized equipment in a sound-treated test booth
	Extended Speech Intelligibility Index (ESII)	Soli et al., 2018a,b	Audiologist	Using software to estimate the impact of an audiogram in real-world listening conditions
	Speech recognition in noise testing	Giguère et al., 2008; Laroche et al., 2003	Audiologist	Using recorded speech materials delivered at a calibrated level in an audiometric test booth
	Live-voice speech testing	MacPhee et al., 1988	Anyone	A person says words or phrases that the listener must repeat back
	Internet- and telephone-based screening	Smits et al., 2004; Watson et al., 2012	None	Speech in noise is presented either over the phone or through an Internet-based application

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
3 minutes per ear	Sensitivity: 0.8–0.96; specificity: 0.77–0.89.	No	Older children, adults, elderly	
3–5 minutes per ear	Sensitivity: 0.91–0.98; specificity: 0.62–0.86.	No	Any age	
30–45 minutes	Test-rest reliability of 0.8, with differences of 1–3 degrees.	No	Adults	Not widely implemented because of technical challenges.
20 minutes	Test-retest reliability of 0.9, with differences of 3–5 dB.	No	Any age	
5 minutes	Test-retest reliability ranging from 0.78 to 0.97.	No	Any age	Software for calculating ESII is not currently available.
15 minutes	Test-retest reliability of 0.85 or higher for recorded materials.	Yes, depending on the stimuli	Any age	
1–2 minutes	Sensitivity of 1 for hearing loss and specificity of 0.84.	No	Any age	Numerous limitations related to lack of calibration and standardization across talkers.
5–10 minutes	Sensitivity of 0.80 for hearing loss and specificity of 0.84.	Yes	Adults	

continued

ANNEX TABLE 5-7
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Hearing (*continued*)	Hearing questionnaires: Hearing Handicap Inventory for Adults (HHIA) Speech, Spatial and Qualities of Hearing Scale (SSQ)	 Newman et al., 1990 Gatehouse and Noble, 2004	None	Can be administered as an interview or self-administered by the listener	

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
10 minutes	HHIA—α = 0.93; test-retest reliability of 0.97. SSQ—test-retest reliability of 0.84 for interview and 0.72 for self-administration.	No	Adults	

ANNEX TABLE 5-8
Selected Speech/Fluency and Voice Assessments

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Speech and Fluency Assessments					
Intelligibility	Speech/ Phoneme Intelligibility Test (SIT) Assessment of Intelligibility in Dysarthric Speech (AIDS)	Yorkston et al., 2007 Yorkston et al., 1984	MA, MS, SLPD, PhD; CCC, SLP	Record utterances; transcribe and calculate intelligibility, communication efficiency	
	Phoneme Intelligibility Test	Yorkston et al., 2007	MA, MS, SLPD, PhD; CCC, SLP	Record utterances; transcribe and calculate intelligibility in words	
Motor Speech Disorders	Frenchay Dysarthria Assessment, 2nd Edition	Enderby and Palmer, 2008	MA, MS, SLPD, PhD; CCC, SLP	Rate performance on speech tasks, structures, and functions	
	Apraxia Battery for Adults, 2nd Edition	Dabul, 2000	MA, MS, SLPD, PhD; CCC, SLP	6 subtests with varying levels of speech production complexity	
	Apraxia of Speech Rating Scale	Strand et al., 2014	MA, MS, SLPD, PhD; CCC, SLP	16 items; scored 0–4 for characteristics, prominence, severity on speech production tasks	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
	30 minutes		Madonna Rehabilitation Institute, Lincoln, Nebraska ProEd	Literate, English	SIT is computer based. AIDS is paper based.
	30 minutes			Literate, English	
	20 minutes		ProEd	Ages 12–97, English	
	20 minutes	Reliability: α = 0.83–0.99; validity: content, criterion, construct.		Adolescents, adults, English	
	30 minutes	Reliability: ICC = 0.94; validity: specificity 100%, sensitivity 75–96%.	*Journal of Communication Disorders*	Adults, English	

continued

ANNEX TABLE 5-8
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered
Functional Communication	Communication Effectiveness Checklist		PRO	
	Communication Participation Item Bank	Baylor et al., 2013	PRO	
	Levels of Speech Usage	Baylor et al., 2008	PRO	5-category scale rates type, frequency, amount, and perceived importance of daily speaking situations
	Fatigue Severity Scale	Krupp et al., 1989	PRO	9 items; 7-point Likert scale
	Quality of Communication Life Scale	Paul et al., 2004	MA, MS, SLPD, PhD; CCC, SLP	Psychosocial, vocational, educational effects of communication impairment
	ASHA Functional Assessment of Communication Skills for Adults	Frattali et al., 2017	MA, MS, SLPD, PhD; CCC, SLP	43 items

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
			English	
			English	
			English	
5 minutes	Valid, reliable.	*Archives of Neurology*, open access	Adults; multiple sclerosis, traumatic brain injury (TBI), Parkinson's disease; variety of populations	Several language versions validated.
15 minutes	Validated.	American Speech-Language-Hearing Association (ASHA)	Adults with neurogenic communication disorders, English	
	Validated.	ASHA	Adults with stroke, TBI; English	Downloadable Excel spreadsheet to calculate scores.

continued

ANNEX TABLE 5-8
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Fluency	Behavior Assessment Battery for Adults who Stutter	Brutten and Vanryckeghem, 2003, 2007	PRO	Individual indicates strength of various negative emotional reactions to situations	
	Overall Assessment of the Speaker's Experience of Stuttering-Adult	Yaruss, 2010; Yaruss and Quesal, 2006	MA, MS, SLPD, PhD; CCC, SLP; PRO	Individual responds to each question on 5-point Likert scale	
	Stuttering Severity Instrument, 4th Edition	Riley and Bakker, 2009	MA, MS, SLPD, PhD; CCC, SLP	Collect speech using pictures and readings provided; score frequency and severity	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
		Validity: construct, content; reliability: internal (α = 0.96–0.98).	Plural Publishing, Inc.	Adults who stutter	
	20 minutes	Validated; criterion-referenced.	Stuttering Therapy Resources, Inc.	Ages 18+, English	Electronic and paper versions; also validated in Dutch, German, Hebrew, Portuguese, Spanish; based on *International Classification of Functioning, Disability and Health*.
	20 minutes	Descriptive; norm-referenced (N = 60 adults).	ProEd	Ages 2–adult	Has a computerized scoring option.

continued

ANNEX TABLE 5-8
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Voice Assessments					
Patient-Reported Assessment/ Outcomes (PRO)	Voice Outcome Survey	Gliklich et al., 1999	PRO	5-point Likert scale	
	Voice-Related Quality of Life	Hogikyan and Sethuraman, 1999	PRO	5-point Likert scale	
	Communication Participation Item Bank	Baylor et al., 2013	PRO	5-point Likert scale	
	Linear Analog Scale Assessment of Voice Quality	Llewellyn-Thomas et al., 1984	PRO	Visual analog scale	
	Voice Handicap Index 10	Arffa et al., 2012; Rosen et al., 2004	PRO	5-point Likert scale	
Auditory-Perceptual	Consensus Auditory-Perceptual Evaluation of Voice	ASHA DIV3, n.d.; Zraick et al., 2011	MA, MS, PhD; CCC, SLP	Visual analog scales, measure mm (0–100)	
	Grade, Roughness, Breathiness, Asthenia, Strain (GRBAS) scale	Carding et al., 2009; Hirano, 1981	MA, MS, PhD; CCC, SLP	4-point grading scale (0 = normal/absent deviance; 1 = slight deviance; 2 = moderate deviance; 3 = severe deviance)	

	Time to Administer	Psychometric Properties (reliability = $r \geq 0.70$, validity)	Proprietary	Population to Which Applies	Comments
	5 items	Validity: content, construct, longitudinal, convergent; reliability +.	No	5th-grade literacy, English, unilateral vocal fold paralysis	
	10 items	Validity: content, longitudinal, convergent; reliability +.	No	English, voice disorders	
	48 items, long version; 10 items, short version	Validity: content, construct, convergent; reliability +.	No	English, communication disorders	
	16 items	Validity: content, longitudinal, convergent; reliability +.	No	English, laryngeal cancer	
	10 items	Validity: construct, convergent; reliability +.	No	5th-grade literacy, English, voice disorders	
	Speech tasks: vowels, sentences, running speech; 6 items completed after person performs voice tasks	Construct validity, reliability +.	ASHA, available at http://www.asha.org	Voice disorders	
	Assess current conversational speech or reading aloud	Valid and reliable.			

continued

ANNEX TABLE 5-8
Continued

Physical Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Visual-Perceptual	Laryngoscopy, endoscopy, stroboscopy	Poburka, 1999	MD; MA, MS, PhD; CCC, SLP	Examine structures at rest and during speech	
Physiological	Acoustic analyses; Multidimensional Voice Program (MDVP)		MA, MS, PhD; CCC, SLP		
	Aerodynamic analyses				

NOTE: α = Cronbach's alpha; CCC, SLP = certificate of clinical competence, speech-language pathologist; ICC = intraclass correlation coefficient; MA, MS, SLPD, PhD = master of arts, master of science, doctor of speech-language pathology, doctor of philosophy; PRO = patient-reported outcome measure.

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
30 minutes				View structures.
		Kay Elemetrics		Habitual pitch, loudness, ranges.
				Average pressure, airflow.

ANNEX TABLE 5-9
Selected Language Assessments

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Language and Communication	Boston Diagnostic Aphasia Examination-Third Edition (BDAE-3) Component: Boston Naming Test, Second Edition (BNT-2)	Goodglass et al., 2000	MA, MS, PhD; CCC, SLP	Long version: every test in the inventory; short version: selected items from test inventory; extended testing probes language within each area; assessment of aphasia and related disorders includes aphasia definition and characteristics, normative basis, administration and interpretation instructions, severity rating, visuospatial and quantitative skills after brain injury	
	Western Aphasia Battery-Revised	Kertesz, 2007	MA, MS, PhD; CCC, SLP	Individually administered subtests	
	Arizona Battery for Communication Disorders of Dementia	Bayles et al., 1989, 1993	MA, MS, PhD; CCC, SLP	17 subtests (speech discrimination, visual perception, literacy)	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
	Long version: 60–120 minutes; short version: 30–45 minutes	Reliability: 0.68–0.98; demonstrates good internal consistency; validity: discriminant analysis yielded no misclassification; normative data: 242 people with aphasia.	https://www.linguisystems.com (accessed June 21, 2019)	Adult, aphasia, brain injury	
	Bedside screen: 15 minutes; full: 45 minutes, 8 tests, 105 minutes, 8+ supplementary	Reliability: test-retest, inter- and intrajudge; validity: face, content, construct.	Kertesz, 1982, 2007	Adult, aphasia, brain injury	Even when considered nonaphasic, has ongoing discourse impairments; high scores correlate with Communicative Effectiveness Index functional communication; evidence of racial/ethnic differences on selected subtests.
	45–90 minutes	Validity: criterion, construct.		Adult (15+ yrs), Alzheimer's, Parkinson's disease, traumatic brain injury (TBI)	

continued

ANNEX TABLE 5-9
Continued

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Language and Communication (*continued*)	Porch Index of Communicative Ability-Revised	Porch, 2001	MA, MS, PhD; CCC, SLP certified in scoring system	Multidimensional scoring designed to measure level of performance; 18 subtests	
	Psycho-linguistic Assessments of Language Processing in Aphasia (PALPA)	Kay et al., 1992		60 subtests (auditory processing, reading, spelling, naming, comprehension)	
	National Institutes of Health (NIH) Stroke Scale	NINDS, 2011	Certified health care provider (participate in online training)	15 items (consciousness, language, neglect, visual field, extraocular movement, motor strength, ataxia, dysarthria, sensory)	
	Comprehensive Aphasia Test (CAT)	Swinburn et al., 2004	MA, MS, PhD; CCC, SLP	34 subtests (language comprehension, repetition, spoken language, reading, writing)	
	Peabody Picture Vocabulary Test, 5th Edition	Dunn, 2018	MA, MS, PhD; CCC, SLP	Select image from 4 displayed	
	Expressive Vocabulary Test, 3rd Edition	Williams, 2018	MA, MS, PhD; CCC, SLP		

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
		http://www.picaprograms.com/pica-test-materials-and-pricing.html (accessed April 4, 2019)	Adult, aphasia	Evidence that scoring is not of equal intervals.
90–120 minutes	None			
<10 minutes	Reliability: interrater reliability = 0.62; validity: content, responsiveness.	NINDS, 2011	Stroke, aphasia	Dutch, English, Italian, Portuguese, Spanish.
90–120 minutes	Validity: none; reliability: interrater (0.90).		Adult aphasia	
10–15 minutes	Validity; reliability.	Pearson	Ages 26–90+	Receptive vocabulary; correlates with cognition; online scoring system available.
15–20 minutes	Validity; reliability.	Pearson		Expressive vocabulary; online scoring available.

continued

ANNEX TABLE 5-9
Continued

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Language and Communication (*continued*)	Cognitive Linguistic Quick Test-Plus	Helm-Estabrooks, 2017	MA, MS, PhD; CCC, SLP	5 communication domains (attention, memory, executive functions, language, visuospatial skills)	
	Communicative Activities of Daily Living-2	Holland et al., 1999	MA, MS, PhD; CCC, SLP	50 items (social interaction, communication, sequential relationships, humor/ metaphor/ absurdity)	
	Pragmatic Protocol	Prutting and Kirchner, 1987	MA, MS, PhD; CCC, SLP	30 communication abilities observed	
	Assessment of Pragmatic Abilities and Cognitive Substrates (APACS)	Arcara and Bambini, 2016	MA, MS, PhD; CCC, SLP	2 domains: discourse and nonliteral language	
	Communication Effectiveness Inventory	Lomas et al., 1989	PRO	16 communication situations	
	La Trobe Communication Questionnaire	Douglas et al., 2000	2 forms: PRO, clinician	30-item Likert-type ratings	
	Communication Checklist-Adult	Whitehouse and Bishop, 2009	Well-known informant	70 items	

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
15–30 minutes	One pilot (N = 13) and three studies (N = 92, 154, 119, respectively) established reliability and validity; criterion-referenced; descriptive severity.	Pearson	Adult ages 18–90, aphasia, brain injury	English and Spanish versions.
30 minutes	Reliability: coefficient (0.93), test-retest (0.85), interrater (0.99); validity: 0.66.		Adult	
15 minutes	Reliability: interrater (0.90).			
35–45 minutes	Reliability: test-retest; internal consistency (α = 0.60–0.70); validity: construct, content.	Download	Adults ages 19–89	
10 minutes	Internal consistency (α = 0.90), test-retest (ICC = 0.94), interrater (ICC = 0.73); validity: construct.		Adults, literate, aphasia	
PRO: 30 minutes; informant: 15 minutes	Internal consistency (α = 0.8596); reliability: test-retest (r = 0.7558).		Adolescents ages 13–17, adults ages 18–64; TBI	See Donovan et al., 2008.
		Pearson	Adults, developmental and acquired	

continued

ANNEX TABLE 5-9
Continued

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Language and Communication (*continued*)	Stroke and Aphasia Quality of Life Scale-39	Hilari et al., 2003	PRO	39 items	
	Assessment of Living with Aphasia	Simmons-Mackie et al., 2014			
	Aphasia Communication Outcome Measure	Hula et al., 2015	PRO	59 items; domains: talking, comprehension, writing	
	American Speech-Language-Hearing Association Functional Assessment of Communication Skills	Frattali et al., 1995, 2017	MA, MS, PhD; CCC, SLP; PRO	43 items, 4 domains (social, basic need, written language, and daily planning)	
	ASHA Quality of Communication Life Scale	Paul et al., 2004	MA, MS, PhD; CCC, SLP; PRO	ICF participation focus	

NOTE: α = Cronbach's alpha; CCC, SLP = certificate of clinical competence, speech-language pathologist; ICC = intraclass correlation coefficient; MA, MS, PhD = master of arts, master of science, doctor of philosophy.

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which Applies	Comments
	Reliability: internal consistency (α = 0.74–0.94), test-retest (ICC = 0.89–0.98); validity: acceptability, construct (*r* = 0.38–0.58; convergent, *r* = 0.55–0.67; discriminant, *r* = 0.02–0.27).		Adults with stroke, aphasia	
	Internal consistency reliability coefficient >0.80.			
	Communicative functioning in aphasia (validity: factor analyses supported a coherent measurement model, items functioned similarly across demographic and clinical subgroups, and scores showed convergence with related constructs).	JSLHR, 2015	Adult, aphasia	
20 minutes	Reliability: interrater (0.88–0.95); validity: external (0.73).	Rockville, MD; ASHA	Adult, U.S. American	A software application available from the first author may be used to administer and score the assessment.
15 minutes	Reliability; validity.	Rockville, MD; ASHA	Adult, neurogenic communication impairments	

6

Selected Instruments for Assessment of Mental Functional Abilities Relevant to Work Requirements

This chapter reviews instruments available for measuring mental abilities relevant to work requirements. The framework described in Chapter 2 (see Figure 2-3) provides a way to organize specific types of instruments for assessing mental function and helping determine an individual's ability to perform specific work tasks. For example, computer-assisted testing can be used to assess mental abilities and provide information on an individual's ability to complete work tasks. However, it is important to ensure that the assessment considers task sequencing and coordination, and the ability to do sustained work on a regular and continuing basis. The chapter also addresses functional assessment in people who have mental impairments.

The U.S. Social Security Administration's (SSA's) adult *Listing of Impairments* for mental disorders includes neurocognitive disorders such as dementia; schizophrenia spectrum and other psychotic disorders; depressive, bipolar, and related disorders; intellectual disorders; anxiety and obsessive-compulsive disorders; somatic symptom and related disorders; personality and impulse-control disorders; autism spectrum disorder; neurodevelopmental disorders; eating disorders; and trauma- and stressor-related disorders.[1] Community functioning and mental illness are closely associated in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5) definition of mental illness, in which "clinically significant distress or disability" is a key diagnostic criterion for all mental disorders. The definition of what are probably considered the most severe of the major mental illnesses, schizophrenia spectrum disorders, incorporates impairment and

[1]This text has been revised since prepublication release.

its persistence in psychosocial functioning as part of establishing the diagnosis. In addition to the diagnosis requiring functional limitations, diagnostic symptoms—most notably disorganized speech or behavior and negative symptoms, such as diminished concentration, persistence, and pace; motivation; and goal-directed behavior—also have implications for functioning. And while the majority of individuals with a mental illness diagnosis do not experience significant functional impairment, most psychiatric symptoms, if severe and sufficiently persistent, can cause functional and occupational impairment. The symptoms of depression, for example—such as slowed thought and reduced physical movement that must be observable by others; diminished ability to think, concentrate, or make decisions; and fatigue or loss of energy—have obvious implications for functioning at work and socially and for independent self-care.

An additional important facet of the relationship between mental illness and functioning is that some major mental illnesses are episodic in nature, with the severity of symptoms and functional impairments varying over time, and with periods of greater severity ranging from weeks to months. This variability is particularly important in considering whether an individual can perform substantial gainful activity. For most people with schizophrenia spectrum disorders, for example, acute periods may require temporary hospitalization to prevent harm to self or others, but importantly, even in periods of greater symptom stability, the degree of psychosocial impairment can be significant. Many people do not return to premorbid levels of psychosocial functioning either ever or until many years later in the course of the illness. Mood disorders, such as major depression, are also episodic; in contrast to schizophrenia, however, a return to unimpaired functioning between episodes is possible, including for some people who function at high levels and make significant work contributions, while others display interepisode residual functional impairment. The key point is that it is important to assess the persistence of impairment due to mental disorders associated with episodic or persistent symptoms.

MENTAL FUNCTIONAL ABILITIES RELEVANT TO WORK REQUIREMENTS

Annex Table 6-1 links mental functional domains identified by the committee (defined in Annex Table 6-2) to specific work demands. The committee identified the following mental functional domains: general cognitive/intellectual ability, language and communication, learning and memory, attention and vigilance, processing speed, executive functioning, adaptability, and work-related personal interactions.[2] These domains are adapted from the *Report of the Mental Cognitive Subcommittee of the Occupational Information Development Advisory Panel* (OIDAP, 2009), *Psychological Testing in the Service of Disability Determination* (IOM, 2015), and *Informing Social Security's Process for Financial Capability Determination* (NASEM, 2016) and from the Paragraph B criteria in SSA's adult *Listing of Impairments* for mental disorders (SSA, n.d.-a). Annex Table 6-1 links these eight mental functional domains to mental abilities listed in three SSA forms used to collect functional information when making a disability determination: the Psychiatric Review Technique Form (PRTF), the Function Report, and the Mental Residual Functional Capacity Assessment (MRFC). The functional domains identified by the committee are also linked to SSA's Paragraph B criteria, to mental abilities in the Occupational Requirements Survey (ORS), and to work activities relevant to mental processes in Occupational Information Network (O*NET).

The functional domain of *language and communication*, for example, refers to "receptive and expressive language abilities" and "how well a person can understand spoken or written language, communicate his or her thoughts, and follow directions" (OIDAP, 2009, pp. C21 and C23).

In the PRTF, the language and communication domain links to *language* and *interacting with others*; the PRTF lists *language* under neurocognitive disorders.[3] In the function report, this domain links to *getting along with others, understanding,* and *completing tasks* (SSA, 2015); the Function Report allows for self- and third-party reports. Links to this domain in the MRFC form include both *social interaction* and *understanding and memory* (SSA, n.d.-b). If the medical evidence provided shows that an individual has the ability to ask simple questions or request assistance, the claimant may not demonstrate a limitation under the category of *social interaction* (SSA, n.d.-b). For *understanding and memory*, the MRFC defines the ability to understand and remember through assessing short and simple instructions (SSA, n.d.-b). The ORS's cognitive demand element most relevant to *language and communication* is work-related personal interactions, defined as "the requirement of the job to cooperate with others, handle conflict, and respond to social cues, requests, and criticism" (DOL, 2017, p. 64). O*NET's cognitive abilities are "abilities that influence the acquisition and application of knowledge in

[2]The domains of adaptability and work-related personal interactions are included in the cognitive demand elements in the July 2017 version of the Occupation Requirements Survey (ORS) Collection Manual. The cognitive elements were updated in an August 2018 version of the ORS Collection Manual, which became available following the committee's work on this chapter. Annex Table 6-1 includes relevant cognitive elements from the 2018 ORS Collection Manual as well as those from the 2017 manual.

[3]Form SSA-2506-BK (01-2017) UF, obtained via personal communication with Joanna Firmin, U.S. Social Security Administration, February 23, 2018.

problem solving" (O*NET, 2018a). *Oral comprehension, oral expression, written comprehension,* and *written expression* link to the *language and communication* domain. The relevant O*NET work activity that links to this domain is *processing information*, described as compiling, coding, categorizing, calculating, tabulating, auditing, or verifying information or data (O*NET, 2018b). Thus, Annex Table 6-1 shows that a limitation in the functional domain of language and communication can limit specific job demands related to interacting with others, processing information, and completing tasks.

Finally, SSA uses the *Listing of Impairments* at step 3 of its process for determining whether a claimant qualifies for benefits (see Chapter 2). The claimant's mental disorder must satisfy requirements listed in both paragraphs A and B (or C for listings with a paragraph C).[4] Paragraph A includes the medical criteria that must be present in a claimant's medical evidence. Paragraph B provides information on the functional criteria assessed to determine how a mental disorder limits functioning. Four criteria represent the areas of mental functioning a person uses in a work setting: *understand, remember, or apply information; interact with others; concentrate, persist, or maintain pace;* and *adapt or manage oneself*. The *language and communication* functional domain links to the criteria *understand, remember, or apply information*, which refer to the "abilities to learn, recall, and use information to perform work activities" (SSA, n.d.-a). Examples include understanding and learning terms, instructions, and procedures; following one- or two-step oral instructions to carry out a task; describing work activity to someone else; asking and answering questions and providing explanations; recognizing a mistake and correcting it; identifying and solving problems; sequencing multistep activities; and using reason and judgment to make work-related decisions. SSA does not require documentation of all of these examples.

With respect to the other functional criteria in the *Listing of Impairments, interact with others* refers to the "abilities to relate to and work with supervisors, co-workers, and the public" (SSA, n.d.-a). Examples listed include cooperating with others; asking for help when needed; handling conflicts with others; stating own point of view; initiating or sustaining conversation; understanding and responding to social cues (physical, verbal, emotional); responding to requests, suggestions, criticism, correction, and challenges; and keeping social interactions free of excessive irritability, sensitivity, argumentativeness, or suspiciousness. Again, SSA does not require documentation of all of these examples.

Concentrate, persist, or maintain pace refers to the "abilities to focus attention on work activities and stay on task at a sustained rate" (SSA, n.d.-a). Examples include initiating and performing a task that you understand and know how to do, working at an appropriate and consistent pace,

[4]This text has been revised since prepublication release.

completing tasks in a timely manner, ignoring or avoiding distractions while working, changing activities or work settings without being disruptive, working close to or with others without interrupting or distracting them, sustaining an ordinary routine and regular attendance at work, and working a full day without needing more than the allotted number or length of rest periods during the day. As with other Paragraph B criteria, SSA does not require documentation of all of these examples.

Adapt or manage oneself refers to the "abilities to regulate emotions, control behavior, and maintain well-being in a work setting" (SSA, n.d.-a). Examples listed include responding to demands, adapting to changes, managing psychologically based symptoms, distinguishing between acceptable and unacceptable work performance, setting realistic goals, making plans for oneself independently of others, maintaining personal hygiene and attire appropriate to a work setting, and being aware of normal hazards and taking appropriate precautions. Again, SSA does not require documentation for all of these examples.

INSTRUMENTS USED TO ASSESS MENTAL FUNCTIONAL ABILITIES RELEVANT TO WORK REQUIREMENTS

The committee applied the following criteria in deciding which instruments to describe in this chapter: (1) sufficient representation in the scientific literature and/or widespread use; (2) evidence of sound psychometric properties, including (when applicable) construct validity, internal consistency, sensitivity to change, test-retest reliability, intra-/interrater agreement (including subject/proxy and telephone/in-person administration); (3) normative data; (4) applicability across a range of conditions and functional levels; (5) availability in the public domain; (6) ease of administration; (7) brevity; (8) availability in multiple languages; (9) validation in subpopulations; (10) multiple administration formats (telephone interview versus in-person administration; self- versus proxy respondent); and (11) availability of alternative forms to minimize the risk of practice effects for performance measures. Some of the instruments discussed do not fulfill all of these criteria, but they are included because they illustrate the range of potential assessment instruments. Discussion of the instruments turns first to general assessment tools (see Annex Table 6-3), then to neuropsychological testing used to assess the mental functional domains described above (see Annex Table 6-4), and finally to measures of disorder severity and work-related functional impairment (see Annex Table 6-5). Annex Tables 6-3, 6-4, and 6-5 provide information on selected functional assessment tools for mental abilities, including qualifications to administer, how to administer, time to administer, psychometric properties, proprietary considerations, and the populations to which they apply.

General Assessment Tools

The instruments described below are used to assess mental function across multiple domains.

Work Disability Functional Assessment Battery (WD-FAB) Mental Health Measures

To support its disability determination process, SSA funded a contract to the National Institutes of Health (NIH) to develop the WD-FAB, a claimant-reported measure of mental health. Marfeo and colleagues (2018) developed four mental health measures assessing cognition and communication (68 items), self-regulation (34 items), resilience and sociability (29 items), and mood and emotions (34 items). To develop these measures, the authors collected data from a random, stratified sample of 1,695 SSA claimants and a general-population sample of 2,025 working-age adults (Marfeo et al., 2018). To expand the WD-FAB scales, 169 new items were developed, and responses were analyzed using factor analysis and item response theory (IRT) analysis to construct unidimensional scales (Marfeo et al., 2018). In addition, computer adaptive testing (CAT) simulations were conducted to examine the instrument's psychometric properties. Results of confirmatory factor analysis revealed acceptable fit statistics across all mental health subdomains in both samples and for all scales (root mean square error of approximation ≤ 0.08, comparative fit index and Tucker-Lewis index ≥ 0.9) (Marfeo et al., 2018). Correlations between the CAT simulations and the full item bank exceeded 0.95. Differential item functioning related to age, sex, and race was minimal in both samples. The authors concluded that all four scales displayed acceptable psychometric properties. Further, results of a recent study of the psychometric properties of the WD-FAB demonstrated reliability and construct validity in a large group of working-age adults ($N = 335$), as well as in adults unable to work because of permanent physical ($N = 375$) or mental ($N = 296$) disability (Meterko et al., 2018). The WD-FAB is unique in its use of IRT and CAT.

World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0)

Validity and reliability have been demonstrated for WHODAS 2.0[5] through extensive field testing in international, multicenter studies, and it has demonstrated robust factor structure for general disability and specific life domains (WHO, 2010). An IRT-based scoring method is used that

[5]Refer to Chapter 4 for a detailed description of WHODAS 2.0.

accounts for item difficulty and employs an algorithm (available from WHO) to determine a summary score that is converted into a metric rating from 0 (no disability) to 100 (full disability). Its advantages include public availability, brevity (20-minute administration), a simple scoring algorithm (items scored as 0 [none], 1 [mild], 2 [moderate], and 3 [extreme], with scores summed across items); multiple administration modalities (interviewer, self-, or proxy report), and telephone administration by an interviewer with basic skills. A 12-item, 5-minute version provides a brief assessment of global functioning that explains 81 percent of the variance of the 36-item version (WHO, 2010).

University of California, San Diego, Performance-Based Skills Assessment (UPSA)

Ratings of functional capacity are of increasing interest in studies of people with mental illness because they are highly correlated with cognition and some aspects of community functioning. In particular, the UPSA has shown some promise in being able to predict real-world functioning in middle-aged and older adults with schizophrenia (Mausbach et al., 2011) and mood disorders (Bowie et al., 2006; Mausbach et al., 2010). Specifically, the UPSA has demonstrated high correlations with measures of personal care skills, interpersonal skills, and community activities (Mausbach et al., 2010). It has high interrater and cross-temporal reliability, as well as demonstrated validity in assessing functional skills in healthy adults and the elderly. Disadvantages are that it is not strongly associated with employment and does not serve as a predictive factor for employment (Mausbach et al., 2011). Furthermore, it is not publicly available; it requires training to administer; and it is not validated in neurological disorders, including stroke and dementia.

Occupational Functioning Scale (OFS)

The OFS is an observer rating scale of work ability in people with mental health disorders. Its validity was established by comparison with other work ability measures (e.g., SAS-Work, Work Ability Index, sickness absence) and other measures not related to work ability, such as the SCL-90-GSI and Inventory of Interpersonal Problems (Hannula et al., 2006). Acceptable interrater reliability was demonstrated (intraclass correlation coefficient [ICC] = 0.91), as was validity, with the strongest relationships found with other measures of work compared with symptoms or interpersonal problems (Hannula et al., 2006).

Global Assessment of Functioning (GAF) Scale

The GAF is a scoring system for the severity of illness. The rating on the scale is derived from a clinician's judgment of a person's ability to function in daily life based on a composite of psychological symptoms and social and occupational functioning. It does not take into account impairment in function caused by physical or environmental limitations. Ratings range from 0 to 100, with the lowest score consistent with the worst area. The GAF has limited reliability and predictive validity because the domains assessed do not vary together. Problems with predictive utility arise from the tendency of clinicians to overweigh symptoms. Because the GAF is an unstandardized, unreliable rating of disability, SSA no longer uses it in the assessment of disability claims, and it is considered only to the extent that it is consistent with other evidence. Prior to DSM-5, the GAF was endorsed as an assessment of functioning and reported on the fifth axis of a multiaxial system. DSM-5 no longer maintains a multiaxial system and eliminated the use of the GAF "for several reasons, among which were its lack of conceptual clarity (i.e., including symptoms, suicide risk, and disabilities in its descriptors) and questionable psychometrics in routine practice" (APA, 2013, p. 16). These limitations motivated efforts to develop alternative ways of measuring social and occupational functioning separate from psychological symptoms, but inclusive of physical impairment.

Social and Occupational Functioning Assessment Scale (SOFAS)

SOFAS is derived from the GAF (DSM Axis V) (Morosini et al., 2000). It focuses on social and occupational functioning without the influence of psychological symptoms. The influence of general medical conditions is considered in SOFAS ratings if impairment is due to the consequences of physical and/or mental health problems; not considered is the lack of opportunity or environmental limitations. The SOFAS describes functioning during the current period or a specified period of time, such as the highest level of functioning in the past 12 months. Its ratings are based on a scale of 100–0, reflecting a range from excellent to grossly impaired functioning: 100 = superior functioning in a range of activities; 60 = moderate difficulty in social and occupational or school functioning (few friends, conflict with peers and co-workers); 50 = serious impairment in social and occupational or school functioning (no friends, unable to keep job); 40 = major impairment in several areas, such as work or school and family relations (e.g., depressed individual avoids friends, neglects family, and is unable to work); and 0 = unable to rate because of inadequate information. As with the GAF, evidence of reliability and validity is limited by the confounding of ratings that include both social and occupational functioning.

Mental Illness Research, Education and Clinical Center (MIRECC) GAF Social and Occupational Functioning Scales

The MIRECC GAF was developed by the U.S. Veterans Health Administration (VHA) to facilitate clinicians' GAF ratings, which are required every 90 days for patients receiving mental health services (Niv et al., 2007). The MIRECC version of the GAF provides separate ratings of occupational, social, and psychological functioning. It is similar to the GAF in providing scores ranging from 100 to 0, with the following clinical ranges: 70–100 = "fully functional" (e.g., works consistently, socially effective with minimal symptoms); 50–69 = "borderline functional" (e.g., misses work frequently, interpersonal conflicts, mild to moderate symptoms such as moderate depression); 20–49 = "dysfunctional" (e.g., some sheltered work, difficulty with coherent conversation, impairment in reality testing); and 10–19 = "dangerous" (e.g., unable to self-care or interact with others; dangerous to self or others and grossly impaired communication) (Niv et al., 2007). A large (N = 398) multisite study, Enhancing Quality of Care in Psychosis, conducted at three VHA sites, demonstrated adequate concurrent and predictive validity for the MIRECC GAF's three subscales, superior to the concurrent and predictive validity of the GAF (Niv et al., 2007). The strongest convergent validity was in occupational scores, which were strongly correlated with employment in the past month and work status.

Specific Level of Functioning Scale (SLOF)

The SLOF, a multidimensional behavioral rating scale developed in the early 1980s (Schneider and Struening, 1983), is designed to measure directly observable functioning and daily living skills. It consists of 43 behavioral items measured on a 5-point Likert scale and six subscales: physical functioning (e.g., vision, hearing), personal care skills (eating, personal hygiene), interpersonal relationships (interacts with others), social acceptability (acts within bounds of social norms), activities of community living (household responsibilities), and work skills (completes assigned tasks), as well as an "other" item addressing areas of functioning not included on the instrument. Scores range from 43 to 215, with lower scores indicating better functioning. The SLOF requires 15–20 minutes to complete.

The SLOF has separate scales for self- and collateral ratings; the collateral rating scale includes a question about how well the reporter knows the person, rated on a scale from 1 to 5. In a study involving 173 outpatients with schizophrenia, the ICC for the combined scales was $r = 0.62$, and for individual scale items, the ICC range was 0.38–0.80. The same study also measured the ICC of the SLOF in 982 inpatients with schizophrenia; the ICC for the combined scales was $r = 0.42$, and for individual scale items, the ICC range was 0.13–0.72 (Schneider and Struening, 1983), indicating better correlations among scale items in outpatients versus inpatients with schizophrenia. In a study of 221 community-dwelling and nursing home patients, the ICC range was 0.74–0.85. The convergent validity of the SLOF was assessed with the UPSA measure of functional capacity and found to be highest in 78 community-dwelling people with schizophrenia for community activities ($r = 0.61$, $p < 0.01$), but also significant for interpersonal skills ($r = 0.34$, $p < 0.01$) and work skills ($r = 0.54$, $p < 0.01$). In addition, the SLOF was significantly correlated with cognitive functioning as measured by tests of problem solving, inhibition, information processing speed, object recognition, attention, and praxis, as follows: interpersonal skills, $r = 0.23$, $p < 0.05$; community activities, $r = 0.50$, $p < 0.01$; and work skills, $r = 0.41$, $p < 0.01$ (Bowie et al., 2006). The SLOF was also found to be significantly correlated with symptom measures, including the Positive and Negative Symptom Scale ($r = 0.5$, $p < 0.0001$) (Cramer et al., 2000). Significant interrater reliability was demonstrated in one study in schizophrenia using self-ratings from 67 patients and ratings from their case managers (overall $r = 0.28$, $p < 0.01$) (Bowie et al., 2007). Some sensitivity to functional change in the context of treatment was demonstrated in 60 people with schizophrenia in residential treatment and enrolled in an Assertive Community Treatment program. Ratings were obtained at baseline and 1 year later: $t = 4.024$; $df = 29$, $p = 0.0004$ (Chandler et al., 1999). In comparison with the "gold standard" Personal and Social Performance scale, the SLOF was found to be valid and reliable in a large sample of Italian people with serious psychiatric disorders (Mucci et al., 2014). In a study of people with schizophrenia and schizoaffective disorder, significant relationships were found between poorer cognition and overestimation of work function, as well as between higher depression levels and underestimation of interpersonal function (Ermel et al., 2017). Overall, the SLOF had significantly stronger correlations with interpersonal and work function compared with the other areas of function.

Neuropsychological Testing[6]

Neuropsychological testing provides valuable information regarding functional capacity in the domain of cognitive functioning. Relevant to SSA's considerations, cognitive functioning includes intellectual capacity,

[6]Neuropsychology is a subspecialty of psychology with defined training. A clinical neuropsychologist is "a professional psychologist trained in the science of brain-behavior relationships [and] specializes in the application of assessment and intervention principles based on the scientific study of human behavior across the lifespan as it relates to normal and abnormal functioning of the central nervous system" (Bieliauskas, 1998, p. 161).

attention and concentration, processing speed, language and communication, visual-spatial abilities, and memory (IOM, 2015). Most tests of cognitive functioning require the test taker to complete timed tasks in a controlled testing environment. Neuropsychologists interpret results relative to population norms and in terms of the test taker's pattern of relative strengths and weaknesses across cognitive domains. Neuropsychological testing allows SSA to evaluate the severity of cognitive impairments and claimants' residual functional capacity.

A wide variety of performance-based neuropsychological tests can be used to assess a claimant's level of cognitive functioning. Numerous performance and symptom validity measures are in use that can assist professionals in interpreting the validity of psychological test results (IOM, 2015). Described below are several commonly used tests within domains of cognitive functioning that are relevant to the mental listings' Paragraph B criteria. (See Lezak et al. [2012] and Strauss et al. [2006] for a comprehensive perspective on performance-based cognitive tests.)

General Cognitive/Intellectual Ability

General cognitive/intellectual ability encompasses reasoning, problem solving, and meeting cognitive demands from basic to high levels of complexity. The most widely used test of cognitive/intellectual functioning is the Wechsler Adult Intelligence Scale, fourth edition (WAIS-IV) (Wechsler, 2008).

Montreal Cognitive Assessment (MoCA) The MoCA is a 30-item screener used to detect cognitive impairment. It assesses orientation (time and place), attention (target detection using tapping, digit span forward), working memory (serial subtraction, digit span backward), verbal short-term memory (two acquisition trials of five nouns with a 5-minute delayed-recall trial), executive functioning (shortened, adapted version of Trail Making Test B), and language (phonemic fluency, confrontation naming, and complex sentence repetition) (Nasreddine et al., 2005). In a validation trial, the MoCA has demonstrated high sensitivity and specificity (approximately 0.9) for detection of mild cognitive impairment[7] (cutoff score of 22; behavioral correlates of mild cognitive impairment include complaints about memory and memory deficits without notable functional impairment) and dementia (cutoff score of 16; behavioral correlates of dementia include notable cognitive and functional impairment) (Nasreddine et al., 2005), as well

[7]Mild cognitive impairment is "an intermediate clinical state between normal cognitive aging and dementia, and it precedes and leads to dementia in many cases" (Nasreddine at al., 2005, p. 695).

as acceptable reliability (internal consistency [Cronbach's alpha = 0.83]). Using these cutoffs, sensitivity and specificity were consistently higher for the MoCA than for the Mini-Mental State Exam (MMSE). The MoCA requires in-person administration because it includes performance-based tasks. Advantages include a 10-minute administration time; public availability; translation to more than 40 languages; a large peer-reviewed literature; and evidence of ability to detect early cognitive changes in a range of neurological disorders, including Parkinson's disease (Hu et al., 2014). Disadvantages include limited validation and a lack of precise cutoff scores for non-English versions.

Short Orientation-Memory-Concentration Test of Cognitive Impairment (OMCT) The OMCT is a 6-item brief version of the 26-item Blessed test, designed to assess the cognitive domains of orientation, attention, and working memory (Katzman et al., 1983). The total score is 28; up to six errors are within normal limits, with scores of 20 and below indicating cognitive impairment. Performance scores discriminate among mild, moderate, and severe cognitive deficits. A discriminant analysis in elderly patients indicated that orientation and easier attention items distinguished those with severe levels of impairment from those with no impairment. The more difficult working memory items differentiated mild from no cognitive impairment, although a subgroup of elderly adults living independently in the community made errors on these items. The sensitivity and specificity of the OMCT are high (about 90 percent) for determining the presence of Alzheimer's disease. A postmortem study demonstrated a positive correlation between scores on the OMCT and plaque counts obtained from the cerebral cortex of 38 subjects (Katzman et al., 1983). The OMCT was equivalent to the MMSE in identifying the presence of dementia in a study comparing unimpaired patients with those with vascular or degenerative dementia. In patients with Alzheimer's disease, performance on the OMCT was equivalent to mean values of a simple reaction time, and was correlated with Wechsler global memory quotient and orientation, logical memory, and paired associate items of the scale. Performance scores were reliable a month later, with no evidence of practice effects (Davous et al., 1987). The advantages of the OMCT are similar to those of the MoCA in that it is brief (5–10 minutes administration time), is easy to score, requires no third-party input or special equipment or training, and has high sensitivity and specificity in persons with mild cognitive impairments and dementia. Its disadvantages reflect (1) bias due to education such that OMCT scores are related to years of education and (2) low specificity with elderly black community residents (Fillenbaum et al., 1990). The OMCT is not widely used despite its advantages.

Brief Assessment of Cognition (BAC) The BAC is a neuropsychological battery assessing verbal learning (List Learning), working memory (Digit Sequencing Test), verbal fluency (Category Instances and Controlled Word Association Test), information processing speed (Symbol Coding), motor speed (Token Motor Task), and problem solving (Tower of London) (Keefe et al., 2004). These cognitive domains often are impaired in individuals with schizophrenia and other severe mental illnesses and are correlated highly with community functioning (Keefe et al., 2006). The BAC predicted work outcomes in approximately 900 subjects with schizophrenia, schizoaffective disorders, and mood disorders as part of the Mental Health Treatment Study (McGurk et al., 2018). It was sufficiently sensitive to detect minimal cognitive impairment (low to average range) in a group of Social Security Disability Insurance (SSDI) recipients, with verbal learning being the strongest predictor of work functioning of any cognitive domain or demographic characteristic assessed. The BAC demonstrates evidence of concurrent and predictive validity, test-retest reliability (ICC = 0.79 or greater), and interrater reliability (ICC = 0.89), with established norms for adults with mental illness and normal-functioning adults of all ages (Keefe et al., 2008). Its advantages include standardized administration requiring relatively minimal training, comprehensive cognitive assessment, brief administration time (30 minutes), being well researched with a high uptake in clinical trials and community settings, versions translated into more than eight languages, and significant use in national and international studies. Recently, test developers released a tablet administration version, the BAC App (Atkins et al., 2017). Disadvantages of the BAC are its proprietary nature; limited evidence of cross-cultural validity; and limited use in samples of neurological disorders, including acquired brain injury and dementia.

Cognitive Capacity Screening Examination (CCSE) The CCSE is a motor-free, 30-item mental status screener designed to detect cognitive limitations in individuals not receiving mental health services (Jacobs et al., 1977). A score of less than 20 indicates cognitive impairment. The CCSE assesses orientation, thought content, attention, language ability, general knowledge, short-term memory, abstraction, and judgment (Foreman, 1987; Jacobs et al., 1977). Its reliability, sensitivity, specificity, and convergent validity have been established (Foreman, 1987; Kaufman et al., 1979; Spitzer et al., 1980).

Language and Communication

Language and communication functioning includes receptive and expressive language skills in both spoken and written modalities. The mental listings' Paragraph B criterion *understand, remember, or apply information* (see Annex Table 6-1) requires consideration of language and communication skills as these skills are crucial to effective performance in all jobs. The mental functions associated with language include decoding messages; expressing ideas; and organizing semantic and symbolic meaning, structuring grammar, and producing messages. A variety of tests can be used to assess language abilities, including the Boston Naming Test, Controlled Oral Word Association, and Boston Diagnostic Aphasia Examination (Goodglass and Kaplan, 1983; Kaplan et al., 2001; Spreen and Strauss, 1991).

The NIH Toolbox provides a comprehensive set of performance assessments that allow quick assessment of cognitive, emotional, sensory, and motor functions using a tablet computer (NU, 2018b). The Toolbox includes 100 stand-alone measures and takes 30 minutes to assess cognitive, emotional, sensory, and motor function. It was developed and validated using state-of-the-science methods to enhance its psychometric properties. The Toolbox is normed on a nationally representative sample to enable cross-measure comparisons. It was designed to enable measuring outcomes in longitudinal studies, and is available in English, Spanish, and other languages. The cognition battery assesses the mental processes required to learn, thinking, knowing, remembering, judging, and problem solving (Weintraub et al., 2013). It assesses higher-level functions including language, imagination, perceptions, and the planning and execution of complex behaviors. Highlighted below are tests relevant to SSA's mental functional domains.

NIH Toolbox: Picture Vocabulary This test measures receptive vocabulary (Gershon et al., 2014), with respondents selecting the picture that most closely matches the meaning of a word displayed on a video screen using a multiple-choice option. Carlozzi and colleagues (2017) established its construct validity in a poststroke sample. The average time for its completion is 4 minutes. It is administered using a tablet computer with proprietary software that requires an annual license.

NIH Toolbox: Oral Reading Recognition This test measures reading decoding skills and crystallized cognitive abilities (Gershon et al., 2014). Respondents read aloud and pronounce letters and words as accurately as possible. The average time for completion is 3 minutes.

Learning and Memory

Learning and memory abilities include registering and storing new information and retrieving information. This domain links to the Paragraph B criterion *understand, remember, or apply information*. Memory functions include short- and long-term memory; immediate, recent, and remote memory; memory span; and retrieval of memories. Commonly used tests of learning and memory include the Wechsler Memory Scale (Wechsler, 2009), the Wide Range Assessment of Memory and Learning (Sheslow and Adams, 2003), and the California Verbal Learning Test (Delis et al., 2000; Sheslow and Adams, 2003; Wechsler, 2009).

NIH Toolbox: Picture Sequence Memory Test This test measures episodic memory (Dikmen et al., 2014; Loring et al., 2019). Respondents reproduce a sequence of pictures that are displayed on a video screen. Practice sequences and test items are available for respondents 8 years of age and older. Typical time to complete the test is 7 minutes.

NIH Toolbox: List Sorting Working Memory Test This test measures working memory (Tulsky et al., 2013, 2014). Respondents recall and sequence stimuli they hear read aloud and presented on a video screen. Average time for completion is 7 minutes.

NIH Toolbox: Auditory Verbal Test (Rey) This test measures immediate recall (Weintraub et al., 2013). Respondents listen to words presented via audio recording and recall as many as possible. Respondents with visual limitations that preclude reading may complete the Picture Sequence Memory Test. Average time for completion is 3 minutes.

Attention and Vigilance

Attention and vigilance tests measure the ability to maintain attentional focus despite typical distractions. As shown in Annex Table 6-1, attention and vigilance link to the Paragraph B criterion *concentrate, persist, or maintain pace*. Commonly used tests include the WAIS-IV working memory index, the Paced Auditory Serial Addition Test, and the Continuous Performance Test (Conners and Multi-Health Systems Staff, 2000; Gronwall, 1977; Wechsler, 2009).

NIH Toolbox: Flanker Inhibitory Control and Attention Test This test measures attention and inhibitory control (Akshoomoff et al., 2014; Zelazo et al., 2013). Respondents focus on a visual stimulus displayed on a video screen while inhibiting attention to stimuli flanking it. Average time for completion is 3 minutes.

Processing Speed

Processing speed reflects how long it takes a person to answer questions and process information. This domain links to the Paragraph B criterion *concentrate, persist, or maintain pace*. Tests of processing speed include the WAIS-IV processing speed index and the Trail Making Test Part A (Reitan and Wolfson, 1993; Wechsler, 2008).

NIH Toolbox: Pattern Comparison Processing Speed Test This test measures speed of processing (Carlozzi et al., 2015). Respondents discern whether a sequence of two simple pictures presented side by side are the same or different in 85 seconds. Average time for completion is 4 minutes.

NIH Toolbox: Oral Symbol Digit Test This test also measures speed of processing (Denboer et al., 2014). Respondents view symbol–number pairs on a video screen. They are then asked to press a number on a keyboard to indicate the number that is associated with a symbol. A Pattern Comparison Processing Speed Test is available for respondents whose motor skills preclude key pressing.

Executive Functioning

Executive functioning reflects complex cognitive abilities, including planning, prioritizing, organizing, decision making, task switching, responding to feedback, correcting errors, inhibiting behavior, and mental flexibility. Tests that assess aspects of executive functioning include the Trail Making Test Part B, the Wisconsin Card Sorting Test, and the Delis-Kaplan Executive Function System (Delis et al., 2001; Heaton, 1993; Reitan, 1992).

The standard error of IRT-delivered measures provides clues to exaggeration of symptoms or attempts to "fake bad." A large standard error suggests that test scores may not be valid. The NIH Toolbox application produces standard errors for most measures. Tests built on IRT help identify people who answer difficult items correctly but then fail easier items. Standard errors are likely to be inflated when test takers "fake bad" in an inconsistent manner.

NIH Toolbox: Flanker Inhibitory Control and Attention Test This test measures attention and inhibitory control (Weintraub et al., 2013; Zelazo et al., 2013). Respondents focus on a visual stimulus displayed on a video screen while inhibiting attention to stimuli flanking it. Average time for completion is 3 minutes.

NIH Toolbox: Dimensional Change Card Sort Test This test measures cognitive flexibility and attention (Weintraub et al., 2013; Zelazo et al., 2013). Respondents view pictures that vary in two dimensions, such as shape and color. A word displayed on the video screen cues them as to which dimension they should use to sort stimuli. Average time for completion is 4 minutes.

Adaptability and Work-Related Personal Interactions

Adaptability "measures characteristics of an occupation that cause a worker to adjust to changes in work routines" (DOL, 2017, p. 61) and links to the Paragraph B criterion *adapt or manage oneself*. Work-related personal interactions include cooperating with others; handling conflict; and responding to social cues, requests, and criticism (DOL, 2017, p. 64). This domain links to the Paragraph B criterion *interact with others*.

WD-FAB Behavioral Health: Social Interactions Scale This scale is grounded in a theoretic framework intended to distinguish five domains of behavioral health functioning: behavioral control, temperament and personality, adaptability, basic interactions, and workplace behaviors (Marfeo et al., 2013a,c). Four domains (self-efficacy, mood and emotions, behavioral control, and social interactions) are supported empirically (Marfeo et al., 2013b,c, 2014). The four item banks demonstrate strong reliability, accuracy, and breadth of coverage, as well as large correlations between simulated 5- or 10-item CATs and the full item bank. The six items making up the social interaction factors demonstrate excellent goodness-of-fit indices in unidimensional confirmatory factor analyses (Marfeo et al., 2013b). The correlation between a four-item CAT and all six items was 0.99 in a sample of 1,015 SSDI claimants and a comparative sample of 1,000 adults in the United States (Marfeo et al., 2013b). Administration requires less than 2 minutes.

Personal and Social Performance Scale (PSP) The PSP was derived from the SOFAS to measure social functioning. It assesses four domains: socially useful activities (e.g., work and school), personal and social relationships, self-care, and disturbing and aggressive behavior (Sivec et al., 2017). Its validity and reliability have been demonstrated in outpatients diagnosed with schizophrenia (Kawata and Revicki, 2008).

Measures of Disorder Severity and Work-Related Functional Impairment

Psychiatric disorders are generally defined by the presence of specific symptoms. Clinical care is built around psychiatric diagnosis; therefore, it is important to consider the possible link between diagnosis and work-related functional impairment. This link would be at least partially mediated by the severity of symptoms related to the specific diagnosis. As already noted, for some psychiatric diagnoses, such as schizophrenia, impairment specifically refers to a significantly reduced capacity to participate in social relationships; care for oneself; or meet basic role obligations such as those of a worker, student, or parent. Impairment sustained for a minimum of 6 months is required for the diagnosis. Although not part of the diagnostic criteria for schizophrenia, the disorder is characterized by a decline in premorbid cognitive functioning, such that cognitive levels of individuals with schizophrenia are typically below those of the general population and of other psychiatric populations, such as individuals with bipolar disorder and major depression, which can also have a very serious course (Rosenheck et al., 2006). The level of psychosocial impairment in schizophrenia is also generally lower than that of individuals with other psychiatric disorders, including those having conditions that result in significant disability, such as bipolar disorder and major depression. Despite the fact that on average, schizophrenia is the most severe of the mental illnesses, the degree of impairment can be highly variable across different areas of functioning, as well as across people with the disorder, some of whom are capable of working part- or full-time. A number of reports suggest that the employment rate among individuals diagnosed with schizophrenia ranges from 10 to 20 percent for those not receiving supported employment services (Rosenheck et al., 2006). It should be noted that better symptom and functional outcomes are often associated with the receipt of evidence-based pharmacological and psychosocial treatments, but for others, it is important to consider the possibility of significant work-related functional impairment.

This section examines whether there are symptom assessments related to some common psychiatric diagnoses that have been found to be associated with work-related functional impairment. One such diagnosis is major depression, measures for which are discussed in Chapter 7. Disorders addressed below are anxiety disorders (see *Desk Reference to the Diagnostic Criteria from DSM-V*, e.g., agoraphobia [pp. 121–122]; generalized anxiety disorder [GAD] [pp. 122–123]), obsessive-compulsive and related disorders (obsessive-compulsive disorder [OCD] [pp. 129–130]), trauma-related disorders (posttraumatic stress disorder [PTSD] [pp. 143–149]); and autism spectrum disorder [ASD] [APA, 2014]).

Anxiety Disorder

In a review of functional outcomes and anxiety symptoms, McKnight and colleagues (2016) identified 83 articles examining the relationship between common anxiety-related disorders and functional impairment, including occupational impairment. Of these articles, 40 consider PTSD, 17 OCD, 13 social anxiety disorder, 9 GAD, 6 panic disorder, and 7 agoraphobia. The review produced a total of 497 correlations between individual disorders and functional impairment. The authors found that these anxiety-related disorders were only modestly correlated with functional outcomes (social, occupational, and physical functioning). An important question is whether any specific measures used for these disorders are useful in estimating the extent of occupational impairment. The following discussion focuses on OCD and PTSD measures, as measures of other anxiety-related disorders showed weak correlations with occupational impairment.

Obsessive-Compulsive Disorder

The intrusive thoughts and compulsive behaviors found in OCD can adversely affect work performance and functioning. Mancebo and colleagues (2008) evaluated the relationship between OCD severity and functional outcomes in 238 individuals with OCD from the Brown Longitudinal OCD Study. They found that OCD severity as measured by the Yale-Brown Obsessive Compulsive Scale (Y-BOCS) was the greatest predictor of occupational disability. Specifically, occupational disability rose 2.26 times for each standard deviation score increase (5.83) on the Y-BOCS (Mancebo et al., 2008). Eisen and colleagues (2006) recruited and interviewed 197 individuals who were part of a larger-scale OCD study to evaluate functional impairment and quality of life among individuals with this disorder. They found that higher scores on the Y-BOCS were related to poorer quality-of-life measures. Specifically, a score of 20 or higher on the Y-BOCS "appeared to be an inflection point" at which impairment became significantly more pronounced. Poor work outcomes were associated with greater severity of compulsions.

Posttraumatic Stress Disorder

PTSD can reduce work functioning through limitations in the ability to sustain attention on tasks, get along with peers, and leave the safety of one's home. These effects can be worse if the trauma occurred in the workplace. Smith and colleagues (2005) conducted a study evaluating the relationship of symptom severity on the PTSD Checklist (PCL) and Clinician-Administered PTSD Scale (CAPS) with unemployment among 325 adult male Vietnam War veterans with PTSD participating in a randomized controlled trial of two different therapies for this disorder. They found that PCL and CAPS scores were 8 to 13 percent higher for individuals who were unemployed than for employed individuals. A score increase of 10 points on the CAPS was associated with a 5.9 percent increase in the probability of an individual's not working. The mean score for workers on the CAPS was 76.62 (standard deviation [SD] = 18.78), compared with 84.09 (SD = 17.74) for nonworkers (Smith et al., 2005). On the CAPS subscale for reexperiencing symptoms, workers had a mean score of 20.48 (SD = 7.02), compared with 23.2 (SD = 6.92) for nonworkers (Smith et al., 2005). On the avoidance scale, the mean score for workers was 31.72 (SD = 9.58) and for nonworkers was 34.27 (SD = 8.81). On the hyperarousal scale, workers had a mean score of 24.42 (SD = 6.07), compared with 26.62 (SD = 5.97) for nonworkers. Similarly, the mean scores on the PCL differed for workers (58.76, SD = 12.73) and nonworkers (64.03, SD = 10.66) (Smith et al., 2005). Taylor and colleagues (2006) found associations between specific PTSD symptoms and work disability (inability to work due to disability from PTSD). Hyperarousal and reexperiencing symptoms as reported via the CAPS were the symptoms most highly correlated with collecting benefits as the result of an inability to work because of PTSD.

Autism Spectrum Disorder

With respect to ASD, its distinguishing diagnostic features include impairments in social communication and social interaction across multiple contexts, coupled with restricted, repetitive patterns of behavior, interests, or activities that result in significant difficulties with current social, occupational, or community functioning (APA, 2013). Several studies have documented worse employment-related outcomes for adults with autism compared with adults with other types of mental or developmental impairments and matched on general measures of health and socioeconomic status. For instance, Roux and colleagues (2013) examined whether young adults had ever worked for pay since high school, whether they were currently employed, and whether their employment status was full-time. They compared youth with autism with four other groups: youth with an intellectual disability, youth with severe mental illness, youth with learning disabilities, and youth with language impairment. Covariates included measures of sex, age, ethnicity, household income, overall health, conversation ability, and a scale of functional skills. The adjusted odds were significant for 10 of the 12 comparisons. Youth with autism had worse outcomes on every measure for every comparison.

Few studies have examined the link between the severity of ASD's distinguishing features and the ability to work, the likelihood of employment, or work performance. Most extant studies of these employment-related outcomes among people on the autism spectrum exclude validated measures of these core autistic features and instead focus on such factors as IQ and verbal ability (which is not synonymous with social communication).

A few recent studies have begun to examine the linkages between distinctly autistic impairments and employment-related outcomes, with mixed results. For example, a nationally representative study of postsecondary outcomes among young adults with autism who had formerly received special education services examined the association between parent-rated conversation ability and whether youth had ever had any paid employment since high school. Among youth with "no trouble" conversing, 72 percent (95% confidence interval [CI]: 52.3–86.2) had ever held a job, compared with 17 percent (95% CI: 7.1–36.0) of youth who could not converse at all (Shattuck et al., 2012). On the other hand, a 10-year longitudinal study of 161 adults with autism aged 18.4 to 52.1 years at baseline examined the association between severity of autism symptoms (as measured by the Autism Diagnostic Interview-Revised) and changes over time in an ordinal vocational outcomes index. Vocational outcomes were worse for those with an intellectual disability and better for those with higher levels of independence in activities of daily living. However, there was no significant association between autistic impairments and vocational outcomes (Taylor and Mailick, 2014).

The committee found no studies examining how autistic strengths might affect employment-related outcomes, despite influential review articles suggesting this as an area for future inquiry (Scott et al., 2018).

FINDINGS AND CONCLUSIONS

Findings

6-1. It is important to assess the persistence of impairment due to mental disorders, given the possibility of episodic or persistent symptoms.

6-2. When assessing mental functional abilities relevant to work requirements, it is important to assess the following domains: general cognitive/intellectual ability, language and communication, learning and memory, attention and vigilance, processing speed, executive functioning, adaptability, and work-related personal interactions.

6-3. The Work Disability Functional Assessment Battery (WD-FAB) is unique in its use of item response theory and computer adaptive testing.

6-4. The Mental Illness Research, Education, and Clinical Center (MIRECC) Global Assessment of Functioning (GAF) provides separate scores for symptoms, social functioning simulations, and occupational functioning, and has demonstrated the strongest convergent validity in occupational scores, which were strongly correlated with employment in the past month and work status.

6-5. While the Social and Occupational Functioning Assessment Scale (SOFAS) advances measurement by separating the original GAF into one scale for symptoms and another for social and occupational functioning, SOFAS scores confound work and social functioning.

6-6. The Specific Level of Functioning Scale demonstrated significantly strong correlations with interpersonal and work function compared with the other areas of function.

6-7. The Brief Assessment of Cognition demonstrates evidence of predictive validity for work outcomes in persons with schizophrenia, schizoaffective disorders, and mood disorders.

6-8. The Yale-Brown Obsessive Compulsive Scale demonstrates evidence of predictive validity for occupational disability in persons with obsessive-compulsive disorder (OCD).

6-9. The National Institutes of Health Toolbox provides a comprehensive set of performance assessments that allow quick assessment of cognitive, emotional, sensory, and motor functions using a tablet computer.

6-10. Anxiety disorders are only modestly correlated with functional outcomes (social, occupational, and physical functioning).

6-11. The intrusive thoughts and compulsive behaviors found in OCD can adversely affect work performance and functioning.

6-12. Posttraumatic stress disorder can reduce work functioning through limitations in the ability to sustain attention on tasks, get along with peers, and leave the safety of one's home.

6-13. Symptoms associated with depression, including fatigue, difficulty concentrating, and slowed response speed, can impair work functioning.

6-14. On average, young adults with autism have worse employment outcomes relative to youth with other types of impairments, after adjusting for a range of covariates.

6-15. Results of cognitive testing are likely to be less stable for individuals whose mental disorders are characterized by an intermittent or fluctuating course than for those with stable conditions.

Conclusions

6-1. There are no conclusive studies examining the association between the severity of impairments specific to autism and abilities relevant to work.

6-2. Understanding the relationship between mental illness and functioning is important because some major mental illnesses are episodic in nature, with severity of symptoms and functional impairments varying over time, and with periods of greater severity ranging from weeks to months.

6-3. There is no single measure that captures all important aspects of mental abilities needed for work, although the WD-FAB, as a self-report battery of relevant questions, shows promise. More development work is needed for the WD-FAB to fulfill its promise for use in disability determination.

6-4. It is important to perform more frequent assessments of disability applicants with mental disorders that are characterized by an intermittent or fluctuating course.

REFERENCES

Akshoomoff, N., E. Newman, W. K. Thompson, C. McCabe, C. S. Bloss, L. Chang, D. G. Amaral, B. J. Casey, T. M. Ernst, J. A. Frazier, J. R. Gruen, W. E. Kaufmann, T. Kenet, D. N. Kennedy, O. Libiger, S. Mostofsky, S. S. Murray, E. R. Sowell, N. Schork, A. M. Dale, and T. L. Jernigan. 2014. The NIH Toolbox Cognition Battery: Results from a large normative developmental sample (PING). *Neuropsychology* 28(1):1–10. doi: 10.1037/neu0000001.

APA (American Psychiatric Association). 2013. *Diagnostic and statistical manual of mental disorders: DSM-5*. Arlington, VA: APA.

APA. 2014. *Desk reference to the diagnostic criteria from DSM-5*. Washington, DC: American Psychiatric Publishing.

Atkins, A. S., T. Tseng, A. Vaughan, E. W. Twamley, P. Harvey, T. Patterson, M. Narashimhan, and R. S. Keefe. 2017. Validation of the tablet-administered Brief Assessment of Cognition (BAC App). *Schizophrenia Research* 181:100–106. doi: 10.1016/j.schres.2016.10.010.

Bagby, R. M., A. G. Ryder, D. R. Schuller, and M. B. Marshall. 2004. The Hamilton Depression Rating Scale: Has the gold standard become a lead weight? *American Journal of Psychiatry* 161(12):2163–2177.

Beck, A. T., N. Epstein, G. Brown, and R. A. Steer. 1988. An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology* 56(6):893–897.

Bieliauskas, L. A. 1998. The Houston conference on specialty education and training in clinical neuropsychology. *Archives of Clinical Neuropsychology* 13(2):160–166.

Bovin, M. J., B. P. Marx, F. W. Weathers, M. W. Gallagher, P. Rodriguez, P. P. Schnurr, and T. M. Keane. 2015. Psychometric properties of the PTSD checklist for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (PCL-5) in veterans. *Psychological Assessment* 28(11):1379–1391. doi: 10.1037/pas0000254.

Bowie, C. R., A. Reichenberg, T. L. Patterson, R. K. Heaton, and P. D. Harvey. 2006. Determinants of real-world functional performance in schizophrenia subjects: Correlations with cognition, functional capacity, and symptoms. *American Journal of Psychiatry* 163(3):418–425.

Bowie, C. R., E. W. Twamley, H. Anderson, B. Halpern, T. L. Patterson, and P. D. Harvey. 2007. Self-assessment of functional status in schizophrenia. *Journal of Psychiatric Research* 41(12):1012–1018.

Carlozzi, N. E., J. L. Beaumont, D. S. Tulsky, and R. C. Gershon. 2015. The NIH Toolbox Pattern Comparison Processing Speed Test: Normative data. *Archives of Clinical Neuropsychology* 30(5):359–368.

Carlozzi, N. E., D. S. Tulsky, T. J. Wolf, S. Goodnight, R. K. Heaton, K. B. Casaletto, A. W. K. Wong, C. M. Baum, R. C. Gershon, and A. W. Heinemann. 2017. Construct validity of the NIH Toolbox Cognition Battery in individuals with stroke. *Rehabilitation Psychology* 62(4):443–454.

Chandler, D., G. Spicer, M. Wagner, and W. Hargreaves. 1999. Cost-effectiveness of a capitated assertive community treatment program. *Psychiatric Rehabilitation Journal* 22(4):327. doi: 10.1037/h0095219.

Conners, C. K., and Multi-Health Systems Staff (Eds.). 2000. *Conners' Continuous Performance Test II: Computer program for windows technical guide and software manual*. North Tonawanda, NY: Multi-Health Systems.

Cramer, J. A., R. Rosenheck, W. Xu, J. Thomas, W. Henderson, and S. Dennis. 2000. Quality of life in schizophrenia: A comparison of instruments. *Schizophrenia Bulletin* 26(3):659–666.

Davous, P., Y. Lamour, E. Debrand, and P. Rondot. 1987. A comparative evaluation of the short orientation memory concentration test of cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry* 50(10):1312–1317.

Delis, D. C., J. H. Kramer, and E. Kaplan. 2000. *California Verbal Learning Test: CVLT-II; adult version; manual*. San Antonio, TX: The Psychological Corporation.

Delis, D. C., E. Kaplan, and J. Kramer. 2001. *Delis-Kaplan Executive Function System*. San Antonio, TX: The Psychological Corporation.

Denboer, J. W., C. Nicholls, C. Corte, and K. Chestnut. 2014. National Institutes of Health Toolbox Cognition Battery. *Archives of Clinical Neuropsychology* 29(7):692–694.

Derogatis, L. R., R. S. Lipman, and I. Covi. 1973. The SCL-90: An outpatient psychiatric rating scale. *Psychopharmacology Bulletin* 9(1):13–28.

Derogatis, L. R., R. S. Lipman, K. Rickels, E. H. Uhlenhuth, and L. Covi. 1974. The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science* 19(1):1–15.

Dikmen, S. S., P. J. Bauer, S. Weintraub, D. Mungas, J. Slotkin, J. L. Beaumont, R. Gershon, N. R. Temkin, and R. K. Heaton. 2014. Measuring episodic memory across the lifespan: NIH Toolbox Picture Sequence Memory Test. *Journal of the International Neuropsychological Society* 20(6):611–619.

DOL (U.S. Department of Labor). 2017. *Occupational requirements survey collection manual*, 3rd ed. https://www.bls.gov/ors/about/information-for-survey-participants/pdf/occupational-requirements-survey-collection-manual-072017.pdf (accessed April 4, 2019).

DOL. 2018. *Occupational requirements survey collection manual*, 4th ed. https://www.bls.gov/ors/about/information-for-survey-participants/pdf/occupational-requirements-survey-collection-manual-082018.pdf (accessed April 4, 2019).

Eisen, J. L., M. A. Mancebo, A. Pinto, M. E. Coles, M. E. Pagano, R. Stouf, and S. A. Rasmussen. 2006. Impact of obsessive-compulsive disorder on quality of life. *Comprehensive Psychiatry* 47(4):270–275.

Ermel, J., C. S. Carter, J. M. Gold, A. W. MacDonald III, J. D. Ragland, S. M. Silverstein, M. E. Strauss, and D. M. Barch. 2017. Self versus informant reports on the Specific Levels of Functioning Scale: Relationships to depression and cognition in schizophrenia and schizoaffective disorder. *Schizophrenia Research: Cognition* 9:1–7.

Fillenbaum, G., A. Heyman, K. Williams, B. Prosnitz, and B. Burchett. 1990. Sensitivity and specificity of standardized screens of cognitive impairment and dementia among elderly black and white community residents. *Journal of Clinical Epidemiology* 43(7):651–660.

Foreman, M. D. 1987. Reliability and validity of mental status questionnaires in elderly hospitalized patients. *Nursing Research* 36(4):216–220.

Gershon, R. C., K. F. Cook, D. Mungas, J. J. Manly, J. Slotkin, J. L. Beaumont, and S. Weintraub. 2014. Language measures of the NIH Toolbox Cognition Battery. *Journal of the International Neuropsychological Society* 20(6):642–651.

Goodglass, H., and E. Kaplan. 1983. *Boston Diagnostic Aphasia Examination*. Philadelphia, PA: Lea & Febiger.

Goodman, W. K., L. H. Price, S. A. Rasmussen, C. Mazure, R. L. Fleischmann, C. L. Hill, G. R. Heninger, and D. S. Charney. 1989. Yale-Brown Obsessive Compulsive Scale (Y-BOCS). *Archives of General Psychiatry* 46(11):1006–1011.

Gronwall, D. M. 1977. Paced auditory serial-addition task: A measure of recovery from concussion. *Perceptual and Motor Skills* 44(2):367–373. doi: 10.2466/pms.1977.44.2.367.

Hamilton, M. 1960. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry* 23(1):56–62.

Hannula, J. A., K. Lahtela, A. Jarvikoski, J. K. Salminen, and P. Makela. 2006. Occupational Functioning Scale (OFS)—An instrument for assessment of work ability in psychiatric disorders. *Nordic Journal of Psychiatry* 60(5):372–378.

Heaton, R. K. 1993. *Wisconsin Card Sorting Test: Computer version*. Odessa, FL: Psychological Assessment Resources.

Hu, M. T., K. Szewczyk-Królikowski, P. Tomlinson, K. Nithi, M. Rolinski, C. Murray, K. Talbot, K. P. Ebmeier, C. E. Mackay, and Y. Ben-Shlomo. 2014. Predictors of cognitive impairment in an early stage Parkinson's disease cohort. *Movement Disorders* 29(3):351–359.

Huppert, J. D., M. R. Walther, G. Hajcak, E. Yadin, E. B. Foa, H. B. Simpson, and M. R. Liebowitz. 2007. The OCI-R: Validation of the subscales in a clinical sample. *Journal of Anxiety Disorders* 21(3):394–406.

IOM (Institute of Medicine). 2015. *Psychological testing in the service of disability determination*. Washington, DC: The National Academies Press.

Jacobs, J. W., M. R. Bernhard, A. Delgado, and J. J. Strain. 1977. Screening for organic mental syndromes in the medically ill. *Annals of Internal Medicine* 86(1):40–46.

Kaplan, E., H. Goodglass, and S. Weintraub. 2001. *Boston Naming Test*. Austin, TX: PRO-ED.

Katzman, R., T. Brown, P. Fuld, A. Peck, R. Schechter, and H. Schimmel. 1983. Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. *American Journal of Psychiatry* 140(6):734–739.

Kaufman, D. M., M. Weinberger, J. J. Strain, and J. W. Jacobs. 1979. Detection of cognitive deficits by a brief mental status examination: The cognitive capacity screening examination, a reappraisal and a review. *General Hospital Psychiatry* 1(3):247–255.

Kawata, A. K., and D. A. Revicki. 2008. Psychometric properties of the Personal and Social Performance scale (PSP) among individuals with schizophrenia living in the community. *Quality of Life Research* 17(10):1247–1256.

Keefe, R. S., T. E. Goldberg, P. D. Harvey, J. M. Gold, M. P. Poe, and L. Coughenour. 2004. The Brief Assessment of Cognition in Schizophrenia: Reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophrenia Research* 68(2–3):283–297.

Keefe, R. S., M. Poe, T. M. Walker, and P. D. Harvey. 2006. The relationship of the Brief Assessment of Cognition in Schizophrenia (BACS) to functional capacity and real-world functional outcome. *Journal of Clinical and Experimental Neuropsychology* 28(2):260–269.

Keefe, R. S., P. D. Harvey, T. E. Goldberg, J. M. Gold, T. M. Walker, C. Kennel, and K. Hawkins. 2008. Norms and standardization of the Brief Assessment of Cognition in Schizophrenia (BACS). *Schizophrenia Research* 102(1–3):108–115.

Kroenke, K., R. L. Spitzer, and J. B. Williams. 2001. The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine* 16(9):606–613.

Lezak, M., D. Howieson, E. Bigler, and D. Tranel. 2012. *Neuropsychological assessment*, 5th ed. New York: Oxford University Press.

Loring, D. W., S. C. Bowden, E. Staikova, J. A. Bishop, D. L. Drane, and F. C. Goldstein. 2019. NIH Toolbox Picture Sequence Memory Test for assessing clinical memory function: Diagnostic relationship to the Rey Auditory Verbal Learning Test. *Archives of Clinical Neuropsychology* 34(2):268–276. doi: 10.1093/arclin/acy028.

Mancebo, M. C., B. Greenberg, J. E. Grant, A. Pinto, J. L. Eisen, I. Dyck, and S. A. Rasmussen. 2008. Correlates of occupational disability in a clinical sample of obsessive-compulsive disorder. *Comprehensive Psychiatry* 49(1):43–50.

Marfeo, E. E., S. M. Haley, A. M. Jette, S. V. Eisen, P. Ni, K. Bogusz, M. Meterko, C. M. McDonough, L. Chan, D. E. Brandt, and E. K. Rasch. 2013a. Conceptual foundation for measures of physical function and behavioral health function for Social Security work disability evaluation. *Archives of Physical Medicine and Rehabilitation* 94(9):1645–1652.

Marfeo, E. E., P. Ni, S. M. Haley, A. M. Jette, K. Bogusz, M. Meterko, C. M. McDonough, L. Chan, D. E. Brandt, and E. K. Rasch. 2013b. Development of an instrument to measure behavioral health function for work disability: Item pool construction and factor analysis. *Archives of Physical Medicine and Rehabilitation* 94(9):1670–1678.

Marfeo, E. E., P. Ni, S. M. Haley, K. Bogusz, M. Meterko, C. M. McDonough, L. Chan, E. K. Rasch, D. E. Brandt, and A. M. Jette. 2013c. Scale refinement and initial evaluation of a behavioral health function measurement tool for work disability evaluation. *Archives of Physical Medicine and Rehabilitation* 94(9):1679–1686.

Marfeo, E. E., P. Ni, L. Chan, E. K. Rasch, and A. M. Jette. 2014. Combining agreement and frequency rating scales to optimize psychometrics in measuring behavioral health functioning. *Journal of Clinical Epidemiology* 67(7):781–784.

Marfeo, E. E., P. Ni, C. McDonough, K. Peterik, M. Marino, M. Meterko, E. K. Rasch, L. Chan, D. Brandt, and A. M. Jette. 2018. Improving assessment of work related mental health function using the Work Disability Functional Assessment Battery (WD-FAB). *Journal of Occupational Rehabilitation* 28(1):190–199.

Mausbach, B. T., C. R. Bowie, P. D. Harvey, E. W. Twamley, S. R. Goldman, D. V. Jeste, and T. L. Patterson. 2008. Usefulness of the UCSD Performance-Based Skills Assessment (UPSA) for predicting residential independence in patients with chronic schizophrenia. *Journal of Psychiatric Research* 42(4):320–327.

Mausbach, B. T., P. D. Harvey, A. E. Pulver, C. A. Depp, P. S. Wolyniec, M. H. Thornquist, J. R. Luke, J. A. McGrath, C. R. Bowie, and T. L. Patterson. 2010. Relationship of the Brief UCSD Performance-Based Skills Assessment (UPSA-B) to multiple indicators of functioning in people with schizophrenia and bipolar disorder. *Bipolar Disorders* 12(1):45–55.

Mausbach, B. T., C. A. Depp, C. R. Bowie, P. D. Harvey, J. A. McGrath, M. H. Thronquist, J. R. Luke, P. S. Wolyniec, A. E. Pulver, and T. L. Patterson. 2011. Sensitivity and specificity of the UCSD Performance-Based Skills Assessment (UPSA-B) for identifying functional milestones in schizophrenia. *Schizophrenia Research* 132(2–3):165–170.

McGurk, S. R., R. E. Drake, H. Xie, J. Riley, R. Milfort, T. W. Hale, and W. Frey. 2018. Cognitive predictors of work among Social Security Disability Insurance beneficiaries with psychiatric disorders enrolled in IPS supported employment. *Schizophrenia Bulletin* 44(1):32–47.

McKnight, P. E., and T. B. Kashdan. 2009. The importance of functional impairment to mental health outcomes: A case for reassessing our goals in depression treatment research. *Clinical Psychology Review* 29(3):243–259.

McKnight, P. E., S. S. Monfort, T. B. Kashdan, D. V. Blalock, and J. M. Calton. 2016. Anxiety symptoms and functional impairment: A systematic review of the correlation between the two measures. *Clinical Psychology Review* 45:115–130. doi: 10.1016/j.cpr.2015.10.005.

Meterko, M., M. Marino, P. Ni, E. Marfeo, C. M. McDonough, A. Jette, K. Peterik, E. Rasch, D. E. Brandt, and L. Chan. 2018. Psychometric evaluation of the improved Work-Disability Functional Assessment Battery. *Archives of Physical Medicine and Rehabilitation*. In press. Corrected proof available online December 19, 2018. doi.org/10.1016/j.apmr.2018.09.125.

Morosini, P. L., L. Magliano, L. Brambilla, S. Ugolini, and R. Pioli. 2000. Development, reliability and acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale (SOFAS) to assess routine social funtioning. *Acta Psychiatrica Scandinavica* 101(4):323–329.

Mosbach, P., S. Campanelli, A. D'Anello, L. Klaus, E. Aberger, J. Pelletier, P. Quinn, V. Sundar, A. Houtenville, B. Ogut, H. Brotsos, and O. Urdapilleta. 2018. Mental cognitive demands: Contract No. 28321318A00040029/28321318FA0011262. Summit Consulting.

Mucci, A., P. Rucci, P. Rocca, P. Bucci, D. Gibertoni, E. Merlotti, S. Galderisi, and M. Maj. 2014. The Specific Level of Functioning Scale: Construct validity, internal consistency and factor structure in a large Italian sample of people with schizophrenia living in the community. S*chizophrenia Research* 159(1):144–150.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2016. *Informing Social Security's process for financial capability determination*. Washington, DC: The National Academies Press.

Nasreddine, Z. S., N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, and H. Chertkow. 2005. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53(4):695–699.

Niv, N., A. N. Cohen, G. Sullivan, and A. S. Young. 2007. The MIRECC version of the Global Assessment of Functioning Scale: Reliability and validity. *Psychiatric Services* 58(4):529–535.

NU (Northwestern University). 2018a. *Cognition measures: NIH Toolbox Cognition Batteries*. http://www.healthmeasures.net/explore-measurement-systems/nih-toolbox/intro-to-nih-toolbox/cognition (accessed April 4, 2019).

NU. 2018b. *Why use NIH Toolbox?* http://www.healthmeasures.net/explore-measurement-systems/nih-toolbox (accessed April 4, 2019).

OIDAP (Occupational Information Development Advisory Panel). 2009. *Mental cognitive subcommittee: Content model and classification recommendations*. http://www.ssa.gov/oidap/Documents/AppendixC.pdf (accessed April 4, 2019).

O*NET (National Center for O*NET Development). 2018a. *Cognitive abilities*. https://www.onetonline.org/find/descriptor/browse/Abilities/1.A.1 (accessed April 4, 2019).

O*NET. 2018b. *Mental processes*. https://www.onetonline.org/find/descriptor/browse/Work_Activities/4.A.2 (accessed April 4, 2019).

Patterson, T. L., S. Goldman, C. L. McKibbin, T. Hughs, and D. V. Jeste. 2001. UCSD Performance-Based Skills Assessment: Development of a new measure of everyday functioning for severely mentally ill adults. *Schizophrenia Bulletin* 27(2):235–245.

Rapp, A. M., R. L. Bergman, J. Piacentini, and J. F. McGuire. 2016. Evidence-based assessment of obsessive-compulsive disorder. *Journal of Central Nervous System Disease* 8:13–29. doi: 10.4137/JCNSD.S38359.

Reitan, R. M. 1992. *Trail Making Test: Manual for administration and scoring*. Mesa, AZ: Reitan Neuropsychology Laboratory.

Reitan, R. M., and D. Wolfson. 1993. *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation*, 2nd ed. Tucson, AZ: Neuropsychology Press.

Rosenheck, R., D. Leslie, R. Keefe, J. McEvoy, M. Swartz, D. Perkins, S. Stroup, J. K. Hsiao, J. Lieberman, and CATIE Study Investigators Group. 2006. Barriers to employment for people with schizophrenia. *American Journal of Psychiatry* 163(3):411–417.

Roux, A. M., P. T. Shattuck, B. P. Cooper, K. A. Anderson, M. Wagner, and S. C. Narendorf. 2013. Postsecondary employment experiences among young adults with an autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry* 52(9):931–939.

Rybarczyk, B. 2011. Social and Occupational Functioning Assessment Scale (SOFAS). In *Encyclopedia of Clinical Neuropsychology*. New York: Springer. P. 2313.

Schneider, L. C., and E. L. Struening. 1983. SLOF: A behavioral rating scale for assessing the mentally ill. *Social Work Research and Abstracts* 19(3):9–21.

Scott, M., B. Milbourn, M. Falkmer, M. Black, S. Bölte, A. Halladay, M. Lerner, J. L. Taylor, and S. Girdler. 2018. Factors impacting employment for people with autism spectrum disorder: A scoping review. *Autism* August 3:1362361318787789. doi: 10.1177/1362361318787789.

Shattuck, P. T., S. C. Narendorf, B. P. Cooper, P. R. Sterzing, M. Wagner, and J. L. Taylor. 2012. Postsecondary education and employment among youth with an autism spectrum disorder. *Pediatrics* 129(6):1042–1049.

Sheslow, D., and W. Adams. 2003. *Wide Range Assessment of Memory and Learning Second Edition: Administration and technical manual*. Lutz, FL: Psychological Assessment Resources.

Sivec, H. J., V. L. Montesano, D. Skubby, K. A. Knepp, and M. R. Munetz. 2017. Cognitive Behavioral Therapy for Psychosis (CBT-p) delivered in a community mental health setting: A case comparison of clients receiving CBT informed strategies by case managers prior to therapy. *Community Mental Health Journal* 53(2):134–142.

Smith, M. W., P. P. Schnurr, and R. R. Rosenheck. 2005. Employment outcomes and PTSD symptom severity. *Mental Health Services Research* 7(2):89–101.

Spielberger, C. D. 1983. *Manual for the State-Trait Anxiety Inventory STAI (form Y) ("self-evaluation questionnaire")*. https://ubir.buffalo.edu/xmlui/handle/10477/1873 (accessed April 24, 2019).

Spielberger, C. D. 1989. *State-Trait Anxiety Inventory: A comprehensive bibliography*. Palo Alto, CA: Consulting Psychologists Press.

Spitzer, R. L., K. Kroenke, and J. B. Williams. 1980. *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association.

Spreen, O., and E. Strauss. 1991. *Controlled oral word association (word fluency)*. Oxford, UK: Oxford University Press.

SSA (U.S. Social Security Administration). 2015. *Function report adult form SSA-3373-BK*. https://www.ssa.gov/forms/ssa-3373-bk.pdf (accessed April 4, 2019).

SSA. n.d.-a. *Disability evaluation under Social Security: 12.00 mental disorders adult*. https://www.ssa.gov/disability/professionals/bluebook/12.00-MentalDisorders-Adult.htm (accessed April 4, 2019).

SSA. n.d.-b. *Mental residual functional capacity assessment form SSA-4734-F4-SUP*. http://ssaconnect.com/tfiles/SSA-4734-F4-SUP.pdf (accessed April 4, 2019).

Strauss, E., E. M. Sherman, and O. Spreen. 2006. *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford, UK: Oxford University Press.

Taylor, J. L., and M. R. Mailick. 2014. A longitudinal examination of 10-year change in vocational and educational activities for adults with autism spectrum disorders. *Developmental Psychology* 50(3):699–708. doi: 10.1037/a0034297.

Taylor, S., J. Wald, and G. J. G. Asmundson. 2006. Factors associated with occupational impairment in people seeking treatment for posttraumatic stress disorder. *Canadian Journal of Community Mental Health* 25(2):289–301.

Tulsky, D. S., N. E. Carlozzi, N. Chevalier, K. A. Espy, J. L. Beaumont, and D. Mungas. 2013. NIH Toolbox Cognition Battery (CB): Measuring working memory. *Monographs of the Society for Research in Child Development* 78(4):70–87.

Tulsky, D. S., N. Carlozzi, N. D. Chiaravalloti, J. L. Beaumont, P. A. Kisala, D. Mungas, K. Conway, and R. Gershon. 2014. NIH Toolbox Cognition Battery (NIHTB-CB): List sorting test to measure working memory. *Journal of the International Neuropsychological Society* 20(6):599–610.

Weathers, F. W., D. D. Blake, P. P. Schnurr, D. G. Kaloupek, B. P. Marx, and T. M. Keane. 2013a. *The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5)*. https://www.ptsd.va.gov/professional/assessment/adult-int/caps.asp (accessed April 4, 2019).

Weathers, F. W., B. T. Litz, T. M. Keane, P. A. Palmieri, B. P. Marx, and P. P. Schnurr. 2013b. *The PTSD Checklist for DSM-5 (PCL-5)*. https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp (accessed April 4, 2019).

Wechsler, D. 2008. *Wechsler Adult Intelligence Scale (WAIS-IV)*, 4th ed. San Antonio, TX: NCS Pearson.

Wechsler, D. 2009. *WMS-IV: Wechsler Memory Scale—Administration and scoring manual.* San Antonio, TX: The Psychological Corporation.

Weintraub, S., S. S. Dikmen, R. K. Heaton, D. S. Tulsky, P. D. Zelazo, P. J. Bauer, N. E. Carlozzi, J. Slotkin, D. Blitz, K. Wallner-Allen, N. A. Fox, J. L. Beaumont, D. Mungas, C. J. Nowinski, J. Richler, J. A. Deocampo, J. E. Anderson, J. J. Manly, B. Borosh, R. Havlik, K. Conway, E. Edwards, L. Freund, J. W. King, C. Moy, E. Witt, and R. C. Gershon. 2013. Cognition assessment using the NIH Toolbox. *Neurology* 80(11 Suppl. 3):S54–S64.

WHO (World Health Organization). 2010. *Measuring health and disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0)*. Geneva, Switerzerland: WHO.

Zelazo, P. D., J. E. Anderson, J. Richler, K. Wallner-Allen, J. L. Beaumont, and S. Weintraub. 2013. II. NIH Toolbox Cognition Battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development* 78(4):16–33.

ANNEX TABLE 6-1
Mental Functional Domains Relevant to Work Requirements

Mental Functional Domains (identified by the committee)	Paragraph B Criteria	Occupational Requirements Survey	Occupational Information Network (O*NET)	
			Abilities	Work Activities
General Cognitive/ Intellectual Ability		Decision making (DOL, 2017); Problem solving (DOL, 2018)	Inductive reasoning Mathematical reasoning Oral comprehension Written comprehension	Analyzing data or information Judging the qualities of things, services, or people Developing objectives and strategies Making decisions and solving problems Organizing, planning, and prioritizing work
Language and Communication	Understand, remember, or apply information; Interact with others[a]	Work-related personal interactions (DOL, 2017); Personal contacts: verbal interactions and people skills (DOL, 2018)	Oral comprehension Oral expression Written comprehension Written expression	Processing information
Learning and Memory	Understand, remember, or apply information		Memorization Number facility Oral comprehension Written comprehension	Evaluating information to determine compliance with standards Updating and using relevant knowledge

[a] This text has been revised since prepublication release.

	SSA Psychiatric Review Technique Form (SSA-2506-BK)	SSA Function Report-Adult and Third Party (SSA-3373-BK/ SSA-3380-BK)	SSA Mental Residual Functional Capacity Assessment (SSA-4734-F4-SUP)
	Executive function	Completing tasks	
	Understand, remember, or apply information		
	Language	Getting along with others	Social interaction
	Understand, remember, or apply information[a]	Understanding	
			Understanding and memory
	Interact with others	Completing tasks	
	Learning and memory	Memory	Understanding and memory
	Understand, remember, or apply information	Understanding	

continued

ANNEX TABLE 6-1
Continued

Mental Functional Domains (identified by the committee)	Paragraph B Criteria	Occupational Requirements Survey	Occupational Information Network (O*NET)	
			Abilities	Work Activities
Attention and Vigilance	Concentrate, persist, or maintain pace	Adaptability (DOL, 2017)	Problem sensitivity	Evaluating information to determine compliance with standards
				Analyzing data or information
				Organizing, planning, and prioritizing work
				Updating and using relevant knowledge
			Selective attention	Scheduling work and activities
				Making decisions and solving problems
				Developing objectives and strategies

	SSA Psychiatric Review Technique Form (SSA-2506-BK)	SSA Function Report-Adult and Third Party (SSA-3373-BK/ SSA-3380-BK)	SSA Mental Residual Functional Capacity Assessment (SSA-4734-F4-SUP)
	Complex attention	Concentration	Sustained concentration and persistence
	Concentrate, persist, or maintain pace		

continued

ANNEX TABLE 6-1
Continued

Mental Functional Domains (identified by the committee)	Paragraph B Criteria	Occupational Requirements Survey	Occupational Information Network (O*NET)		
			Abilities	Work Activities	
Processing Speed	Concentrate, persist, or maintain pace	Adaptability (DOL, 2017)	Perceptual speed	Evaluating information to determine compliance with standards	
			Fluency of ideas	Organizing, planning, and prioritizing work	
		Pace (DOL, 2017, 2018)			
			Speed of closure	Processing information	

	SSA Psychiatric Review Technique Form (SSA-2506-BK)	SSA Function Report-Adult and Third Party (SSA-3373-BK/ SSA-3380-BK)	SSA Mental Residual Functional Capacity Assessment (SSA-4734-F4-SUP)
	Perceptual-motor	Understanding	Understanding and memory
		Following instructions	
		Completing tasks	

continued

ANNEX TABLE 6-1
Continued

Mental Functional Domains (identified by the committee)	Paragraph B Criteria	Occupational Requirements Survey	Occupational Information Network (O*NET)		
			Abilities	Work Activities	
Executive Functioning		Decision making (DOL, 2017); Problem solving (DOL, 2018)	Category flexibility	Analyzing data or information	
				Updating and using relevant knowledge	
			Problem sensitivity	Developing objectives and strategies	
			Inductive reasoning	Thinking creatively	
				Judging the qualities of things, services, or people	
		Adaptability (DOL, 2017)	Deductive reasoning	Evaluating information to determine compliance with standards	
			Information ordering	Making decisions and solving problems	
				Organizing, planning, and prioritizing work	
			Flexibility of closure	Scheduling work and activities	

	SSA Psychiatric Review Technique Form (SSA-2506-BK)	SSA Function Report-Adult and Third Party (SSA-3373-BK/ SSA-3380-BK)	SSA Mental Residual Functional Capacity Assessment (SSA-4734-F4-SUP)
	Executive function	Completing tasks	Adaptation
	Understand, remember, or apply information		
	Complex attention	Following instructions	
	Adapt or manage oneself		

continued

ANNEX TABLE 6-1
Continued

Mental Functional Domains (identified by the committee)[b]	Paragraph B Criteria	Occupational Requirements Survey	Occupational Information Network (O*NET)		
			Abilities	Work Activities	
Adaptability	Adapt or manage oneself	Adaptability (DOL, 2017)			
Work-Related Personal Interactions	Interact with others	Work-related personal interactions (DOL, 2017); Personal contacts: verbal interactions and people skills (DOL, 2018)			

[b]The domains of "adaptability" and "work-related personal interactions" are included in the cognitive demand elements in the July 2017 version of the Occupation Requirements Survey (ORS) Collection Manual. The table was updated to include both the 2017 elements and the revised elements from the August 2018 version of the ORS Collection Manual.

	SSA Psychiatric Review Technique Form (SSA-2506-BK)	SSA Function Report-Adult and Third Party (SSA-3373-BK/ SSA-3380-BK)	SSA Mental Residual Functional Capacity Assessment (SSA-4734-F4-SUP)
	Adapt or manage oneself	Completing tasks	Adaptation
		Following instructions	
		Getting along with others	
	Interact with others	Getting along with others	Social interaction

ANNEX TABLE 6-2
Definitions of Mental Functional Domains

Mental Functional Domains[a]	Definition
General Cognitive/ Intellectual Ability	How well a person can reason, solve problems, and meet cognitive demands of varied complexity (OIDAP, 2009, p. C-21)
Language and Communication	How well a person can understand spoken or written language, communicate his or her thoughts, and follow directions (OIDAP, 2009, p. C-21)
Learning and Memory	How well a person can learn and remember new information (OIDAP, 2009, p. C-21)
Attention and Vigilance	How well a person can sustain the focus of attention in a work environment with ordinary distractions (OIDAP, 2009, p. C-22)
Processing Speed	How quickly a person can respond to questions and process information (OIDAP, 2009, p. C-22)
Executive Functioning	How well a person can plan, prioritize, organize, sequence, initiate, and execute multistep procedures (OIDAP, 2009, p. C-22)
Adaptability	Measures characteristics of an occupation that cause a worker to adjust to changes in work routines (DOL, 2017, p. 61)
Work-Related Personal Interactions	The requirement of a job to cooperate with others; handle conflict; and respond to social cues, requests, and criticism (DOL, 2017, p. 64)

[a]The domains of "adaptability" and "work-related personal interactions" are included in the cognitive demand elements in the July 2017 version of the Occupation Requirements Survey (ORS) Collection Manual. The cognitive elements were updated in an August 2018 version of the ORS Collection Manual, which became available following the committee's work on this chapter.

SOURCES: DOL, 2017; OIDAP, 2009.

CHAPTER 6 ANNEX TABLES CONTINUE ON THE NEXT PAGE

ANNEX TABLE 6-3
Selected General Assessments for Mental Function

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Adaptability/ Work-Related Personal Interactions	World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0)	WHO, 2010	None	Self-report; 36 items and past 30 days	
	University of California, San Diego, Performance-Based Skills Assessment (UPSA)	Mausbach et al., 2008; Patterson et al., 2001	Trained rater	Paper-and-pencil administered	
	Occupational Functioning Scale	Hannula et al., 2006	Mental health worker (minimal qualifications)	Observer-rated	
	Social and Occupational Functioning Assessment Scale (SOFAS)	Rybarczyk, 2011	Trained rater	Clinician-rated (0–100), similar to Global Assessment of Functioning (GAF)	
	Mental Illness Research, Education, and Clinical Center (MIRECC) Global Assessment of Functioning (GAF) Social and Occupational Functioning Scales	Niv et al., 2007	Trained rater	Clinician-rated (0–100), similar to the GAF	
	Specific Level of Functioning Scale (SLOF)	Schneider and Struening, 1983	Trained rater	43-item scale; 5-point rating scale (43–215); lower scores reflect worse functioning	

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
5–20 minutes		Need to sign an agreement with World Health Organization	Generic	
30 minutes	Excellent interrater reliability and criterion validity.	Proprietary	Psychosis, mood disorder, healthy aging, dementia; cognitive and functional impairments related to medical disorders	Available in Spanish.
	Interrater reliability and criterion validity.			English and Finnish versions.
	Reliable and valid in schizophrenia.			
45 minutes	Reliable and valid in schizophrenia.	In public domain	Psychiatric illness	3 subscales: occupational, social, symptoms; all tested in schizophrenia.
60 minutes	Reliable and valid in schizophrenia.	In public domain	Schizophrenia	

ANNEX TABLE 6-4
Selected Psychological Assessments

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered
General Cognitive or Intellectual Ability	Montreal Cognitive Assessment (screen)	Nasreddine et al., 2005	Minimal qualifications	Paper-and-pencil test
General Cognitive or Intellectual Ability; Learning and Memory; Attention and Vigilance	Short Orientation-Memory-Concentration Test of Cognitive Impairment (OMCT)	Katzman et al., 1983		
General Cognitive or Intellectual Ability	Brief Assessment of Cognition (BAC)	Keefe et al., 2004, 2008	Bachelor-level education recommended; some training required	Paper-and-pencil test
General Cognitive or Intellectual Ability; Learning and Memory; Attention and Vigilance	Cognitive Capacity Screening Examination (CCSE)	Jacobs et al., 1977		
Learning and Memory; Attention and Vigilance	National Institutes of Health (NIH) Toolbox Cognition Battery: episodic memory, executive function and attention, working memory, language, processing speed, immediate recall	NU, 2018a	"Approval is granted to researchers and clinicians with knowledge of how to use neuropsychological tests"	Performance tests

Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
10 minutes	Valid and reliable for mild cognitive impairment (MCI).	Public domain		
5–10 minutes	Valid for MCI and dementia; reliability unknown.	Public domain		
30 minutes	Valid and reliable for mental illness; healthy controls; aging.	For purchase from Neurocog Trials		
	Established reliability and validity.	Public domain		
30 minutes	Excellent, normed on U.S. general population.	Yes	Generic	Available in English, Spanish, and other languages; assesses mental processes required to learn, thinking, knowing, remembering, judging, and problem solving; brief, reliable, valid, general population norms; requires iPad, license, annual fee, training.

continued

ANNEX TABLE 6-4
Continued

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Adaptability; Work-Related Personal Interactions	WD-FAB Behavioral Health: self-efficacy, mood and emotions, behavioral control, social interactions	Marfeo et al., 2013b	None	Self-report	
	Personal and Social Performance Scale	Morosini et al., 2000	Mental health worker (minimum qualifications)	Record review	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
	~15 minutes	Excellent; developed with large sample of Social Security Disability Insurance (SSDI) claimants.	No	Generic; individuals with self-reported mental disabilities	
	10 minutes	Reliable and valid in schizophrenia.	Public domain; requires a license agreement		

ANNEX TABLE 6-5
Selected Measures of Disorder Severity and Work-Related Functional Impairment

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Obsessive-Compulsive Disorder					
Attention and Vigilance	Yale-Brown Obsessive Compulsive Scale (Y-BOCS)	Goodman et al., 1989	Advanced graduate-level training in administration and interpretation of psychodiagnostic assessment instruments	Clinician-administered	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
	60 minutes	From Rapp et al. (2016): good to fair internal consistency (α = 0.78–0.89). Good short-term test-retest reliability (r = 0.8869). Good convergent validity: total severity score correlates with clinician-rated measures of obsessive compulsive-disorder (OCD) severity (r = 0.75–0.79). Good discriminant validity: moderate correlations with measures of worry (r = 0.44–0.48).	Available free online	Clinical and nonclinical samples	There are other versions of this scale, including a self-report measure and one for children.

continued

ANNEX TABLE 6-5
Continued

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Attention and Vigilance (*continued*)	Obsessive-Compulsive Inventory-Revised	Huppert et al., 2007	None	Self-report (18 items)	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
	10 minutes	From Rapp et al. (2016): good internal consistency (α = 0.81–0.88). Good to adequate test-retest reliability (r = 0.70–0.84). Good to fair convergent validity: total score correlates with clinician-rated measures of OCD severity (r = 0.41–0.66). Fair to poor discriminant validity: moderate-to-large correlations with depression (r = 0.39–0.70), anxiety (r = 0.47), and worry (r = 0.42).	Available free online	Clinical and nonclinical samples	

continued

ANNEX TABLE 6-5
Continued

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Posttraumatic Stress Disorder					
Attention and Vigilance	PTSD Checklist for DSM-5 (PCL)	Weathers et al., 2013b	Self-report, but should be interpreted by a clinician	Self-report (20 items), though needs to be administered with a brief assessment of Criterion A for posttraumatic stress disorder (PTSD) (trauma exposure)	
	Clinician-Administered PTSD Scale for DSM-5 (CAPS-5)	Weathers et al., 2013a	Advanced graduate-level training in administration and interpretation of psychodiagnostic assessment instruments	Structured interview	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
	5–10 minutes	From Bovin et al. (2015): good internal consistency (0.96), test-retest reliability (*r* = 0.84), and convergent and discriminant validity in sample of veterans.	Public domain	Any population, though psychometric properties tested in veteran populations	20-item measure looking at PTSD symptom severity; often used with Life Events Checklist, which includes traumatic events an individual may have experienced (Criterion A).
	45–60 minutes	From ptsd.va.gov: "The CAPS is the gold standard in PTSD assessment"; strong interrater reliability = 0.78 to 1.00 and test-retest reliability = 0.83.	Created by staff at the U.S. Department of Veterans Affairs (VA) National Center for PTSD; to obtain this scale, must complete online request form; can access it online without using the form, but the VA suggests using the form to verify qualifications to administer	Any population, though psychometric properties tested in veteran populations	

continued

ANNEX TABLE 6-5
Continued

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Depression and Anxiety Disorders					
Attention and Vigilance	Patient Health Questionnaire-9 (PHQ-9)	Kroenke et al., 2001; Mosbach et al., 2018	Self-report, but should be interpreted by a clinician	Self-report (9 items)	
	Hopkins Symptom Check List-20 items (SCL-20)	Derogatis et al., 1974	None	Self-report (20 items)	
	Hopkins Symptom Check List-90 items (SCL-90)	Derogatis et al., 1973	None	Self-report (90 items)	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
	3 minutes	From Kroenke et al. (2001): good internal reliability: (α = 0.89 in PHQ Primary Care study and 0.86 in PHQ Ob-Gyn Study); good test-retest reliability. Correlations between PHQ-9 completed in clinic and administered by phone was 0.84. Construct validity: strong association between increasing PHQ-9 severity scores and worsening function on the SF-20 Health Survey.	Freely available online	Clinical and nonclinical (used to make diagnoses of depression and look at severity)	
	Not given	McKnight and Kashdan (2009): internal consistency (α = 0.92).	Can be obtained by purchasing SCL-90-R at https://www.pearsonclinical.com/psychology/products/100000645/symptom-checklist-90-revised-scl90r.html (accessed April 11, 2019)	Clinical and nonclinical samples	
	12–15 minutes	McKnight and Kashdan (2009): internal consistency (α = 0.86); test-retest reliability (r = 0.81).	Available for purchase at https://www.pearsonclinical.com/psychology/products/100000645/symptom-checklist-90-revised-scl90r.html (accessed April 11, 2019)	Clinical and nonclinical samples	

continued

ANNEX TABLE 6-5
Continued

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Attention and Vigilance (*continued*)	Hamilton Depression Rating Scale (HAM-D)	Hamilton, 1960	Must be administered by a clinician	Clinician-administered (17 items)	
	State-Trait-Anxiety Inventory (STAI)	Spielberger, 1983	None	Self-report	

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
	15–20 minutes	McKnight and Kashdan (2009): Internal consistency (α = 0.89). From Bagby et al. (2004): internal reliability ranged from 0.46 to 0.97. Interrater reliability: Pearson's *r* ranged from 0.82 to 0.98. Retest reliability ranged from 0.81 to 0.98.	Freely available online	Should be used only with patients already diagnosed with a depressive affective disorder	
	10–20 minutes	Summary from the American Psychological Association: "Internal consistency coefficients for the scale have ranged from .86 to .95; test-retest reliability coefficients have ranged from .65 to .75 over a 2-month interval (Spielberger, 1983). Test-retest coefficients for this measure in the present study ranged from .69 to .89. Considerable evidence attests to the construct and concurrent validity of the scale (Spielberger, 1989)."	Can be obtained from the publisher, Mind Garden, 855 Oak Grove Avenue, Suite 215, Menlo Park, CA 94025 (http://www.mindgarden.com/index.htm [accessed April 11, 2019])	Research and clinical populations	

continued

ANNEX TABLE 6-5
Continued

Mental Functional Domain	Functional Assessment Tool	References	Qualifications to Administer	How It Is Administered	
Attention and Vigilance (*continued*)	Beck Anxiety Inventory (BAI)	Beck et al., 1988	None for self-administration; some training for verbal administration, but no set qualifications	Self-administered or verbally by a trained administrator (21 items)	

NOTE: α = Cronbach's alpha.

	Time to Administer	Psychometric Properties (reliability, validity)	Proprietary	Population to Which It Applies	Comments
	5–10 minutes	High internal consistency (α = 0.94). Good test-retest reliability (1 week) (0.75). Validity: moderate convergent validity (0.51), and mild divergent validity with Hamilton Depression Scale (0.25) (Beck et al., 1988).	Pearson (will also provide scoring and reporting when the measure is purchased)	Research and clinical populations	

7

Selected Impairments and Limitations in Functional Abilities Relevant to Work

The committee's Statement of Task (see Box 1-1 in Chapter 1) includes "in the context of disability assessment, describ[ing] the spectrum of changes to functional abilities relevant to work requirements related to the progression of common disease processes in example impairments," which "could include, but are not limited to back disorders, cardiac impairments, or depression." Chapter 7 addresses this charge for several common conditions. These illustrations of disease trajectory, treatment, and related disability are provided as though each condition is present in isolation. Yet more commonly, these pure trajectories are modified by their intersections with each other, by aging, and by comorbidities that influence the impact of disease even when not themselves a major cause of limitation.

For each condition addressed, the committee was asked to

- Identify where along the spectrum an individual's ability to perform functions relevant to work requirements is affected;
- Describe whether SSA [U.S. Social Security Administration] could expect improvement, no improvement, or progressive worsening in the example impairments;
- Describe the efficacy of medications and other treatments on an individual's ability to perform functional abilities relevant to work requirements for these examples, and whether that treatment causes its own subset of medical and/or psychological problems that negatively affect an individual's functioning and how SSA could request an appropriate assessment of functional changes;

- Describe when significant changes in functional abilities relevant to work requirements may occur through the aging process for these examples, such as for adults with common age-related physical and mental impairments; and
- Describe how the examples are similar to or different from other impairments. (See Box 1-1 in Chapter 1.)

In addition to back disorders, cardiac impairments, and depression, the committee elected to address traumatic brain injury (TBI) because of its prevalence and complexity and the associated high rates of cognitive impairment and work disability. In this chapter, the committee discusses for each condition the above five bullet points under three headings: (1) effects on the ability to perform work and expectations for improvement, (2) effects of treatment, and (3) effects of aging and comorbidities. This discussion is followed by a summary and comparison of the trajectories of the four conditions, the effects of treatment, and the effects of aging and comorbidities. The chapter ends with findings and conclusions.

BACK DISORDERS

Effects on the Ability to Perform Work and Expectations for Improvement

As the number of persons reporting disability among adults in the United States continues to rise (CDC, 2009), musculoskeletal conditions such as chronic back pain are among the most common associated conditions (U.S. Burden of Disease Collaborators, 2018). Only ischemic heart disease, lung cancer, chronic obstructive pulmonary disease, and diabetes mellitus surpass chronic low back pain with respect to the number of years lived with a disability (U.S. Burden of Disease Collaborators, 2018). Adults in the United States with chronic back pain are more likely to be socioeconomically disadvantaged, to be covered by government-sponsored health insurance, and to have more frequent health care visits (Shmagel et al., 2016). Early predictors for work-related disability for adults with chronic back pain include injury severity, recent prior job-related injury, and length of time off work (Turner et al., 2008).

Most individuals with back pain improve substantially over the course of a few weeks (Pengel et al., 2003). Gurcay and colleagues (2009) evaluated 91 participants with acute back pain using multiple instruments assessing severity of pain, specific and general health, and depression. At 2 weeks, 57 percent had fully recovered, and only 9 percent went on to develop chronic back pain (Gurcay et al., 2009).

The initial presentation to a health care provider may determine the short-term outcome for individuals with acute back pain. Positive prognostic factors for persons with chronic nonspecific low back pain (CNSLBP) include younger age, lack of comorbid clinical conditions, and a lower level or shorter duration of pain intensity at initial assessment. Verkerk and colleagues (2013) evaluated 1,760 individuals with CNSLBP (defined as back pain without a specific etiology, such as radiculopathy, infection, or trauma) who received multidisciplinary therapy. Participants were evaluated for their course of disability and prognostic factors at baseline and at 2-, 5-, and 12-month follow-up evaluations. At the 12-month follow-up evaluation, prognostic factors for recovery included younger age, shorter duration of pain complaint at baseline evaluation, and lack of underlying mental health comorbidity.

Menezes and colleagues (2012) conducted a meta-analysis of inception cohort studies to evaluate the prognosis for acute and chronic back pain. They demonstrated that most individuals with back pain were able to realize substantial improvement in approximately 6 weeks, with more gradual improvement over the next 10 to 11 months. Measures of pain intensity and disability indicate gradual improvement in low back pain. Steffens and colleagues (2014) evaluated 118 consecutive individuals with CNSLBP who were enrolled in a group exercise program that also incorporated cognitive-behavioral therapy. Outcome measures of pain intensity (based on visual analogue scale pain scores of 0–10) and disability (based on the Roland Morris Disability Questionnaire) were made at 12 months. These measures showed that pain intensity had decreased by 39 percent, and disability and function had improved by 60 and 72 percent, respectively.

Although the majority of patients with a new diagnosis of back pain improve substantially, some go on to experience chronic back pain, which can lead to substantial functional incapacity and disability. In one study, approximately 500 participants were evaluated in a primary care setting. Multiple potential predictive factors in relation to low back pain were evaluated (demographic, physical, psychological, and occupational), and participants were again evaluated at 6 months and 5 years. Baseline pain intensity and the individuals' belief that their back pain would be long-standing were the factors most associated with poor outcomes (Campbell et al., 2013). Other factors found to be associated with a poor prognosis for recovery from back pain include older age, sciatica, functional disability, poor general health, increased psychological stress, negative cognitive characteristics, poor colleague relations, heavy physical work demands, and the presence of compensation (Hayden et al., 2010). Heymans and colleagues (2010) found that having no clinically relevant change in pain intensity and disability status and a high level of pain after the first 3 months were strongly related to developing chronic back pain. In a more recent study, Steenstra and colleagues (2017) evaluated prognostic factors for returning to work after 12 weeks of sick leave among individuals with subacute and chronic back pain. Such factors as age, functional status, delay in treatment referral, receipt of workers' compensation, and attorney involvement were negatively associated with returning to work.

Persistent back pain is clearly associated with developing chronic functional disability. Mehling and colleagues (2012) initially evaluated 605 individuals who consulted their primary care providers about acute low back pain. The participants had no history of back pain in the prior 12 months and no history of spine surgery. Follow-up telephone interviews regarding current symptoms and work status were conducted at 6 months and 2 years. The McGill Pain Questionnaire and Roland-Morris Disability Questionnaire were used to assess the participants. The prevalence of chronic back pain and disability was higher than would be expected based on the published medical literature, with 19 percent of the 443 participants who remained in the study reporting chronic back pain and disability. According to Chou and Shekelle (2010), negative predictors for developing disabling chronic back pain are maladaptive pain coping behaviors, nonorganic signs (physical findings without an underlying physiologic or anatomic etiology), functional impairment, poor general health status, and the presence of psychiatric comorbidities.

Effects of Treatment

Treatment for back pain depends on the individual's clinical presentation, associated findings on physical examination, and pathology as demonstrated by radiological and/or electrodiagnostic tests. Treatment varies from noninvasive approaches such as nonsteroidal anti-inflammatory medications and physical therapy to more aggressive treatments, such as spine injection procedures and surgical intervention (Patrick et al., 2014). Initial treatment for individuals with acute back pain (without neurological deficits or acute radiographic pathology) is typically noninvasive, consisting primarily of nonpharmacologic treatments such as exercise; general physical rehabilitation; and such modalities as acupuncture, progressive relaxation, and biofeedback. Nonpharmacologic treatments for chronic low back pain, including exercise, yoga, psychological therapies, and acupuncture, have demonstrated clinical effectiveness (Chou et al., 2017). Multidisciplinary rehabilitation treatment for persons with low back pain has been shown to be particularly effective (Stein and Miclescu, 2017), with demonstrated improvements for pain management, functional restoration, and quality-of-life measures (Moradi et al., 2012; Morone et al., 2011).

Nonsteroidal anti-inflammatory and muscle relaxant medications are a first-line choice for pharmacologic management (Chung et al., 2013; Qaseem et al., 2017). The use of opioid analgesic medications for the treatment of acute and chronic back pain remains controversial (Deyo et al., 2015). Recently, prescription of opioid analgesic medications in hospital emergency departments has been discouraged (Lee et al., 2016). Indeed, Krebs and colleagues report that treatment with opioids for individuals with back pain was not superior to treatment with nonopioid medications for improving pain-related function (Krebs et al., 2018), and individuals with a higher rate of prescribed opioid medications are more likely to have associated depression (Smith et al., 2017).

Spine injection procedures utilizing corticosteroids are usually reserved for individuals whose clinical presentation is consistent with radiculopathy (Benoist et al., 2012). However, the benefits of this treatment are often limited and short term (Choi et al., 2013; Chou et al., 2015). Epidural steroid injections may increase function and reduce the need for surgical intervention for individuals with back pain (Bicket et al., 2015; Choi et al., 2016). Surgical intervention is reserved for individuals with structural spinal pathology and associated clinical presentations. It is more likely to be effective if an individual's symptoms are associated with a structural abnormality, such as spondylolisthesis, stenosis, or disk herniation, and surgical intervention for radiographically demonstrated intervertebral disc herniation, degenerative spondylolisthesis, and spinal stenosis has been shown to be clinically appropriate, with demonstrated efficacy (Abraham et al., 2016), as well as superior to nonsurgical medical management for appropriately selected individuals (Jacobs et al., 2013; Parker et al., 2014).

Multiple factors can determine when and/or whether an individual will return to work after spine surgery for back pain. Huysmans and colleagues (2018a) conducted a systematic review of the medical literature to identify the factors that influence the duration of an individual's absence from and eventual return to work. They found that important factors diminishing the likelihood of returning to work include older age, female gender, longer duration of preoperative symptoms, and shorter time at preoperative employment. They also found that symptoms of depression after surgery prolong the time before resuming work. Anderson and colleagues (2015) confirmed these predictors, specifically emphasizing clinical depression as a strong predictor of both the outcome of lumbar spine surgery and eventual return to work. Lee and colleagues (2017) reviewed individuals receiving workers' compensation prior to lumbar spine surgery, and found that patients who had been working at least 3 months prior to their surgery were more likely to resume working within 1 year. Other factors, such as lower education, more physically demanding work, and low income, were also negative predictors of return to work after lumbar spine surgery (Truszczyńska et al., 2013). Although many patients do return to work, prediction of postsurgical disability is limited even with extensive evaluation. A detailed model using 39 clinical assessment variables and 38 clinical questionnaire items, for example, explained only half of the observed variation in postsurgical outcomes (McGirt et al., 2015).

Return to work after an acute episode of low back pain that resolves is generally quite straightforward. Indeed, the individual may never have been removed from work, depending on the severity of symptoms, and likely symptoms resolved with conservative therapy or none at all. However, in cases where back pain is persistent and the patient is removed from the work environment, there are opportunities for intervention beyond the pharmacologic, nonpharmacologic, and surgical interventions discussed above. Physical therapy as a modality toward strengthening the core and other muscle groups, as well as reducing pain and increasing flexibility, has met with success (Shipton, 2018).

In the work arena, work hardening, also known as work conditioning, is a further extension of physical therapy and is often utilized in the work setting. Work conditioning is usually employed when the injured worker remains unable to work because of deconditioning and/or functional deficits that remain despite physical therapy and other noninvasive therapy. Work conditioning involves simulation of the work environment and job tasks in which the individual will be engaged. This is carried out in a highly structured environment with the end goal of functional restoration and return to work (Schonstein et al., 2003). The degree of physicality of the job and the age of the worker are among the factors that can influence the ability to return to work (Huysmans et al., 2018b).

The biopsychosocial approach has also been employed in this arena, based on the tenet that the relationship between pain and disability is not predictive and that an individual's reaction to injury is mediated to some extent by psychosocial factors (Schultz et al., 2000, 2007). Cognitive-behavioral therapy has also been used to address some of the other factors affecting return to work, including fear avoidance and pain catastrophizing (Besen et al., 2015).

Returning an individual to work in some capacity before full recovery in the work environment has been found to be associated with more rapid improvement in acute low back pain and functional recovery (Shaw et al., 2018). If modified duty is available, it can be used to return the individual to work early with resultant reduced number of days of disability and earlier return to function both at home and at work. A Cochrane database of systematic reviews by Schonstein and colleagues (2003) found evidence that physical conditioning programs that included cognitive-behavioral therapy coupled with intensive physical training by a physiotherapist or a multidisciplinary team encompassing aerobic capacity, muscle strength and endurance, and coordination was associated with reduced number of sick days due to chronic low back pain.

Effects of Aging and Comorbidities

Low back pain is the most common musculoskeletal condition affecting older adults. Approximately 36 percent of individuals aged 65 and older living in the community experience an episode of back pain each year (Cayea et al., 2006), which in the majority of cases is thought to be of mechanical or soft tissue etiology (Weiner et al., 2006). As previously noted, most cases of low back pain resolve over several weeks. However, low back pain tends to be more persistent in older adults. Cassidy and colleagues (2005) conducted a survey by mail to estimate the prevalence of severity-graded low back pain in the general adult population and reported the results after a 1-year follow-up. They found that most episodes of low back pain were mild and tended to resolve in a matter of weeks, but older individuals were more likely to have persistent low back pain and only partial resolution of their symptoms.

The prevalence of chronic low back pain increases with age, with the highest likelihood in the fifth and sixth decades of life (Shmagel et al., 2016). However, the intensity of low back pain has a stronger correlation with disability in younger adults (Houde et al., 2016). In the study by Houde and colleagues (2016), pain measurements were obtained during the initial physician evaluation using a visual analog scale (0 = no pain; 10 = worst pain imaginable) for individuals who presented to a spine center for treatment. Disability also was assessed during the initial visit using the Oswestry Disability Index (ODI). The authors found a significant, positive association between reported pain intensity as measured on the visual analog scale and disability as measured with the ODI for both age groups. The correlation was stronger in the younger ($r = 0.66$; $p < 0.01$) than in the older ($r = 0.44$; $p < 0.01$) group (Fisher $Z = 2.03$; $p < 0.05$), and the linear regression model showed the slope of the relationship to be steeper in the younger group ($p < 0.05$) (Houde et al., 2016).

CARDIAC IMPAIRMENTS

Effects on the Ability to Perform Work and Expectations for Improvement

Heart Failure

The term "heart failure" refers to any primary heart muscle disease (cardiomyopathy) or secondary impairment of heart function that can arise from multiple causes, most commonly coronary artery disease. To encourage awareness, the current American College of Cardiology/American Heart Association staging system for heart failure defines an asymptomatic Stage B that can be detected as decreased heart function, usually left ventricular ejection fraction seen on echocardiography or nuclear imaging, but has not yet resulted in clinical symptoms (Yancy et al., 2013). This extension of the term "heart failure" has enhanced recognition and treatment of early disease to prevent progression, but can create unnecessarily negative expectations about the imminence and inevitability of disability and death (Cleland et al., 2017; Stevenson, 2017).

Individuals may occasionally be diagnosed in the asymptomatic stage (Stage B) when decreased cardiac function is recognized incidentally during general screening, during evaluation of a family with known genetic disease, or during evaluation of an acute symptom such as chest pain or palpitations. At this early stage, it would be uncommon for usual activity to be limited, except perhaps severe sustained exertion, such as at the level of competitive sports. Occasionally, asymptomatic cardiomyopathy associated with a risk of sudden life-threatening arrhythmias may warrant prohibition of sudden strenuous exertion or employment requiring commercial vehicles, machinery, or working at heights (Banning and Ng, 2012). Otherwise, Stage B heart failure would not be expected to limit employment. Progression of disease from Stage B is uncertain and unpredictable, with fewer than 5 percent of cases progressing to major symptomatic limitation during the next 4 years (Young et al., 2017).

At the time of first heart failure diagnosis, many individuals experience shortness of breath occurring at rest or upon modest exertion (Class III or IV). These symptoms often result from an accumulation of excess body fluid and improve dramatically within days of instituting diuretic therapy, usually during hospitalization (Kato et al., 2012). Improvement of the underlying cardiac condition over the next few months can represent spontaneous resolution of a reversible cardiomyopathy, such as a viral-triggered inflammation of the heart or the cardiomyopathy associated with pregnancy. Improvement also can occur with correction of a structural abnormality, such as heart valve disease, or a functional abnormality, such as a persistently rapid heart rhythm. Even without such correctable causes, establishment of the recommended regimen of medications and pacing devices can improve heart function and functional capacity in some individuals even after years of heart failure symptoms, with about 15 percent improving to near-normal levels of function (Cleland et al., 2017).

Most issues regarding employability with cardiomyopathy or heart failure are likely to arise in Stage C, which encompasses individuals who have ever had symptoms of heart failure but whose symptoms are not yet considered to be refractory. It is in this stage that most individuals are first diagnosed with heart failure and that much of the impact of modern therapies has been achieved. Within Stage C, symptom severity varies markedly, as described by the New York Heart Association's (NYHA's) classification ranging from Class I, with no limitation on activity, to Class IV, with limitation on any physical movement (see Figure 7-1).

Early in the course of symptomatic heart disease, at the level of NYHA Class II symptoms only with "more than usual activity," individuals would generally be unable to sustain strenuous physical labor (e.g., moving furniture or leading exercise classes), associated on exercise testing with 7 or more metabolic equivalents (METS) of energy expenditure (see Figure 7-1 and Chapter 5). Current SSA listing criteria for cardiac functional impairment for both heart failure and ischemic heart disease can be met with inability to perform 5 METS of activity (IOM, 2010, pp. 86, 118). The energy

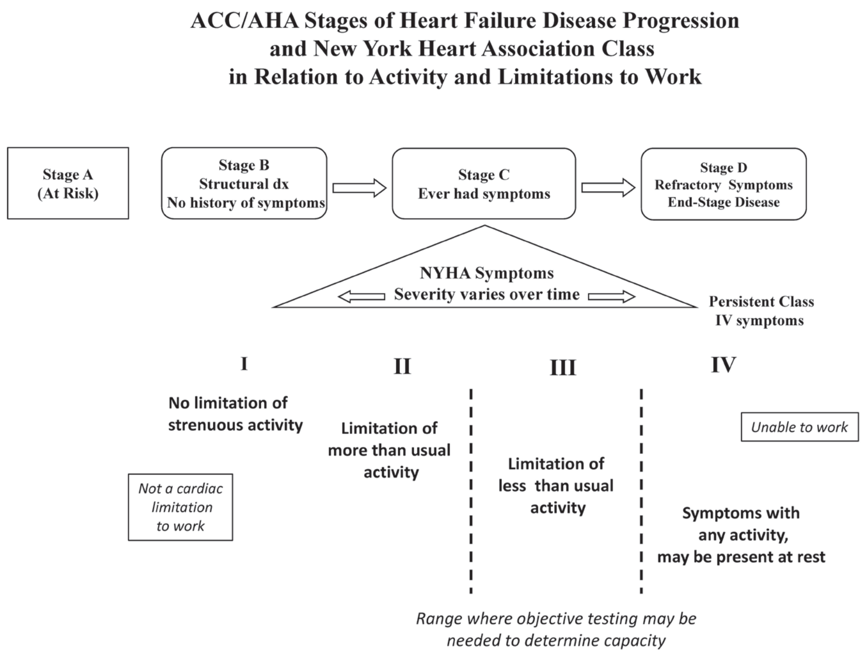

FIGURE 7-1 Alignment of different staging systems to describe heart failure. NOTES: Stages A to D describe patients in terms of disease progression from no history of symptoms through refractory symptoms, with the general implication that most patients move only from left to right. In contrast, the New York Heart Association's (NYHA's) symptom classification (Roman numerals) is dynamic, such that patients frequently shift from no limitation to severe limitation and back to mild limitation, usually as the result of changes in therapy. NYHA class is defined in terms of severity of activity limitation. The range between Classes II and III is where objective testing may be needed to determine actual functional capacity, as an individual in Class I would not be impaired on a cardiac basis, and Class IV symptoms would preclude usual employment. ACC = American College of Cardiology; AHA = American Heart Association; dx = diagnosis.

level of 5 METS is the estimated requirement for doing, for example, light work in a bakery or small carpentry at a bench. This leaves only a narrow, uncertain margin within which typical cardiac conditions would limit an individual's ability to perform common indoor occupations without reaching the SSA listing criteria for disability.

Commonly cited is the statistic that 50 percent of individuals hospitalized with heart failure will die within 5 years, but national data in the United States are dominated by Medicare populations, in which the average age of individuals with heart failure is more than 75. By contrast, in the nationwide Danish registry of 21,455 individuals aged 18 to 60 at the time of their first hospitalization for heart failure, 55 percent were working, and 68 percent of them had returned to the workforce within 12 months (Rørth et al., 2016). Return to work was more common with younger age, male gender, and higher education, but it was substantially less common in the presence of comorbidities, particularly kidney or lung disease (Rørth et al., 2016). In general, the more severe the clinical symptoms and impairment of cardiac function at presentation, the less likely it is that substantial improvement will occur. A recent large prospective 3-year study following Stage C heart failure typical of individuals under age 70 demonstrated that most individuals remained stable without progressive deterioration of symptoms during the 3 years. Each year, about 5 percent died without prior change in their disease stage, and 5 percent deteriorated to Stage D (Kalogeropouolos et al., 2017), the end stage of heart failure.

Individuals with Class IV symptoms that no longer improve with therapy but have become refractory are considered to have progressed to Stage D, with expected survival of less than 1 year (Yancy et al., 2013). Highly selected patients with good self-care and without comorbidities may be appropriate for consideration of cardiac transplantation, durable mechanical circulatory support devices, palliative care, home intravenous inotropic therapy, or hospice care (Yancy et al., 2013). As contemporary therapies for heart failure have slowed the progression of heart failure, they have also markedly reduced unexpected sudden deaths prior to deterioration to Stage D disease, such that the prevalence and duration of severely symptomatic heart failure continue to increase (Udelson and Stevenson, 2016).

Coronary Artery Disease

Symptoms of angina during exercise or other stress arise most commonly from narrowing of the coronary arteries due to atherosclerosis. Most individuals with potentially limiting symptoms resulting from coronary artery disease undergo procedural interventions to enlarge significantly narrowed coronary arteries or surgery to bypass them, assuming they have no comorbidities that limit eligibility for major cardiac surgery.

Evaluation of functional capacity usually occurs following initial appropriate interventions.

Many individuals with coronary artery disease present first with myocardial infarction (heart attack). A previous myocardial infarction that was small or that was effectively treated early is a common cause for Stage B heart failure. Even before the modern era of effective early intervention, the majority of individuals returned to work within 3 months following myocardial infarction (Froelicher et al., 1994). Nonetheless, while coronary artery disease can lead to functional impairment due to angina or myocardial infarction or after cardiac surgery, much of the disability resulting from coronary artery disease relates to a progressive decrease in heart function after myocardial infarction, which is the most common cause for heart failure (Cahill and Kharbanda, 2017).

Of interest, a strong predictor of return to work was an individual's prediction of return to work made before coronary artery bypass surgery or in the hospital early after myocardial infarction (Shanfield, 1990; Sivarajan and Newton, 1984). Indeed, the capacity to continue or return to work with disease does not correlate well with the measured severity of the disease (Sullivan et al., 1997). For patients at multiple stages of coronary artery disease, depression is a key factor predicting impairment and disability (Papakostas, 2009). Depression is one of the most intensely studied predictors of outcome, present in about one of every five patients hospitalized with myocardial infarction, most of whom continue to have depression months later (Bush et al., 2005). Depression after myocardial infarction is associated with poorer quality of life and increased risk of death, with a less consistent relationship being demonstrated with nonfatal cardiac events.

For patients with coronary artery disease, concern may arise regarding the effect of perceived stress in the work environment. Work stress has been associated with a 10 to 40 percent increase in the incidence of a first heart attack, but work stress is difficult to isolate as patients likely to perceive work stress may have other factors that increase their cardiac risk. This association has not been considered sufficient to warrant any targeted recommendations for patients at risk for cardiac disease (Kivimäki and Kawachi, 2015). Patients who perceived major job stress were less likely to return to work after a myocardial infarction, but it was not possible to isolate that effect as it was inversely correlated with job satisfaction. A large prospective study continued over 2 years demonstrated that perception of high stress at the time of initial return to work after a heart attack did not have a major impact unless stress was very high. However, the maintenance of stress over time was associated with almost twice the risk of recurrent heart attack, and almost eight-fold higher risk in patients who also had major impairment in cardiac function after the initial heart attack (Aboa-Éboulé et al., 2007).

Effects of Treatment

Current assessment of physical function and survival with heart failure and coronary artery disease reflects widespread use of medications, devices, and procedures recommended for these diagnoses (Fihn et al., 2014; Yancy et al., 2013). Many medications prescribed for heart failure and coronary artery disease overlap, particularly for individuals with hypertension. Individuals with recent myocardial infarction receive therapies to delay or prevent progression to heart failure. As most of these therapies improve disease prognosis and many help decrease the symptoms of heart failure, their use is expected to improve rather than limit the ability to work. Medications may occasionally cause a transient drop in blood pressure for a few seconds upon standing after prolonged sitting, which may impair the ability to move rapidly at short notice. Individuals with implantable pacemakers and/or defibrillators in their shoulder area will have some restriction on vigorous pulling and stretching with their arm on that side. The most common treatment requiring work adaptation is the use of diuretics by individuals with fluid retention in heart failure. The appropriate timing of diuretic administration may require individuals to use the bathroom several times during the morning, and the rapidity of drug action may require that they be located close to bathroom facilities during these periods. In terms of the conceptual framework described in Chapter 2 (see Figure 2-3), these factors associated with the use of diuretics would be considered interrupters, because they interfere with the ability to perform sustained work activities on a regular and continuing basis.

Return to work after revascularization procedures and cardiac surgery is less frequent than would be expected given the favorable impact of these treatments on reducing cardiac symptoms. Following revascularization with either surgery or catheter-based procedures as performed before 2000, almost one-third of individuals did not return to work, and about one in five rated their health as fair or poor (BARI, 1997), this despite the relief of angina in more than 70 percent of individuals having undergone these procedures to improve coronary artery blood flow (BARI, 2007). In another report on coronary artery bypass surgery, 90 percent of individuals described symptom relief, but only 50 percent returned to work. Of those who did not return to work, only 30 to 40 percent gave “heart problems” as the reason. The likelihood of returning to work increased with higher socioeconomic status, more education, and “looking forward to returning to work” (Mital et al., 2004).

Cardiac rehabilitation is routinely recommended for patients after myocardial infarction or coronary artery bypass surgery and more recently approved for patients with heart failure. Despite the clear benefits to improve survival and quality of life, fewer than one-third of eligible patients participate in cardiac rehabilitation programs (Jolliffe et al., 2001). Some of the limitations are insurance coverage and access to nearby facilities. Exercise capacity and quality of life have frequently improved in large studies. Return to work has occasionally been documented in the United States (Williams et al., 2006), but many patients are already past retirement age at the time they acquire a cardiac indication for rehabilitation. Rates of return to work are more often reported for studies outside the United States, and are higher with earlier enrollment in rehabilitation (Samkange-Zeeb et al., 2006). As with other populations, return to work is more common with younger age and higher patient expectations to return to work.

Effects of Aging and Comorbidities

In considering the impact of cardiac conditions on exercise capacity, it should be noted that the expected "normal" levels of exercise capacity decline substantially with older age. For instance, an exercise performance of 7 METS, needed for heavy carpentry, would be 50 percent of normal for a 40-year-old man and 90 percent of normal for a 70-year-old man (Froelicher et al., 1992). Most cardiovascular exercise impairments are described in terms of proportional decrease such that the absolute diminution is greater from a higher predicted normal value. However, little work has been done to compare cardiac impairment in exercise capacity across age groups.

In general, the limitation imposed by the severity of the cardiac condition outweighs the contribution from chronologic age. However, age is associated with increasing comorbidities, which substantially increase the physical limitation and likelihood of disability due to heart disease (Forman et al., 2018). Atrial fibrillation is a cardiac comorbidity that increases dramatically with age, present in about 0.1 percent of patients aged 40 to 50 and 1.8 percent of those aged 60 to 70, commonly aggravating symptoms of both heart failure and coronary artery disease (Feinberg et al., 1995). Kidney disease and pulmonary disease are the most common medical comorbidities diminishing physical function in people with cardiac disease. The increasing prevalence of diabetes and obesity in older age is anticipated to further limit functional capacity and prognosis for patients with heart failure (Christiansen et al., 2017).

A study following patients after cardiac catheterization showed that physical function as assessed by functional status questionnaires converged to the same level by 6 months regardless of the number of diseased vessels, but remained strongly correlated separately with depression and anxiety (Sullivan et al., 1997). Depression occurs in 20 to 40 percent of patients with heart failure and in a similar proportion of individuals with coronary artery disease (Bush et al., 2005; Mbakwem et al., 2016). The combination is correlated with increased mortality for both conditions. Intervention to relieve depression has been more effective after myocardial infarction than for heart failure, but has not had demonstrable benefit for survival. Unlike most comorbidities, however, concomitant depression with heart failure may be reported less commonly by older individuals than by those who are younger (Gottleib et al., 2004).

DEPRESSION

Effects on the Ability to Perform Work and Expectations for Improvement

Extensive evidence indicates that depression negatively affects functioning relevant to work (Adler et al., 2006; Kessler et al., 2003b; Lerner et al., 2004; Mosbach et al., 2018). In understanding where along the spectrum of depression severity an individual's ability to perform functions relevant to work requirements is affected, it is important to consider how depression can affect job performance. Symptoms of depression include not only sadness but also symptoms that directly reduce functioning, such as inability to attend and concentrate, psychomotor agitation or retardation, and fatigue or loss of energy (Appelbaum, 2018). Preoccupation with death or suicide, guilt, and inability to experience pleasure also may reduce work functioning (Appelbaum, 2018).

Contributors/Correlates of Depression-Related Disability

Given the relatively high prevalence of depressive disorders, it is important to note that the vast majority of individuals experiencing a depressive disorder do not develop severe work-related disability that results in an inability to work. It is thus important to understand and identify the critical factors that contribute to depression-related disability. To this end, one can ask about (1) factors predisposing to depression-related functional impairment, (2) factors related to the depression itself that increase vulnerability to functional impairment, (3) important co-occurring disorders, and (4) factors related to the nature of the job that interact with depression-related work disability.

Predisposing factors A variety of sociodemographic factors are associated with work-related disability among individuals who are depressed. Ervasti and colleagues (2013) found that work-related disability from depression was associated with lower level of education, lower position in job, and renting versus owning a home (Mosbach et al., 2018). Elinson and colleagues (2004) found that among individuals with depression, those aged 18 to 24 were more likely to work than those aged 55 to 69; men and white individuals were more likely to work than women and black individuals; and those with higher levels of education were more likely to work than those with lower levels. The Finnish Vantaa Depression Study (VDS) examined factors related to social and occupational disability, social adjustment, and work disability among individuals with major depressive disorder (MDD) (Rytsälä et al., 2005). Severity and recurrence of depression were the most important factors associated with level of social, functional, and work disability; older age and current Axis I and II comorbidity contributed significantly as well. The study also found that older age, greater sense of hopelessness, worse social and occupational functioning, and persistence of depression predicted receipt of a disability pension (Rytsälä et al., 2007). Sorvaniemi and colleagues (2003) conducted a retrospective study of 213 adult psychiatric outpatients with first-time documented major depression as defined in the *Diagnostic and Statistical Manual of Mental Disorders*, *Third Edition Revised* (DSM-III-R) and found that older age, comorbidity, and lowered self-esteem were strongly associated with being granted a pension. Taken together, these findings indicate that older age is the most consistently observed factor contributing to work-related disability in the context of depression.

Depression-related factors Not surprisingly, overall severity of depression has been associated with loss and impairment of productivity. According to Lerner (2008), severity of depression symptoms has repeatedly been shown to account for some of the variation in work outcomes. A study of the employed baseline sample in the Depression Improvement Across Minnesota: Offering a New Direction (DIAMOND) study ($N = 771$) found that for every increase of 1 point on the Patient Health Questionnaire-9 (PHQ-9), productivity as measured by the Work Productivity and Activity Impairment Questionnaire declined by an additional 1.65 percent (Beck et al., 2011). Consistent with that result, individuals with MDD are most at risk for disability days due to depression, although minor depression can also lead to disability days (Broadhead et al., 1990). As noted above, the VDS found that severity and recurrence of depression were critical predictors of occupational disability and that greater sense of hopelessness and persistence of depression predicted receipt of a disability pension (Rytsälä et al., 2005, 2007). McKnight and Kashdan (2009) conducted an extensive review of the correlation between clinical measures of depression among individuals with MDD and overall functional impairment. They found moderate to strong correlations across clinical samples of all ages. Occupational impairment (Aikens et al., 2008; Hannula et al., 2006; Hirschfield et al., 2002) showed similarly robust correlations with depression severity. Importantly, the association between depression and occupational impairment was found to be greater from assessments obtained during treatment ($b = 0.22$; $t = 3.27$, $p < 0.008$) and posttreatment ($b = 0.33$; $t = 4.16$, $p < 0.002$) than from those obtained at baseline (McKnight and Kashdan, 2009).

Previous research has examined whether specific symptoms of depression negatively affect the degree of work impairment. Studies have found low energy or fatigue, psychomotor disturbance, and low interest or pleasure (Sanderson et al., 2007); difficulty concentrating or being fidgety, feeling tired, or sleep disturbance (Lerner et al., 2004); and memory problems, anxiety, and irritability (Lam, 2012) to be associated with depression-related work outcomes.

Depression severity appears to be the most important clinical depression-related factor associated with disability (Lerner, 2008). It is also clear, however, that levels of depression severity that fail to meet the threshold for a diagnosis can be associated with marked role impairment (Wells et al., 1989).

Co-occurring disorders An emerging body of research documents the combined effects of mental health disorders such as depression and physical health disorders on work-related disability (Kessler and Frank, 1997; Rytsälä et al., 2005; Wells et al., 1989). Buist-Bouwman and colleagues (2005) analyzed data from the Netherlands Mental Health Survey and Incidence Study and found that all physical disorders, except injury caused by accident, were significantly related to anxiety and mood disorders. Both physical and mental disorders were significantly related to work loss; notably, the physical–mental comorbidity was largely additive except for chronic back pain and hypertension, which interacted with mental disorders synergistically. Other studies have documented the association of comorbidity upon receipt of disability pensions (Rytsälä et al., 2007; Sorvaniemi et al., 2003).

Nature of the job In addition to factors relating to the individual, the effects of depression vary by job type. For example, Lerner and colleagues (2004) studied factors related to productivity among employed individuals with depression. They found that reduced productivity was most influenced by depression severity ($p < 0.01$ in 5/5 models); however, deficits increased when employees had occupations requiring proficiency in decision making and communication and/or frequent customer contact ($p < 0.05$ in 3/5 models).

Measuring Depression and Work Impairment

Understanding of where along the spectrum of depression severity an individual's ability to work is affected is limited by the fact that most accepted clinical measures of depression do not assess an individual's functional capacity (Harvey et al., 2017). As is true for all mental disorders, the DSM-5 diagnosis of both major and persistent depressive disorders requires that the “symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning” (APA, 2013, pp. 95, 98). (Persistent depressive disorder, previously called dysthymia, requires fewer depressive symptoms relative to MDD, and is sometimes termed “minor” depression [APA, 2013].) Therefore, functional impairment is a component of the diagnosis of depression, distinct from clinical symptoms. Understanding the extent to which commonly used depression measures provide relevant information about work-related impairment and whether there are measures of work-related functional ability that are useful for individuals experiencing major depression should provide a starting point for understanding where along the spectrum of depression severity work-related disability may occur.

A review by McKnight and Kashdan (2009) identifies three studies that provide correlations of the Hamilton-Depression Rating Scale (HAM-D) (Hirschfeld et al., 2002), the Symptoms Checklist Depression Scale (SCL)-20 (Aikens et al., 2008), and the SCL-90 (Hannula et al., 2006) with functional measures related to occupational functioning. There does not appear to be a consistent correlation or cutpoint over time for the HAM-D, and insufficient data are available to comment on the SCL-90 and SCL-20. Regarding the PHQ-9, the general standard is that a cutpoint of 15 or above is indicative of moderately severe depression, which generally corresponds with work impairment (Kroenke et al., 2001; Lerner et al., 2004, 2012, 2015). Studies by Adler and colleagues (2006) and Lerner and colleagues (2015) have found that a cutpoint of 13 captures additional people with substantial work impairment from depression.

A number of other scales have been used in a limited way to assess the impact of depression on work-related impairment instead of measuring depression per se. None of these scales has achieved widespread use, and all have limited utility. The Lam Employment Absence and Productivity Scale is a measure of work functioning that has been validated for individuals with MDD. It has been shown to have good internal consistency validity and demonstrated to measure “work productivity and troublesome symptoms” among individuals with depression (Lam et al., 2009). The Depression Prognosis Index (DPI) takes various factors into account to predict depression outcomes. Similarly, the 36-Item Short Form Health Survey (SF-36) takes account of social functioning among other factors, within which work functioning is assessed. The DPI has been found to be a valid measure of depression outcomes, including work functioning as measured by the SF-36 (Rubenstein et al., 2007). The Endicott Work Productivity Scale (EWPS) also has been used to assess work function among depressed individuals. Hellerstein and colleagues (2014) used the EWPS to assess the impact of behavioral activation on work productivity among individuals with depression.

A recent study compared individuals with MDD and healthy controls using the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0), which provides measures of "real-world functioning" relevant to work (Milanovic et al., 2018). The "life activities" subscale of the WHODAS was used to measure functional disability and the "getting along with other people" subscale to measure interpersonal functioning. This study evaluated task performance as well as self-rated competence. The MDD group performed worse than controls on competence tasks, and reported greater functional disability and lower self-perception of competence.

Effects of Treatment

A critical factor in addressing the question of expected improvement in depression is the receipt of appropriate treatment. While data vary, a consistent finding across multiple studies is that fewer than half of people with depression receive treatment, and even fewer receive evidence-based treatment (Stewart et al., 2003). Kessler and colleagues (2003a) estimated that only 21.7 percent of Americans with depression received adequate treatment in the National Comorbidity Study Replication. Therefore, the assumption of treatment in determining the expected course of depression-related impairment is not concordant with the average person's experience.

What happens to untreated depression? Stegenga and colleagues (2012) followed a cohort of individuals with MDD in primary care. They found that 17 percent had a chronic and 40 percent had a fluctuating course, while 43 percent remitted. Individuals with chronic courses had more severe depressive symptoms (mean difference 6.54; 95% confidence interval [CI] 4.38–8.70), somatic symptoms (mean difference 3.31; 95% CI: 1.61–5.02), and greater mental dysfunction (mean difference 10.49; 95% CI: –14.42 to –6.57) at baseline relative to those who remitted from baseline, independent of age, sex, level of education, presence of a chronic disease, and a lifetime history of depression (Stegenga et al., 2012).

Schoenbaum and colleagues (2002) underscore the importance of treatment. They evaluated the effects of depression treatment in primary care on individuals' clinical status and employment over 6 months. Data were drawn from a randomized controlled trial of quality improvement for depression that included 938 adults with depressive disorder in 46 managed primary care clinics in five states. At 6 months, individuals with appropriate care had lower rates of depressive disorder (24 versus 70 percent), better mental health–related quality of life, and higher rates of employment (72 versus 53 percent) (each $p < 0.05$) compared with those without such care.

At the same time, symptomatic improvement does not lead automatically to functional improvement. Lerner and colleagues (2015) emphasize the value of high-quality depression care in reducing unemployment, absenteeism, and presenteeism (i.e., at-work performance deficits) (Rost et al., 2004, 2005; Schoenbaum et al., 2002). However, they note that treating depression symptoms does not always restore the ability to work. Studies have shown that even after improvement in depressive symptoms, residual functional limitations persist, including limitations in performance of work activities (Buist-Bouwman et al., 2004; Hirschfeld et al., 2002; Judd et al., 2000; Lerner et al., 2011). The large Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial, for example, found that individuals who responded to first-line antidepressants had reduced work-related impairments. At the same time, however, individuals who did not respond to first-line antidepressants had persistent work-related impairment even when second-line antidepressants eventually reduced their symptoms (Trivedi et al., 2013). Furthermore, other factors, such as chronic general medical and mental comorbidities, influence work-related impairment (Adler et al., 2006; Lerner et al., 2010, 2012).

Adler and colleagues (2006) found further evidence of the persistent impact of depression severity on job performance among employed primary care patients. The study groups included 286 individuals with DSM-IV MDD and/or dysthymia, 93 individuals with rheumatoid arthritis, and 193 depression-free healthy control subjects. Measurements form the Work Limitations Questionnaire showed the depression group had significantly greater deficits in managing mental-interpersonal, time, and output tasks. The rheumatoid arthritis group had significant deficits in managing physical job demands relative to other comparison groups. Although symptom severity helped to predict improvements in job performance, the "job performance of even the clinically improved subset of depressed individuals remained consistently worse than that of the control groups" (Adler et al., 2006).

An additional question regarding the effects of depression treatment is the extent to which treatment improves cognition, which may improve work-related impairment. A large European study found that cognitive symptoms were problematic in depressed persons but did see some improvement with treatment, albeit to a lesser degree than was the case for depressive symptoms (Hammer-Helmich et al., 2018). This large 2-year European observational study of 1,159 outpatients treated with antidepressants found immediate improvement in depressive symptoms and functional impairment following initiation or switch of antidepressant monotherapy, followed by more gradual improvement and long-term stabilization. Improvements in cognitive symptoms were less marked during the acute treatment phase. Functional impairment in individuals with MDD was found not only to be associated with severity of depressive symptoms but also to be independently associated with subjective cognitive symptoms after adjustment for depression severity throughout the 2 years of follow-up. The authors emphasize the importance of cognitive impairment as a limiting factor in recovery of functioning from depression. As discussed further below, Drake and colleagues (2013) found that supported employment services are associated with improved work functioning in people with major depression. Such services can help compensate for the impairing effects of symptoms on the ability to function at work and improve quality of life (McGurk et al., 2018). A recent large randomized controlled trial considered the question of whether cognitive problems improve with antidepressant treatment (Shilyansky et al., 2016). The international Study to Predict Optimized Treatment in Depression (iSPOT-D) assessed the effects of acute antidepressant treatment among 1,008 outpatients, across clinical remission outcomes, on a range of cognitive domains. Impairment in the five domains of attention, response inhibition, verbal memory, decision speed, and information processing speed showed "no relative improvement with acute treatment, irrespective of antidepressant treatment group, even in individuals whose depression remitted acutely according to clinical measures" (Shilyansky et al., 2016, p. 2). In addition, while broader cognitive impairment was associated with greater illness chronicity, it was not associated with symptom severity or previous antidepressant failures. Timing and potential lag as well as extent of improvement are important factors in the impact of depression treatment on functional outcomes. According to McKnight and Kashdan (2009), functioning tends to be less responsive than symptoms to treatment, and therefore functional improvement may take longer than symptom improvement.

In the case of MDD without comorbidity, it appears that improvement can be expected with treatment, although the extent of the improvement and whether it would allow an individual to return to work are uncertain. Persistent symptoms, particularly cognitive symptoms, may impact work-related functioning even in cases of improvement or remission in depressive symptoms. The limited available evidence suggests that the impact of comorbid illnesses is likely to be marked. Evidence suggests that comorbid conditions will increase both the likelihood of depression-related disability and the extent of the work impairment, although the effect in most cases is additive. Without treatment for those conditions, reduction of overall work-related disability is unlikely.

In cases of comorbidity, it is critical to consider the relationship between the comorbid condition and depression. To the extent that depression is exacerbated by or due to another condition, treatment of the second (or third) condition is likely to reduce the severity of depressive symptoms. Notably, the National Comorbidity Study found that almost 80 percent of individuals with 12-month Composite International Diagnostic Interview (CIDI) MDD had comorbid CIDI DSM mental disorders, with MDD only rarely being primary (Kessler et al., 2003a). A recent study of the costs of MDD for treated employed individuals in the United States found that only 40 percent of the incremental costs was due to MDD. The remainder was due to the costs of treatment for other mental health diagnoses (9 percent), as well as non–mental health conditions (29 percent) and associated prescription costs (13 percent) (Greenberg et al., 2015).

Individuals with treatment-resistant depression (including those with persistent depressive disorders) are a key example of how the associated cognitive impairments (i.e., concentration, attention), as well as the physical/somatic manifestations (i.e., disturbed sleep, fatigue, lack of energy), impair individuals' work (and social) functioning, in some cases leading to disability claims. Instruments such as the Work Limitations Questionnaire can be used to assess the specific functional work impairments, and treatments that address these deficits exist (Lerner et al., 2004).

Regarding side effects of antidepressants, Lam (2012) found that individuals reported the medication side effects interfering most with work functioning to be daytime sedation, insomnia, headache, and agitation/anxiety.

As noted above, there is evidence of benefits of vocational rehabilitation, and in particular, supported employment, the only evidence-based practice for improving the capacity for competitive work in people with serious mental illnesses (Drake et al., 2012), including major depression (Drake et al., 2013). Principles of supported employment emphasize rapid deployment of services for the job search, a focus on competitive work, job support services that are consistent with client preferences, integrated mental health and employment services, and time-unlimited follow-along supports. Supported employment produces superior competitive work outcomes as compared with every other form of vocational services. Longer-term outcomes of supported employment do not include routine disengagement from receipt of Social Security Disability Insurance as full-time work is rarely achieved, including in people with severe mental illness, with fear of loss of benefits being a contributor to the predominant part-time work outcomes of supported employment.

Effects of Aging and Comorbidities

Advanced age appears to be a relatively consistent correlate of disability related to depression. This relationship may be due to the additive effects of each episode on work functioning rather than the aging process and its functional decrements, and is independent of other measured factors. As individuals with depressive disorders age, those who have certain chronic comorbid conditions (particularly pain, insomnia, musculoskeletal disorders, and diabetes) find that their depressive symptoms further impair their capacity to function. The key issue is whether the association of increased impairment with depression as people age is due to comorbidities rather than age per se.

MDD occurs more frequently among individuals with chronic medical conditions than among those in the general population (Kang et al., 2015). According to Kessler and colleagues (2005), the depression rate for individuals with medical disorders is approximately two to three times greater than the depression rate for those in the general population. In a 1-year prevalence study conducted by the World Health Organization (WHO), which included 245,400 individuals from 60 countries, 9.3 to 18 percent of those with a single physical disorder experienced depression. Among those with two or more medical conditions, 23 percent experienced depression, compared with just 3.2 percent of medically well individuals (WHO, 2008). According to the Centers for Disease Control and Prevention (CDC, 2012), individuals with the following chronic health disorders experience MDD at higher rates than those in the general population: Alzheimer's disease (11 percent), cardiovascular disease (17 percent), cerebrovascular disease (23 percent), diabetes (27 percent), cancer (42 percent), and Parkinson's disease (51 percent). Strokes are also highly correlated with depression, which is experienced by approximately 31 percent of stroke survivors (Hackett and Pickles, 2014).

While depression often co-occurs with physical health conditions, depression and other mental health disorders also co-occur at high rates. According to Kessler and colleagues (2003a), data from a large-scale national survey revealed that 72.1 percent of individuals who met criteria for lifetime MDD also met criteria for another mental health disorder. The most common of these comorbidities were anxiety disorder (59.2 percent), substance use disorder (24 percent), and impulse control disorder (30 percent).

TRAUMATIC BRAIN INJURY

Traumatic brain injury (TBI) is a leading cause of death and disability worldwide, with estimates of 1.6–3.8 million new cases per year in the United States alone (Shames et al., 2007, as cited in Scaratti et al., 2017). TBI is defined as damage to the brain caused by external mechanical force to the head resulting, for example, from falls, being struck by or against objects, motor vehicle accidents, assaults, sports-related concussions, and blast injuries (Little et al., 2015). Returning to the same occupation where the brain injury occurred may be unsafe. This most commonly occurs in military veterans and first responders, and in such cases, consideration should be given to skill development for a different occupation (Little et al., 2015). The risk of TBI is age dependent, as is the impact on work functioning (Little et al., 2015). For children aged 0–4, TBI is typically due to falls and abuse;

for adolescents, the risk is associated with the legal age of driving; and occupationally related TBI is associated with professional drivers, construction workers, first responders, professional athletes, and military personnel, with the presence of extreme stress at the time of injury diminishing injury recovery in first responders and military personnel (Little et al., 2015). The most prevalent group experiencing TBI is children and youth from infancy through adolescence as the result of falls, abuse, and accidents. TBI is approximately three times more common in men than in women. Sports- and occupationally related TBI is associated with multiple lifetime TBIs. With the exception of blast injuries, TBIs are typically due to acceleration–deceleration, with additional risk of rotational components associated with optic nerve and hippocampal damage and cerebral vasculature bleeds.

TBI severity is classified as mild (less than 30 minutes loss of consciousness [LOC], 13–15 Glasgow Coma Score [GCS], less than 24 hours posttraumatic amnesia [PTA]); moderate (30 minutes–24 hours LOC, 9–12 GCS, 24 hours–7 days PTA); or severe (greater than 24 hours LOC, 3–8 GCS, more than 7 days PTA). TBI severity classification based on these acute TBI variables is only grossly predictive of outcome. In general, persons sustaining milder injuries achieve better recovery and less long-term impairment relative to those with more severe injuries.

Mild TBI, or concussion, is the most common type of TBI, accounting for 70–90 percent of all cases (CDC, 2007). These estimates are imprecise, however, because mild TBI cases often are not documented or consistently diagnosed (Cancelliere et al., 2014).

Effects on the Ability to Perform Work and Expectations for Improvement

TBI is associated with a broad range of motor, sensory, emotional, and cognitive impairment. Symptoms include fatigue, irritability, impulse control problems, self-centeredness, headaches, dizziness, sleep disturbance, problems with balance and coordination, anxiety, and depression, as well as impairment in memory, attention, information processing speed, planning, and problem solving (Little et al., 2015). Most individuals with mild TBI will recover within 1 year, but 5–20 percent of individuals may have persistent problems (Losoi et al., 2016; Rao et al., 2010). Symptoms can resolve in months, but in some cases, sequelae can last years. The course and prognosis of TBI are complicated by multiple trauma, such as damage to the spinal cord, orthopedic injuries, and chest trauma. Cognitive impairment may resolve for people with mild TBI within 3 months, but up to 18 percent of those with mild TBI experience persistent symptoms and functional deficits that can last from months to years postinjury, with as many as 18 percent of those who were employed premorbidly remaining unemployed at 12 months post-TBI (Cancelliere et al., 2014). Individuals with psychiatric comorbidities prior to experiencing TBI were found to be at increased risk for relapse and to have lower return-to-work rates following their injury (Garrelfs et al., 2015).

A substantial literature exists on the relationship between TBI and work functioning, but this literature does not yield definitive conclusions about the impact of TBI on work because of the heterogeneity of injuries, definitions of TBI and its degrees of severity, demographic characteristics of the injured individuals, methods used to measure both the injury and employment, lengths of follow-up, treatments received, and systems of care involved in assessment and treatment (Cancelliere et al., 2014). The WHO Collaborating Center Task Force for Mild Traumatic Brain Injury determined that the studies reviewed addressing return-to-work rates in people with TBI were of low scientific quality, and no acceptable studies on return to work addressed physical, cognitive, or emotional barriers (Cancelliere et al., 2014).

A recent systematic review of the literature from 1993 to 2015 addressed specific work-related difficulties associated with TBI, defined as job instability over 5 years postinjury, as well as determinants of these difficulties, focusing primarily on return to work (Scaratti et al., 2017). The authors first conducted a systematic assessment of the quality of studies, judging them to be poor, acceptable, good, or excellent based on National Institute for Health and Care Excellence guidelines, and excluding poor-quality studies from the review. They identified 42 papers meeting their inclusionary criteria, encompassing a total of 25,756 people: 29 percent were female, the average age was 34.9 years, mean time from acute event was 14.6 months, 51.4 percent had severe TBI, 14 percent had moderate TBI, and 33.5 percent had mild TBI, with multitrauma information being available in 9 of the 42 studies. The mean employment rate postinjury was 42.5 percent for all TBI, although information on employment in these studies was quite limited (Scaratti et al., 2017).

Demographic variables that predicted work-related difficulties included older age at the time of injury (older age being defined as 35–54 versus younger than 34); female gender (females being found to have lower workforce participation, including reduced hours or unemployment); lower education (those without a high school diploma having the greatest difficulties compared with those with a high school diploma, and the least difficulty occurring in those with a college degree); preinjury unemployment status; preinjury unmarried status; being a manual worker; and having a minority status. Features of the injury associated with work-related difficulties 5 years postinjury included its severity as indicated by a low GCS and lengthy PTA; a violent cause of the TBI (as compared with a car accident); and polytrauma (such as spinal cord injury). Functional status at discharge had prognostic implications for return to work, including requiring assistance with functioning and lower scores on measures of motor and cognitive functioning (Scaratti et al., 2017).

The importance of cognition is highlighted by studies indicating that employment at 1 year postinjury was 82 percent lower for those with lower cognitive scores, as determined by the Functional Independence Measure. The presence of factors directly or indirectly related to cognitive functioning, such as self-centeredness, inappropriate social behavior, impulsivity, and irritability, predicted poorer work functioning 1 year postinjury (Scaratti et al., 2017). In addition to cognitive impairment, other factors found to be related to lower rates of return to work were behavioral problems such as confusion, agitation, inappropriate behavior at admission, and discharge to rehabilitation. Individuals with personality changes in particular were found to be 10 times less likely relative to those without such changes to return to work at 18 months postinjury. Apathy, depression, agitation, posttraumatic stress disorder, and psychosis also have been identified as common behavioral sequelae of TBI, with apathy and depression occurring in more than 50 percent of cases (Scaratti et al., 2017).

Posttraumatic pain and headache also have been identified as impacting return to work (Little et al., 2015). Headache is reported as the most common symptom after TBI and is one of a constellation of symptoms of postconcussive syndrome, with highest rates in mild TBI. Because headache is self-rated, episodic, and variable in course within and among individuals, its impact on return to work is largely unknown. Chronic posttraumatic headache (CPTH) reportedly resolves in a few months postinjury, with 25 percent of individuals still having symptoms after 4 years. CPTH—which has been explained by insufficient treatment; analgesic rebound; and psychosocial comorbidities such as anxiety, depression, and insomnia—is associated with occupational disability. Among people with CPTH, 12 percent report missed workdays and annual absenteeism of an average of 27 days, with greater reduced activity level, more disability, and lower physical function relative to those with nontraumatic chronic headache (Little et al., 2015).

Effects of Treatment

Returning to work promotes recovery from TBI and is associated with improved well-being, health status, and quality of life (Little et al., 2015). Successful return to work is aided by having health insurance; receipt of supported employment; tailoring of work duties based on preserved abilities; a socially inclusive work environment; and reduced stress at work, such as avoiding frustration associated with having to relearn prior performed tasks, avoiding the need to multitask, and diminishing the need to learn new skills (Little et al., 2015). Reviews consistently mention the need for vocational rehabilitation to help people with TBI return to work. Supported employment is also encouraged given its supports and services for the range of physical, psychological, and cognitive sequelae associated with TBI. A few studies have shown enhanced return-to-work rates in people receiving such services. However, randomized controlled studies are needed to further understand whether and how vocational rehabilitation enhances return to work in people with TBI (Mani et al., 2017; Scaratti et al., 2017).

Effects of Aging and Comorbidities

Individuals who are older at the time of their TBI have worse outcomes (e.g., are less likely to return to work) compared with younger people with similar TBI, with "older" being defined in some studies as ages 35–54 (Scaratti et al., 2017). The physical effects of aging also may make older workers more vulnerable to injury, including head injury (Kristman et al., 2010). In addition, older people are more likely to have progressive age-related cognitive decline, which also contributes to diminished rates of return to work, as well as greater difficulties on the job, following TBI (Marquez de la Plata et al., 2008).

SUMMARY AND COMPARISON OF COMMON IMPAIRMENTS

Trajectories

The trajectories of disease are variable for most conditions that can severely impair ability to work, as in the examples of back pain, depression, cardiovascular disease, and TBI discussed in this chapter. Figure 7-2 illustrates the effects of a variety of factors on the trajectory of ability to work over time.

Both back pain and depression can improve, worsen, or stabilize, but their course can also be characterized by episodes of exacerbation followed by return to a previous plateau. Back pain, depression, and cardiac disorders can improve markedly with appropriate treatment even late after diagnosis. A different course is tracked for TBI, which, like stroke, often improves early to a plateau with little chance of further improvement, and associated limitations often fall into other specific categories, such as disorders of cognition, coordination, and mood.

For cardiac conditions, the severity of deviation from normal capacity can be measured more objectively than for the other conditions discussed in this chapter, as is also the case for pulmonary disorders. Even for cardiopulmonary disease, however, the associated physiologic decrements do not translate directly to work impairment, as multiple other factors affect the ability and the motivation to work with limitations. In addition, the

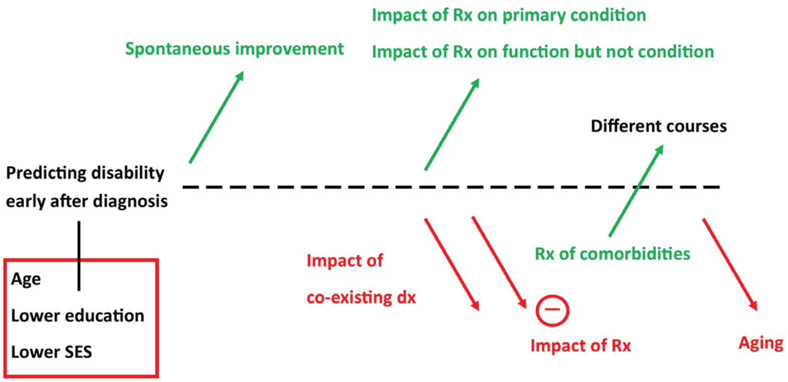

FIGURE 7-2 Illustration of factors with potential positive (green) and negative (red) effects over time on the trajectory of ability to work.
NOTES: Age, lower education, and lower socioeconomic status at the time of diagnosis predict lower likelihood of return to work. After diagnosis, increase in functional capacity can result from spontaneous improvement in the condition, from the impact of therapy on the condition itself, or from the impact of therapy on the ability to adapt to work despite continued presence of the condition. Coexisting diagnoses of comorbidities often limit improvement over time, but treatment of these comorbidities may improve overall function. Aging in general is associated with less improvement over time and greater burden of comorbidities limiting improvement in function. dx = diagnosis; Rx = treatment; SES = socioeconomic status.

multiple physical demands of a work environment are difficult to quantitate, the exception being strictly repetitive physical exertions, which often are not required by the dominant tasks of contemporary employment. Following initiation of therapies, heart failure can also be characterized by episodes of treatable exacerbation, but these usually herald a progressive downhill course. In contrast with the course of musculoskeletal limitations, depression, or most other impairments, the progression of cardiac disease to become the primary limitation to employment is often associated with a decreasing likelihood of surviving the next 2 years.

Impacts of Treatment

Therapies can treat a condition directly or aid adaptation to the work environment without changing the severity of the condition itself. For example, medications can often delay or prevent deterioration of cardiac and peripheral vascular function, while exercise training can improve functional capacity without improving the underlying organ system disease. Likewise, multiple medications can improve the ability to function with major psychiatric diagnoses without necessarily altering the underlying pathology or trajectory of the condition. Diseases differ in the degree to which access to appropriate therapy determines the trajectory of functional capacity. Cardiovascular disease and depression are profoundly influenced by the availability of and compliance with optimal therapy. However, therapies for chronic back pain and for TBI have a limited effect on the condition's trajectory. For conditions, such as cardiac disease and back pain, that may necessitate major surgical procedures, recovery of employability is often influenced by factors other than the primary diagnosis for which the surgery was performed. The efficacy of treatments for comorbidities can also influence the ability to work. In addition, it is important to consider the impact health disparities can have on access to treatment. Factors related to socioeconomic status, race/ethnicity, and language may limit access to treatment, symptom improvement, and thus return to work.

Aging and Comorbidities

Aging affects the physical capacity to repair, recover from, and adapt to most specific conditions that can affect employment. In addition, the age-related decline in reserve capacity for both physical and cognitive tasks decreases the margin for performance of many work-related tasks. Aging is strongly associated with increasing numbers and severity of comorbidities, including physical, mental, and cognitive limitations.

Depression

Depression is the most common comorbidity limiting employment for individuals with a wide range of physical and mental impairments, as well as rehabilitation from such events as surgeries and myocardial infarction. For all of the conditions considered in this chapter, co-occurring depression is frequent and associated with poor outcomes. Of major concern, moreover, it is underrecognized both as a primary diagnosis and as a powerful contributor to impairment from other diagnoses. The impact of treatment is clear for depression as the central diagnosis, but less is known about how to identify and address it as a complicating factor (Anderson et al., 2015; Scarrati et al., 2017; Sullivan et al., 1997).

Other Comorbidities

Although this chapter focuses on several common conditions in isolation, it is increasingly clear that these and other functional limitations often cannot be envisioned as a single impairment, but share common comorbidities that must be considered in evaluation. The landmark Medical

Outcomes Study of ambulatory office visits focused on combinations rather than single diagnoses of physical conditions, including angina, heart failure, myocardial infarction, back problems, gastrointestinal disorders, chronic lung problems, arthritis, and diabetes. Two or more conditions were present in the majority of patients with any one condition, and the number of conditions strongly influenced limitations on physical, role, and social functioning (Stewart et al., 1989). The declining trajectory of physical function in patients with these multiple conditions was documented over a 4-year longitudinal follow-up (Bayliss et al., 2004).

As the prevalence of people with multiple chronic conditions increases, health care delivery is facing new challenges in their evaluation and treatment. A major initiative for 2020 focuses on the need for health services research devoted specifically to "addressing the health care needs of Americans with [multiple chronic conditions] or disabilities" (Iezzoni, 2010, p. 1524).

FINDINGS AND CONCLUSIONS

Findings

7-1. The course of an impairment and where along the trajectory of a condition an individual becomes unable to work vary tremendously and are affected by symptom severity, comorbidities, demographic characteristics, and the job demands associated with the work itself.

7-2. Although the likelihood of return to work for a population is correlated with multiple factors that can be ascertained (as is the case for low back surgery and traumatic brain injury), return to work has not been reliably modeled or predicted with high accuracy.

7-3. Increasing age compounds the effects of the conditions discussed herein, both directly and in association with increasing burden of comorbidities.

7-4. Individuals with a primary cardiac limitation will often meet the U.S. Social Security Administration (SSA) cardiac listings for disability by the time they need to limit their work activities in occupations other than those requiring sustained heavy physical labor.

7-5. Rehabilitation has consistently been shown to improve exercise capacity and quality of life, with frequent benefits for improved mood, but has varying impacts on return to work.

7-6. Individuals with low back pain are more likely to return to work if they are of younger age, have fewer comorbid conditions, and perform less physically demanding work.

7-7. Depression frequently compounds work-related functional limitation in the context of other primary impairments.

7-8. Fewer than half of people with depression receive treatment, and even fewer receive evidence-based treatment.

Conclusions

7-1. Consideration of age and comorbidities is critically important in both the evaluation and the likely trajectory of common conditions and their effects on work-related impairment.

7-2. Depression is particularly important to consider and incorporate into evaluations regarding return to work with a chronic condition that limits physical function.

7-3. Age, presurgical employment, the physical exertion involved in work, and comorbidities are key factors in evaluating individuals' functional abilities relevant to work requirements.

7-4. Evaluation of cardiac limitation to employment may focus primarily on determination of presence of SSA listing criteria for cardiac conditions.

7-5. Rehabilitation offers multiple benefits and is best offered early, but its impact on return to work varies widely.

7-6. Even with extensive assessment in multiple different domains, it is unlikely that multivariate models could be constructed to reliably answer the question of whether an individual will be able to return to work in the future.

7-7. It is advisable to address treatment of comorbidities during treatment of the primary impairment identified, in order to maximize functional capacity and ability to work.

7-8. Given the prevalence and underdetection of depression in general medical care and its impact on functioning, particularly when comorbid with other conditions, it is valuable to screen all disability applicants for depression. Such screening is standard practice in many primary care settings.

REFERENCES

Aboa-Éboulé, C., C. Brisson, E. Maunsell, B. Mâsse, R. Bourbonnais, M. Vézina, A. Milot, P. Theroux, and G. R. Dagenais. 2007. Job strain and risk of acute recurrent coronary heart disease events. *JAMA* 298(14):1652–1660.

Abraham, P., R. C. Rennert, J. R. Martin, J. Ciacci, W. Taylor, D. Resnick, E. Kasper, and C. C. Chen. 2016. The role of surgery for treatment of low back pain: Insights from the randomized controlled Spine Patient Outcomes Research Trials. *Surgical Neurology International* 7:38. http://surgicalneurologyint.com/surgicalint-articles/the-role-of-surgery-for-treatment-of-low-back-pain-insights-from-the-randomized-controlled-spine-patient-outcomes-research-trials (accessed April 4, 2019).

Adler, D. A., T. J. McLaughlin, W. H. Rogers, H. Chang, L. Lapitsky, and D. Lerner. 2006. Job performance deficits due to depression. *American Journal of Psychiatry* 163(9):1569–1576.

Aikens, J., K. Kroenke, D. E. Nease, M. S. Klinkman, and A. Sen. 2008. Trajectories of improvement for six depression-related outcomes. *General Hospital Psychiatry* 30(1):26–31. doi: 10.1016/j.genhosppsych.2007.10.003.

Anderson, J. T., A. R. Haas, R. Percy, S. T. Woods, U. M. Ahn, and N. U. Ahn. 2015. Clinical depression is a strong predictor of poor lumbar fusion outcomes among workers' compensation subjects. *Spine* 40(10):748–756. doi: 10.1097/BRS.0000000000000863.

APA (American Psychiatric Association). 2013. *Diagnostic and statistical manual of mental disorders*, 5th ed. Washington, DC: APA.

Appelbaum, P. S. 2018. *Functional assessment of depression and anxiety disorders relevant to work requirements*. Presentation to the Committee on Functional Assessment for Adults with Disabilities, Washington, DC, April 19. http://nationalacademies.org/hmd/~/media/Files/Agendas/Activity%20Files/SelectPops/Functional%20Assessment/Meeting%203/Functional%20Assessment%20of%20Depression%20%20Anxiety%20%20Appelbaum.pdf (accessed April 4, 2019).

Banning, A. S., and A. Ng. 2012. Driving and arrhythmia: A review of scientific basis for international guidelines. *European Heart Journal* 34(3):236–244. doi.org/10.1093/eurheartj/ehs356.

BARI (Bypass Angioplasty Revascularization Investigation). 1997. Five-year clinical and functional outcome comparing bypass surgery and angioplasty in patients with multivessel coronary disease. Writing group for BARI investigators. *JAMA* 277(9):715–722.

BARI. 2007. Final 10-year follow-up results from the BARI randomized trial. *Journal of the American College of Cardiology* 49(15):1600–1606. doi: 10.1016/j.jacc.2006.11.048.

Bayliss, E. A., M. S. Bayliss, J. E. Ware, and J. F. Steiner. 2004. Predicting declines in physical function in persons with chronic medical conditions. *Health and Quality of Life Outcomes* 2:47–55. doi: 10.1186/1477-7525-2-47.

Beck, A., A. L. Crain, L. I. Solberg, J. Unützer, R. E. Glasgow, M. V. Maciosek, and R. Whitebird. 2011. Severity of depression and magnitude of productivity loss. *Annals of Family Medicine* 9(4):305–311. doi: 10.1370/afm.1260.

Benoist, M., P. Boulu, and G. Hayem. 2012. Epidural steroid injections in the management of low-back pain with radiculopathy: An update of their efficacy and safety. *European Spine Journal* 21(2):204–213.

Besen, E., A. E. Young, and W. S. Shaw. 2015. Returning to work following low back pain: Towards a model of individual psychosocial factors. *Journal of Occupational Rehabilitation* 25(1):25–37.

Bicket, M. C., J. M. Horowitz, H. T. Benzon, and S. P. Cohen. 2015. Epidural injections in prevention of surgery for spinal pain: Systematic review and meta-analysis of randomized controlled trials. *Spine* 15(2):348–362.

Broadhead, W. E., D. G. Blazer, L. K. George, and C. K. Tse. 1990. Depression, disability days, and days lost from work in a prospective epidemiologic survey. *JAMA* 264(19):2524–2528.

Buist-Bouwman, M. A., J. Ormel, R. de Graaf, and W. A. Vollebergh. 2004. Functioning after a major depressive episode: Complete or incomplete recovery? *Journal of Affective Disorders* 82(3):363–371. doi: 10.1016/j.jad.2004.02.007.

Buist-Bouwman, M. A., R. de Graaf, W. A. M. Vollebergh, and J. Ormel. 2005. Comorbidity of physical and mental disorders and the effect on work-loss days. *Acta Psychiatrica Scandinavica* 111(6):436–443.

Bush, D. E., R. C. Ziegelstein, U. V. Patel, B. D. Thombs, D. E. Ford, J. A. Fauerbach, U. D. McCann, K. J. Stewart, K. K. Tsilidis, A. L. Patel, C. J. Feuerstein, and E. B. Bass. 2005. Post-myocardial infarction depression: Summary. *AHRQ Evidence Report Summaries* 123. https://www.ncbi.nlm.nih.gov/books/NBK11913 (accessed April 4, 2019).

Cahill, T. J., and R. K. Kharbanda. 2017. Heart failure after myocardial infarction in the era of primary percutaneous coronary intervention. *World Journal of Cardiology* 9(5):407–415. doi: 10.4330/wjc.v9.i5.407.

Campbell, P., N. E. Foster, E. Thomas, and K. M. Dunn. 2013. Prognostic indicators of low back pain in primary care: Five-year prospective study. *The Journal of Pain* 14(8):873–883. doi: 10.1016/j.jpain.2013.03.013.

Cancelliere, C., V. L. Kristman, J. D. Cassidy, C. A. Hincapié, P. Côté, E. Boyle, L. J. Carroll, B-M. Stålnacke, C. Nygren-de Boussard, and J. Borg. 2014. Systematic review of return to work after mild traumatic brain injury: Results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. *Archive of Physical Medicine and Rehabilitation* 95(3 Suppl. 2):S201–S209.

Cassidy, J. D., P. Côté, L. J. Carroll, and V. Kristman. 2005. Incidence and course of low back pain episodes in the general population. *Spine* 30(24):2817–2823.

Cayea, D., S. Perera, and D. K. Weiner. 2006. Chronic low back pain in older adults: What physicians know, what they think they know, and what they should be taught. *Journal of the American Geriatrics Society* 54(11):1772–1777.

CDC (U.S. Centers for Disease Control and Prevention). 2007. Rates of hospitalization related to traumatic brain injury—Nine states, 2003. *Morbidity and Mortality Weekly Report* 56(8):167–170.

CDC. 2009. Prevalence and most common causes of disability among adults—United States, 2005. *Morbidity and Mortality Weekly Report* 58(16):421–426.

CDC. 2012. *Mental health and chronic diseases*. https://www.cdc.gov/workplacehealthpromotion/tools-resources/pdfs/issue-brief-no-2-mental-health-and-chronic-disease.pdf (accessed April 4, 2019).

Choi, H. J. , S. Hahn, C. H. Kim, B. H. Jang, S. Park, S. M. Lee, J. Y. Park, C. K. Chung, and B. J. Park. 2013. Epidural steroid injection therapy for low back pain: A meta-analysis. *International Journal of Technology Assessment in Health Care* 29(3):244–253.

Choi, J. H., J. Y. Hong, S. W. Suh, J. H. Yang, S. Y. Park, J. H. Park, and S. J. Hong. 2016. What is the role of epidural steroid injections in lumbar spinal disease with moderate disability? *Pain Physician* 19(4):293–298.

Chou, R., and P. Shekelle. 2010. Will this patient develop persistent disabling low back pain? *JAMA* 303(13):1295–1302. doi: 10.1001/jama.2010.344.

Chou, R., R. Hashimoto, J. Friedly, R. Fu, T. Dana, S. Sullivan, C. Bougatsos, and J. Jarvik. 2015. *Pain management injection therapies for low back pain*. Rockville, MD: Agency for Healthcare Research and Quality.

Chou, R., R. Deyo, J. Friedly, A. Skelly, R. Hashimoto, M. Weimer, R. Fu, T. Dana, P. Kraegel, J. Griffin, S. Grusing, and E. D. Brodt. 2017. Nonpharmacologic therapies for low back pain: A systematic review for an American College of Physicians Clinical Practice Guideline. *Annals of Internal Medicine* 166(7):493–505.

Christiansen, M. N., L. Kober, P. Weeke, R. S. Ramachandran, J. L. Jeppesen, J. G. Smith, G. H. Gislason, C. Torp-Pederson, and C. Anderson. 2017. Age specific trends in incidence, mortality, and comorbidities of heart failure in Denmark 1995–2012. *Circulation* 137(25):1214–1223. doi: 10.1161/CIRCULATIONAHA.116.025941.

Chung, J. W., Y. Zeng, and T. K. Wong. 2013. Drug therapy for the treatment of chronic nonspecific low back pain: Systematic review and meta-analysis. *Pain Physician* 16(6):E685–E704.

Cleland, J. G. F., P. Pellicori, A. L. Clark, and M. C. Pettrie. 2017. Time to take the failure out of heart failure. *Journal of the American College of Cardiology* 5(7):538–540. doi: 10.1016/j.jchf.2017.04.003.

Deyo, R. A., M. Von Korff, and D. Duhrkoop. 2015. Opioids for low back pain. *British Medical Journal* 350:g6380. doi: 10.1136/bmj.g6380.

Drake, R. E., G. R. Bond, and D. R. Becker. 2012. *Individual placement and support: An evidence-based approach to supported employment*. New York: Oxford University Press.

Drake, R. E., W. Frey, G. R. Bond, H. H. Goldman, D. Salkever, A. Miller, T. A. Moore, J. Riley, M. Karakus, and R. Milfort. 2013. Assisting Social Security Disability Insurance beneficiaries with schizophrenia, bipolar disorder, or major depression in returning to work. *American Journal of Psychiatry* 170(12):1433–1441.

Elinson, L., P. Houck, S. C. Marcus, and H. A. Pincus. 2004. Depression and the ability to work. *Psychiatric Services* 55(1):29–34.

Ervasti, J., J. Vahtera, J. Pentti, T. Oksanen, K. Ahola, M. Kivimäki, and M. Virtanen. 2013. Depression-related work disability: Socioeconomic inequalities in onset, duration and recurrence. *PLoS ONE* 8(11):e79855. doi: 10.1371/journal.pone.0079855.

Feinberg, W. M., J. L. Blackshear, A. Laupacis, R. Kronmal, and R. G. Hart. 1995. Prevalence, age distribution, and gender of patients with atrial fibrillation. *Archives of Internal Medicine* 155(5):469–473.

Fihn, S. D., J. C. Blankenship, K. P. Alexander, J. A. Bittl, J. G. Byrne, B. J. Fletcher, G. C. Fonarow, R. A. Lange, G. N. Levine, T. M. Maddox, S. S. Naidu, E. M. Ohman, and P. K. Smith. 2014. 2014 ACC/AHA/AATS/PCNA/SCAI/STS focused update of the guideline for the diagnosis and management of patients with stable ischemic heart disease: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines, and the American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *Journal of the American College of Cardiology* 64(18):1929–1949. doi: 10.1016/j.jacc.2014.07.017.

Forman, D. E., M. S. Maurer, C. Boyd, R. Brindis, M. Salive, and M. Rich. 2018. Multimorbidity in older adults with cardiovascular disease. *Journal of the American College of Cardiology* 71(19):2149–2161. doi: 10.1016/j.jacc.2018.03.022.

Froelicher, V. F., C. K. Morris, A. Hideg, T. Kawaguchi, J. Myers, and K. Ueshima. 1992. Nomogram for exercise capacity using METs and age. *American College of Cardiology Learning Highlights* 8:2–5.

Froelicher, E. S., L. L. Kee, M. Newton, B. Lindskog, and M. Livingston. 1994. Return to work, sexual activity, and other activities after acute myocardial infarction. *Heart and Lung* 23(5):423–435.

Garrelfs, S. F., B. H. Donker-Cools, H. Wind, and M. H. Frings-Dresen. 2015. Return-to-work in patients with acquired brain injury and psychiatric disorders as a comorbidity: A systematic review. *Brain Injury* 29(5):550–557.

Gottleib, S. S., M. Khatta, E. Friedmann, L. Einbinder, S. Katzen, B. Baker, J. Marshal, S. Minshall, S. Robinson, M. L. Fisher, M. Potenza, B. Sigler, C. Baldwan, and S. A. Thomas. 2004. The influence of age, gender, and race on the prevalence of depression in heart failure patients. *Journal of the American College of Cardiology* 43(9):1542–1549.

Greenberg, P. E., A. A. Fournier, T. Sisitsky, C. T. Pike, and R. C. Kessler. 2015. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *Journal of Clinical Psychiatry* 76(2):155–162. doi: 10.4088/JCP.14m09298.

Gurcay, E., A. Bal, E. Eksioglu, A. E. Hasturk, A. G. Gurcay, and A. Cakci. 2009. Acute low back pain: Clinical course and prognostic factors. *Disability and Rehabilitation* 31(10):840–845.

Hackett, M. L., and K. Pickles. 2014. Part I: Frequency of depression after stroke: An updated systematic review and meta-analysis of observational studies. *International Journal of Stroke* 9(8):1017–1025. doi: 10.1111/ijs.12357.

Hammer-Helmich, L., J. M. Haro, B. Jönsson, M. A. Tanguy, S. Di Nicola, J. Chollet, D. Milea, B. Rive, and D. Saragoussi. 2018. Functional impairment in patients with major depressive disorder: The 2-year PERFORM study. *Neuropsychiatric Disease and Treatment* 14:239–249. doi: 10.2147/NDT.S146098.

Hannula, J. A., K. Lahtela, A. Järvikoski, J. Salminen, and P. Mäkelä. 2006. Occupational Functioning Scale (OFS)—An instrument for assessment of work ability in psychiatric disorders. *Nordic Journal of Psychiatry* 60(5):372–378. doi: 10.1080/08039480600937140.

Harvey, P. D., E. W. Twamley, A. E. Pinkham, C. A. Depp, and T. L. Patterson. 2017. Depression in schizophrenia: Associations with cognition, functional capacity, everyday functioning, and self-assessment. *Schizophrenia Bulletin* 43(3):575–582.

Hayden, J. A., K. M. Dunn, D. A. van der Windt, and W. S. Shaw. 2010. What is the prognosis of back pain? *Best Practice & Research: Clinical Rheumatology* 24(2):167–179. doi: 10.1016/j.berh.2009.12.005.

Hellerstein, D. J., G. Erickson, J. W. Stewart, P. J. McGrath, K. Hunnicutt-Ferguson, S. K. Reynolds, D. O'Shea, Y. Chen, A. Withers, and Y. Wang. 2014. Behavioral activation therapy for return to work in medication-responsive chronic depression with persistent psychosocial dysfunction. *Comprehensive Psychiatry* 57:140–147. doi: 10.1016/j.comppsych.2014.10.015.

Heymans, M. W., S. van Buuren, D. L. Knol, J. R. Anema, W. van Mechelen, and H. C. de Vet. 2010. The prognosis of chronic low back pain is determined by changes in pain and disability in the initial period. *Spine* 10(10):847–856. doi: 10.1016/j.spinee.2010.06.005.

Hirschfeld, R. M., D. L. Dunner, G. Keitner, D. N. Klein, L. M. Koran, S. G. Kornstein, J. C. Markowitz, I. Miller, C. B. Nemeroff, P. T. Ninan, A. J. Rush, A. F. Schatzberg, M. E. Thase, M. H. Trivedi, F. E. Borian, P. Crits-Christoph, and M. B. Keller. 2002. Does psychosocial functioning improve independent of depressive symptoms? A comparison of nefazodone, psychotherapy, and their combination. *Biological Psychiatry* 51(2):123–133. doi: 10.1016/S0006-3223(01)01291-4.

Houde, F., F. Cabana, and G. Léonard. 2016. Does age affect the relationship between pain and disability? A descriptive study in individuals suffering from chronic low back pain. *Journal of Geriatric Physical Therapy* 39(3):140–145. doi: 10.1519/JPT.0000000000000055.

Huysmans, E., L. Goudman, G. Van Belleghem, M. De Jaeger, M. Moens, J. Nijs, K. Ickmans, R. Buyl, C. Vanroelen, and K. Putman. 2018a. Return to work following surgery for lumbar radiculopathy: A systematic review. *Spine* 18(9):1694–1714. doi: 10.1016/j.spinee.2018.05.030.

Huysmans, E., K. Ickmans, D. Van Dyck, J. Nijs, Y. Gidron, N. Roussel, A. Polli, M. Moens, L. Goudman, and M. De Kooning. 2018b. Association between symptoms of central sensitization and cognitive behavioral factors in people with chronic nonspecific low back pain: A cross-sectional study. *Journal of Manipulative and Physiological Therapeutics* 41(2):92–101.

Iezzoni, L. I. 2010. Multiple chronic conditions and disabilities: Implications for health services research and data demands. *Health Services Research* 45(5):1523–1540.

IOM (Institute of Medicine). 2010. *Cardiovascular disability*. Washington, DC: The National Academies Press.

Jacobs, W. C., S. M. Rubinstein, B. Koes, M. W. van Tulder, and W. C. Peul. 2013. Evidence for surgery in degenerative lumbar spine disorders. *Best Practice & Research: Clinical Rheumatology* 27(5):673–684.

Jolliffe, J. A., K. Rees, R. S. Taylor, D. Thompson, N. Oldridge, and S. Ebrahim. 2001. Exercise-based rehabilitation for coronary heart disease. *The Cochrane Database of Systematic Reviews* (1):CD001800. doi: 10.1002/14651858.CD001800.

Judd, L. L., H. S. Akiskal, P. J. Zeller, M. Paulus, A. C. Leon, J. D. Maser, J. Endicott, W. Coryell, J. L. Kunovac, T. I. Mueller, J. P. Rice, and M. B. Keller. 2000. Psychosocial disability during the long-term course of unipolar major depressive disorder. *Archives of General Psychiatry* 57(4):375–380.

Kalogeropouolos, A. P., A. Samman-Tahhan, J. S. Hedley, A. A. McCue, J. B. Bjork, D. W. Markham, K. N. Bhatt, V. V. Georgiopoulou, A. L. Smith, and J. Butler. 2017. Progression to Stage D heart failure among outpatients with Stage C heart failure and reduced ejection fraction. *Journal of the American College of Cardiology* 5(7):528–537. doi: 10.1016/j.jchf.2017.02.020.

Kang, H.-J., S.-Y. Kim, K.-Y. Bae, S.-W. Kim, I.-S. Shin, J.-S. Yoon, and J.-M. Kim. 2015. Comorbidity of depression with physical disorders: Research and clinical implications. *Chonnam Medical Journal* 51(1):8–18.

Kato, M., L. W. Stevenson, M. Palardy, P. M. Campbell, C. W. May, N. K. Lakdawala, G. Stewart, A. Nohria, J. G. Rogers, J. T. Heywood, M. Gheorghiade, E. F. Lewis, X. Mi, and S. Setoguchi. 2012. The worst symptom as defined by patients during heart failure hospitalization: Implications for response to therapy. *Journal of Cardiac Failure* 18(7):524–533. doi: 10.1016/j.cardfail.2012.04.012.

Kessler, R. C., and R. G. Frank. 1997. The impact of psychiatric disorders on work loss days. *Psychological Medicine* 27(4):861–873.

Kessler, R. C., P. Berglund, O. Demler, R. Jin, D. Koretz, K. R. Merikangas, A. J. Rush, E. E. Walters, and P. S. Wang. 2003a. The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 289(23):3095–3105.

Kessler, R. C., C. Barber, A. Beck, P. Berglund, P. D. Cleary, D. McKenas, N. Pronk, G. Simon, P. Stang, T. B. Üstün, and P. Wang. 2003b. The World Health Organization Health and Work Performance Questionnaire (HPQ). *Journal of Occupational and Environmental Medicine* 45(2):156–174.

Kessler, R. C., W. T. Chiu, O. Demler, K. R. Merikangas, and E. E. Walters. 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* 62(6):617–627.

Kivimäki, M., and I. Kawachi. 2015. Work stress as a risk factor for cardiovascular disease. *Current Cardiology Reports* 17(9):630.

Krebs, E. E., A. Gravely, S. Nugent, A. C. Jensen, B. DeRonne, E. S. Goldsmith, K. Kroenke, M. J. Bair, and S. Noorbaloochi. 2018. Effect of opioid vs. nonopioid medications on pain-related function in patients with chronic back pain or hip or knee osteoarthritis pain: The SPACE randomized clinical trial. *JAMA* 319(9):872–882.

Kristman, V. L., P. Côté, S. Hogg-Johnson, J. D. Cassidy, D. Van Eerd, M. Vidmar M. Rezai, and R. A. Wennberg. 2010. The burden of work disability associated with mild traumatic brain injury in Ontario compensated workers: A prospective cohort study. *The Open Occupational Health & Safety Journal* 2:1–8. https://benthamopen.com/contents/pdf/TOOHSJ/TOOHSJ-2-1.pdf (accessed April 4, 2019).

Kroenke, K., R. L. Spitzer, and J. B. Williams. 2001. The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine* 16(9):606–613.

Lam, R. W. 2012. Which depressive symptoms and medication side effects are perceived by patients as interfering most with occupational functioning? *Depression Research and Treatment* 2012, Article 630206. doi: 10.1155/2012/630206.

Lam, R. W., E. E. Michalak, and L. N. Yatham. 2009. A new clinical rating scale for work absence and productivity: Validation in patients with major depressive disorder. *BMC Psychiatry* 9:78. doi: 10.1186/1471-244X-9-78.

Lee, S. S., Y. Choi, and G. S. Pransky. 2016. Extent and impact of opioid prescribing for acute occupational low back pain in the emergency department. *Journal of Emergency Medicine* 50(3):376–384. doi: 10.1016/j.jemermed.2015.10.015.

Lee, Y. P., S. D. Farhan, D. Kiester, C. Rosen, A. Pendi, and N. Bhatia. 2017. Variables affecting return to work after spinal surgery in a non-workers' compensation population: A retrospective cohort study. *Journal of the American Academy of Orthopaedic Surgeons* 25(12):e282–e288. doi: 10.5435/JAAOS-D-17-00098.

Lerner, D., D. A. Adler, H. Chang, E. R. Berndt, J. T. Irish, L. Lapitsky, M. Y. Hood, J. Reed, and W. H. Rogers. 2004. The clinical and occupational correlates of work productivity loss among employed patients with depression. *Journal of Occupational and Environmental Medicine* 46(6 Suppl.):S46–S55.

Lerner, D., D. A. Adler, W. H. Rogers, H. Chang, L. Lapitsky, T. McLaughlin, and J. Reed. 2010. Work performance of employees with depression: The impact of work stressors. *American Journal of Health Promotion* 24(3):205–213. doi: 10.4278/ajhp.090313-QUAN-103.

Lerner, D., D. A. Adler, R. C. Hermann, W. H. Rogers, H. Chang, P. Thomas, A. Greenhill, and K. Perch. 2011. Depression and work performance: The work and health initiative study. In *Work accommodation and retention in mental health*, edited by I. Z. Schultz and E. S. Rogers. New York: Springer. Pp. 103–120. doi: 10.1007/978-1-4419-0428-7_6.

Lerner, D., D. A. Adler, R. C. Hermann, H. Chang, E. J. Ludman, A. Greenhill, K. A. Perch, W. C. McPeck, and W. H. Rogers. 2012. Impact of a work-focused intervention on the productivity and symptoms of employees with depression. *Journal of Occupational and Environmental Medicine* 54(2):128–135.

Lerner, D., D. A. Adler, W. H. Rogers, H. Chang, A. Greenhill, E. Cymerman, and F. Azocar. 2015. A randomized clinical trial of a telephone depression intervention to reduce employee presenteeism and absenteeism. *Psychiatric Services* 66(6):570–577.

Lerner, M. 2008. What does research tell us about depression, job performance, and work productivity? *Journal of Occupational and Environmental Medicine* 50(4):401–410.

Little, D. M., A. J. Cook, S. B. Morissette, and J. W. Klocek. 2015. Considerations for return to work following traumatic brain injury. *Handbook of Clinical Neurology* 131:465–479. doi: 10.1016/B978-0-444-62627-1.00027-5.

Losoi, H., N. D. Silverberg, M. Wäljas, S. Turunen, E. Rosti-Otajärvi, M. Helminen, T. M. Luoto, J. Julkunen, J. Öhman, and G. L. Iverson. 2016. Recovery from mild traumatic brain injury in previously healthy adults. *Journal of Neurotrauma* 33(8):766–776. doi: 10.1089/neu.2015.4070.

Mani, K., B. Cater, and A. Hudlikar. 2017. Cognition and return to work after mild/moderate traumatic brain injury: A systematic review. *Work* 58(1):51–62.

Marquez de la Plata, C., T. Hart, F. M. Hammond, A. Frol, A. Hudak, C. R. Harper, T. O'Neil-Pirozzi, J. Whyte, M. Carlile, and R. Diaz-Arrastia. 2008. Impact of age on long-term recovery from traumatic brain injury. *Archives of Physical Medicine and Rehabilitation* 89(5):896–903.

Mbakwem, A., F. Aina, and C. Amadi. 2016. Depression in patients with heart failure: Is enough being done? *Cardiac Failure Review* 2(2):110–112. doi: 10.15420/cfr.2016:21:1.

McGirt, M. J., A. Sivaganesan, A. L. Asher, and C. J. Devin. 2015. Prediction model for outcome after low-back surgery. *Neurosurgical Focus* 39(6):1–10. doi: 10.3171/2015.8.FOCUS15338.

McGurk, S. R., R. E. Drake, H. Xie, J. Riley, R. Milfort, T. W. Hale, and W. Frey. 2018. Cognitive predictors of work among Social Security Disability Insurance beneficiaries with psychiatric disorders enrolled in IPS supported employment. *Schizophrenia Bulletin* 44(1):32–37.

McKnight, P. E., and T. B. Kashdan. 2009. The importance of functional impairment to mental health outcomes: A case for reassessing our goals in depression treatment research. *Clinical Psychology Review* 29(3):243–259.

Mehling, W. E., V. Gopisetty, E. Bartmess, M. Acree, A. Pressman, H. Goldberg, F. M. Hecht, T. Carey, and A. L. Avins. 2012. The prognosis of acute low back pain in primary care in the United States: A 2-year prospective cohort study. *Spine* 37(8):678–684. doi: 10.1097/BRS.0b013e318230ab20.

Menezes Costa, L. C., C. G. Maher, M. J. Hancock, J. H. McAuley, R. D. Herbert, and L. O. P. Costa. 2012. The prognosis of acute and persistent low-back pain: A meta-analysis. *CMAJ: Canadian Medical Association Journal* 184(11):E613–E624. doi: 10.1503/cmaj.111271.

Milanovic, M., K. Holshausen, R. Milev, and C. R. Bowie. 2018. Functional competence in major depressive disorder: Objective performance and subjective perceptions. *Journal of Affective Disorders* 234:1–7. doi: 10.1016/j.jad.2018.02.094.

Mital, A., A. Desai, and A. Mital. 2004. Return to work after a coronary event. *Journal of Cardiopulmonary Rehabilitation and Prevention* 24(6):365–373.

Moradi, B., S. Hagmann, A. Zahlten-Hinguranage, F. Caldeira, C. Putz, N. Rosshirt, E. Schönit, A. Mesrian, M. Schiltenwolf, and E. Neubauer. 2012. Efficacy of multidisciplinary treatment for patients with chronic low back pain: A prospective clinical study in 395 patients. *Journal of Clinical Rheumatology* 18(2):76–82.

Morone, G., T. Paolucci, M. R. Alcuri, M. C. Vulpiani, A. Matano, I. Bureca, S. Paolucci, and V. M. Saraceni. 2011. Quality of life improved by multidisciplinary back school program in patıents with chronic non-specific low back pain: A single blind randomized controlled trial. *European Journal of Physical and Rehabilitation Medicine* 47(4):533–541.

Mosbach, P., S. Campanelli, A. D'Anello, L. Klaus, E. Aberger, J. Pelletier, P. Quinn, V. Sundar, A. Houtenville, B. Ogut, H. Brotsos, and O. Urdapilleta. 2018. Mental cognitive demands. Contract No. 28321318A00040029/28321318FA0011262. Summit Consulting.

Papakostas, G. I. 2009. Major depressive disorder: Psychosocial impairment and key considerations in functional improvement. *American Journal of Managed Care* 15(11 Suppl.):S316–S321.

Parker, S. L., S. S. Godil, S. K. Mendenhall, S. L. Zuckerman, D. N. Shau, and M. J. McGirt. 2014. Two-year comprehensive medical management of degenerative lumbar spine disease (lumbar spondylolisthesis, stenosis, or disc herniation): A value analysis of cost, pain, disability, and quality of life: Clinical article. *Journal of Neurosurgery*: *Spine* 21(2):143–149.

Patrick, N., E. Emanski, and M. A. Knaub. 2014. Acute and chronic low back pain. *Medical Clinics of North America* 98(4):777–789.

Pengel, L. H. M., R. D. Herbert, C. G. Maher, and K. M. Refshauge. 2003. Acute low back pain: Systematic review of its prognosis. *British Medical Journal* 327(7410):323. doi: 10.1136/bmj.327.7410.323.

Qaseem, A., T. J. Wilt, R. M. McLean, and M. A. Forciea. 2017. Noninvasive treatments for acute, subacute, and chronic low back pain: A clinical practice guideline from the American College of Physicians. *Annals of Internal Medicine* 166(7):514–530.

Rao, V., M. Bertrand, P. Rosenberg, M. Makley, D. J. Schretlen, J. Brandt, and M. M. Mielke. 2010. Predictors of new-onset depression after mild traumatic brain injury. *The Journal of Neuropsychiatry and Clinical Neurosciences* 22(1):100–104.

Rørth, R., C. Wong, K. Kragholm, E. L. Fosbøl, U. M. Mogensen, M. Lamberts, M. C. Petrie, P. S. Jhund, T. A. Gerds, C. Torp-Pedersen, G. H. Gislason, J. J. McMurray, L. Køber, and S. L. Kristensen. 2016. Return to the workforce after first hospitalization for heart failure. *Circulation* 134(14):999–1009.

Rost, K., J. L. Smith, and M. Dickinson. 2004. The effect of improving primary care depression management on employee absenteeism and productivity: A randomized trial. *Medical Care* 42(12):1202–1210.

Rost, K., J. Fortney, and J. Coyne. 2005. The relationship of depression treatment quality indicators to employee absenteeism. *Mental Health Services Research* 7(3):161–169.

Rubenstein, L. V., N. R. Rayburn, E. B. Keeler, D. E. Ford, K. M. Rost, and C. D. Sherbourne. 2007. Predicting outcomes of primary care patients with major depression: Development of a depression prognosis index. *Psychiatric Services* 58(8):1049–1056. doi: 10.1176/appi.ps.58.8.1049.

Rytsälä, H. J., T. K. Melartin, U. S. Leskelä, T. P. Sokero, P. S. Lestelä-Mielonen, and E. T. Isometsä. 2005. Functional and work disability in major depressive disorder. *The Journal of Nervous and Mental Disease* 193(3):189–195.

Rytsälä, H. J., T. K. Melartin, U. S. Leskelä, T. P. Sokero, P. S. Lestelä-Mielonen, and E. T. Isometsä. 2007. Predictors of long-term work disability in major depressive disorder: A prospective study. *Acta Psychiatrica Scandinavica* 115(3):206–213.

Samkange-Zeeb, F., T. Altenhöner, G. Berg, and T. Schott. 2006. Predicting non-return to work in patients attending cardiac rehabilitation. *International Journal of Rehabilitation Research* 29(1):43–49.

Sanderson, K., E. Tilse, J. Nicholson, B. Oldenburg, and N. Graves. 2007. Which presenteeism measures are more sensitive to depression and anxiety? *Journal of Affective Disorders* 101(1–3):65–74.

Scaratti, C., M. Leonardi, D. Sattin, S. Schiavolin, M. Willems, and A. Raggi. 2017. Work related difficulties in patients with traumatic brain injury: A systematic review on predictors and associated factors. *Disability and Rehabilitation* 39(9):847–855. doi: 10.3109/09638288.2016.1162854.

Schoenbaum, M., J. Unützer, D. McCaffrey, N. Duan, C. Sherbourne, and K. B Wells. 2002. The effects of primary care depression treatment on patients' clinical status and employment. *Health Services Research* 37(5):1145–1158. doi: 10.1111/1475-6773.01086.

Schonstein, E., D. T. Kenny, J. L. Keating, and B. W. Koes. 2003. Work conditioning, work hardening and functional restoration for workers with back and neck pain. *Cochrane Database of Systematic Reviews* 3:CD001822. doi: 10.1002/14651858.CD001822.

Schultz, I. Z., J. Crook, K. Fraser, and P. W. Joy. 2000. Models of diagnosis and rehabilitation in musculoskeletal pain-related occupational disability. *Journal of Occupational Rehabilitation* 10(4):271–293.

Schultz, I. Z., A. W. Stowell, M. Feuerstein, and R. J. Gatchel. 2007. Models of return to work for musculoskeletal disorders. *Journal of Occupational Rehabilitation* 17(2):327–352.

Shames, J., I. Treger, H. Ring, and S. Giaquinto. 2007. Return to work following TBI: Trends and challenges. *Disability and Rehabilitation* 29(17):1387–1395.

Shanfield, S. B. 1990. Return to work after an acute myocardial infarction: A review. *Heart and Lung* 19(2):109–117.

Shaw, W. S., C. C. Nelson, M. J. Woiszwillo, B. Gaines, and S. E. Peters. 2018. Early return to work has benefits for relief of back pain and functional recovery after controlling for multiple confounds. *Journal of Occupational and Environmental Medicine* 60(10):901–910.

Shilyansky, C., L. M. Williams, A. Gyurak, A. Harris, T. Usherwood, and A. Etkin. 2016. Effect of antidepressant treatment on cognitive impairments associated with depression: A randomised longitudinal study. *The Lancet Psychiatry* 3(5):425–435.

Shipton, E. A. 2018. Physical therapy approaches in the treatment of low back pain. *Pain and Therapy* 7(2):127–137.

Shmagel, A., R. Foley, and H. Ibrahim. 2016. Epidemiology of chronic low back pain in U.S. adults: Data from the 2009–2010 National Health and Nutrition Examination Survey. *Arthritis Care Research (Hoboken)* 68(11):1688–1694.

Sivarajan, E. S., and K. M. Newton. 1984. Exercise, education, and counseling for patients with coronary artery disease. *Clinics in Sports Medicine* 3(2):349–369.

Smith, J. A., R. L. Fuino, I. Pesis-Katz, X. Cai, B. Powers, M. Frazer, and J. D. Markman. 2017. Differences in opioid prescribing in low back pain patients with and without depression: A cross-sectional study of a national sample from the United States. *PAIN Reports* 2(4):e606.

Sorvaniemi, M., H. Helenius, and R. K. Salokangas. 2003. Factors associated with being granted a pension among psychiatric outpatients with major depression. *Journal of Affective Disorders* 75(1):43–48.

Steenstra, I. A., C. Munhall, E. Irvin, N. Oranye, S. Passmore, D. Van Eerd, Q. Mahood, and S. Hogg-Johnson. 2017. Systematic review of prognostic factors for return to work in workers with sub acute and chronic low back pain. *Journal of Occupational Rehabilitation* 27(3):369–381. doi: 10.1007/s10926-016-9666-x.

Steffens, D., M. J. Hancock, C. G. Maher, J. Latimer, R. Satchell, M. Ferreira, P. H. Ferreira, M. Partington, and A. L. Bouvier. 2014. Prognosis of chronic low back pain in patients presenting to a private community-based group exercise program. *European Spine Journal* 23(1):113–119. doi: 10.1007/s00586-013-2846-x.

Stegenga, B. T., M. H. Kamphuis, M. King, I. Nazareth, and M. I. Geerlings. 2012. The natural course and outcome of major depressive disorder in primary care: The PREDICT-NL study. *Social Psychiatry and Psychiatric Epidemiology* 47(1):87–95. doi:10.1007/s00127-010-0317-9.

Stein, K. F., and A. Miclescu. 2017. Effectiveness of multidisciplinary rehabilitation treatment for patients with chronic pain in a primary health care unit. *Scandinavian Journal of Pain* 4(4):190–197.

Stevenson, L. W. 2017. Who would be branded with failure? *Circulation* 136(15):1359–1361. doi: 10.1161/CIRCULATIONAHA.117.029667.

Stewart, A., S. Greenfield, R. D. Hays, K. Wells, W. H. Rogers, S. D. Berry, E. A. McGlynn, and J. E. Ware. 1989. Functional status and well-being of patients with chronic conditions. *JAMA* 262(7):907–913.

Stewart, W. F., J. A. Ricci, E. Chee, S. R. Hahn, and D. Morganstein. 2003. Cost of lost productive work time among U.S. workers with depression. *JAMA* 289(23):3135–3144.

Sullivan, M. D., A. Z. Larois, C. Caum, L. C. Grothaus, and W. J. Katon. 1997. Functional status in coronary artery disease: 1-year prospective study of the role of anxiety and depression. *American Journal of Medicine* 103(5):348–356.

Trivedi, M. H., D. W. Morris, S. R. Wisniewski, I. Lesser, A. A. Nierenberg, E. Daly, B. T. Kurian, B. N. Gaynes, G. K. Balasubramani, and A. J. Rush. 2013. Increase in work productivity of depressed individuals with improvement in depressive symptom severity. *American Journal of Psychiatry* 170(6):633–641. doi: 10.1176/appi.ajp.2012.12020250.

Truszczyńska, A., K. Rapała, O. Truszczyński, A. Tarnowski, and S. Łukawski. 2013. Return to work after spinal stenosis surgery and the patient's quality of life. *International Journal of Occupational Medicine and Environmental Health* 26(3):394–400. doi: 10.2478/s13382-013-0105-3.

Turner, J. A., G. Franklin, D. Fulton-Kehoe, L. Sheppard, B. Stover, R. Wu, J. V. Gluck, and T. M. Wickizer. 2008. Early predictors of chronic work disability: A prospective, population-based study of workers with back injuries. *Spine* 33(25):2809–2818.

Udelson, J. E., and L. W. Stevenson. 2016. Future of heart failure diagnosis, therapy, and management. *Circulation* 133(25):2671–2686.

U.S. Burden of Disease Collaborators, A. H. Mokdad, K. Ballestros, M. Echko, S. Glenn, H. E. Olsen, E. Mullany, A. Lee, A. R. Khan, A. Ahmadi, A. J. Ferrari, A. Kasaeian, A. Werdecker, A. Carter, B. Zipkin, B. Sartorius, B. Serdar, B. L. Sykes, C. Troeger, C. Fitzmaurice, C. D. Rehm, D. Santomauro, D. Kim, D. Colombara, D. C. Schwebel, D. Tsoi, D. Kolte, E. Nsoesie, E. Nichols, E. Oren, F. J. Charlson, G. C. Patton, G. A. Roth, H. D. Hosgood, H. A. Whiteford, H. Kyu, H. E. Erskine, H. Huang, I. Martopullo, J. A. Singh, J. B. Nachega, J. R. Sanabria, K. Abbas, K. Ong, K. Tabb, K. J. Krohn, L. Cornaby, L. Degenhardt, M. Moses, M. Farvid, M. Griswold, M. Criqui, M. Bell, M. Nguyen, M. Wallin, M. Mirarefin, M. Qorbani, M. Younis, N. Fullman, P. Liu, P. Briant, P. Gona, R. Havmoller, R. Leung, R. Kimokoti, S. Bazargan-Hejazi, S. I. Hay, S. Yadgir, S. Biryukov, S. E. Vollset, T. Alam, T. Frank, T. Farid, T. Miller, T. Vos, T. Bärnighausen, T. T. Gebrehiwot, Y. Yano, Z. Al-Aly, A. Mehari, A. Handal, A. Kandel, B. Anderson, B. Biroscak, D. Mozaffarian, E. R. Dorsey, E. L. Ding, E. K. Park, G. Wagner, G. Hu, H. Chen, J. E. Sunshine, J. Khubchandani, J. Leasher, J. Leung, J. Salomon, J. Unutzer, L. Cahill, L. Cooper, M. Horino, M. Brauer, N. Breitborde, P. Hotez, R. Topor-Madry, S. Soneji, S. Stranges, S. James, S. Amrock, S. Jayaraman, T. Patel, T. Akinyemiju, V. Skirbekk, Y. Kinfu, Z. Bhutta, J. B. Jonas, and C. J. L. Murray. 2018. The state of U.S. health, 1990–2016: Burden of diseases, injuries, and risk factors among U.S. states. *JAMA* 319(14):1444–1472.

Verkerk, K., P. A. Luijsterburg, M. W. Heymans, I. Ronchetti, A. L. Pool-Goudzwaard, H. S. Miedema, and B. W. Koes. 2013. Prognosis and course of disability in patients with chronic nonspecific low back pain: A 5- and 12-month follow-up cohort study. *Physical Therapy* 93(12):1603–1614.

Weiner, D. K., S. Sakamoto, S. Perera, and P. Breuer. 2006. Chronic low back pain in older adults: Prevalence, reliability, and validity of physical examination findings. *Journal of the American Geriatrics Society* 54(1):11–20.

Wells, K. B., A. Stewart, R. D. Hays, M. A. Burnam, W. Rogers, M. Daniels, S. Berry, S. Greenfield, and J. Ware. 1989. The functioning and well-being of depressed patients. Results from the Medical Outcomes Study. *JAMA* 262(7):914–919.

WHO (World Health Organization). 2008. *The global burden of disease: 2004 update*. Geneva, Switzerland: WHO Press.

Williams, M. A., P. A. Ades, L. F. Hamm, S. J. Keteyian, T. P. LaFontaine, J. L. Roitman, and R. W. Squires. 2006. Clinical evidence for a health benefit from cardiac rehabilitation: An update. *American Heart Journal* 152(5):835–841.

Yancy, C. W., M. Jessup, B. Bozkurt, J. Butler, D. E. Casey, Jr., M. H. Drazner, G. C. Fonarow, S. A. Geraci, T. Horwich, J. L. Januzzi, M. R. Johnson, E. K. Kasper, W. C. Levy, F. A. Masoudi, P. E. McBride, J. J. V. McMurray, J. E. Mitchell, P. N. Peterson, B. Riegel, F. Sam, L. W. Stevenson, W. H. W. Tang, E. J. Tsai, and B. L. Wilkoff. 2013. 2013 ACCF/AHA Guideline for the management of heart failure. *Journal of the American College of Cardiology* 62(16):e150–e239.

Young, K., C. Scott, R. Rodeheffer, and H. Chen. 2017. Preclinical heart failure: Progression of Stage A and Stage B heart failure in a community population. *European Heart Journal* 38(Suppl. 1):ehx504.P3382. doi: 10.1093/eurheartj/ehx504.P3382.

8

Review of Selected Disability Benefit Programs

As part of the Statement of Task for this study, the U.S. Social Security Administration (SSA) asked the committee to "provide an overview of the functional assessment processes in at least three similar benefit programs that assess disability or vocational capabilities (national and state government programs, private sector programs, and foreign programs as applicable) and provide examples of forms, tools, guides, examinations, and other resources used by benefit programs that assess functional aspects of disability and vocational capabilities." This chapter presents the committee's response to that request.

To meet this objective, in addition to reviewing relevant policy and procedural documents, the committee spoke to representatives of the U.S. Department of Veterans Affairs' (VA's) Veterans Benefits Administration (VBA); Chesapeake Employers' Insurance Company (the State of Maryland's workers' compensation program); and two private disability insurance providers, Prudential Financial and Sun Life Financial. This chapter also provides information on additional benefit programs, including the Washington State Department of Labor and Industries (L&I), the Canada Pension Plan (CPP), and the Disability Support Pension (DSP) (Australia).[1] The chapter presents a brief overview of

[1]The amount of information provided to or obtained by the committee varied among the programs discussed in this chapter. There is much publicly available information about the programs and overall disability determination processes of the VBA, the CPP, and DSP. With respect to functional assessment in the programs' disability determination processes, the committee had to rely more heavily on the written or oral information provided by each entity. Significantly less information is publicly available on the disability determination processes and procedures of private disability insurers. For this reason, the committee had to rely primarily on the information provided by the representatives of the two private disability insurance providers discussed in this chapter.

each of these programs, along with descriptions of their disability determination processes and examples of the functional information they collect during assessment. It should be noted that the organizations and agencies described in this chapter are not intended to be a representative sample of the large numbers of such entities conducting disability assessment. Likewise, the information on foreign agencies was derived largely from a website and literature review, and is not intended to be representative of international practices. Annex Table 8-1 at the end of this chapter summarizes elements of each of the programs, allowing for comparisons among them.

OVERVIEW OF SELECTED BENEFIT PROGRAMS

Selected Federal Disability Benefit Programs

Veterans Benefits Administration

The VBA, one of three administrations forming the VA, provides benefits and services to service members, veterans, and their families. The VBA's mission "is to serve as a leading advocate for Servicemembers, Veterans, their families and Survivors, delivering with excellence Veteran-centered and personalized benefits and services that honor their service, assist in their readjustment, enhance their lives, and engender their full trust" (VA, 2017b, p. 2). In fiscal year 2017, the VBA spent approximately $2.6 billion to administer benefits and services through six major program areas: compensation, pension and fiduciary, education, insurance, home loan guaranty, and vocational rehabilitation and employment (VA, 2017b). For fiscal year 2017, the total number of recipients for each program was as follows: compensation, 4,964,209; pension, 478,003; fiduciary, 211,282; education, 946,829; insurance, 6,007,606; home loan guaranty, 740,389; and vocational rehabilitation and employment, 132,218 (VA, 2017b).

Canada Pension Plan

Service Canada, which is part of Employment and Social Development Canada, provides Canadians with "a single point of access to a wide range of government services and benefits," including Old Age Security, the CPP, and private pensions and savings (Government of Canada, 2018f). Funded through worker and employer contributions, the CPP provides benefits to contributors and their families in the event of retirement, disability, or death (Government of Canada, 2018a). The CPP provides benefits to residents living throughout Canada with the exception of Quebec, where the Quebec Pension Plan provides similar benefits. In fiscal year 2017, approximately 5.6 million CPP beneficiaries, including 335,000 people with disabilities and 83,000 of their children, received payments (Government of Canada, 2017). Disability benefits represented $4.3 billion (10 percent) of total CPP benefits paid out of $47.0 billion in employee and employer contributions during that fiscal year (Government of Canada, 2017). In 2018, monthly CPP disability benefits start at $485.20 and, depending on the individual's previous CPP contributions, increase to a maximum of $1,335.83, with a monthly average of $971.23 (Government of Canada, 2018d). Eligible dependent children can receive a children's benefit amounting to $244.64 monthly (Government of Canada, 2018d).

Disability Support Pension (Australia)

DSP, administered by the Australian Government Department of Human Services, provides financial support for people with a permanent physical, intellectual, or psychiatric condition that prevents them from working. From 2017 to 2018, DSP processed 104,000 claims, 31,000 of which were granted (Australian Government Department of Human Services, 2018g). Payments, provided every 2 weeks, depend on the individual's age and living conditions. The maximum basic rates are as follows for those aged 21 or over with or without children or under 21 with children: $834.40 (single), $629.00 (couple, each), and $1,258.00 (couple, combined) (Australian Government Department of Human Services, 2018f). For individuals under 21 without children the maximum rates are as follows: $371.20 (single, under 18, and living at home); $572.90 (single, under 18, and independent); $420.70 (single, aged 18–20, and living at home); $572.90 (single, aged 18–20, and independent); and $572.90 (a member of a couple, aged 20 or younger) (Australian Government Department of Human Services, 2018f).

Selected State Workers' Compensation Programs

Chesapeake Employers' Insurance Company

Chesapeake Employers' Insurance Company, formerly known as the Injured Workers' Insurance Fund, provides workers' compensation insurance to Maryland businesses. Chesapeake's mission is "to provide Maryland businesses with a readily available source for workers' compensation insurance that features high quality products and services at a fair price, and to protect workers and employers by championing workplace safety" (Chesapeake Employers' Insurance, 2018a). In 2017, Chesapeake covered 20,889 injured workers for medical treatment, indemnity (lost time), or both (Chesapeake Employers' Insurance, 2017).

Washington State Department of Labor and Industries

L&I "provides no-fault industrial insurance coverage for most employers and workers in Washington State" (L&I, 2018f). This coverage includes medical treatment and limited wage-replacement benefits for workers with job-related injuries and illnesses (L&I, 2018a). Claim managers work with doctors, employers, and counselors to help injured employees return to work.

The department receives more than 150,000 claims each year (L&I, 2018h). Of accepted claims, 72 percent are administered by the State Fund Program (L&I, 2018h). The remaining 28 percent are self-administered by employers or contracted third-party administrators (L&I, 2018h); approximately 400 self-insured employers administer their own claims (L&I, 2018a). In fiscal year 2017, the State Fund Program insured 176,000 employers, covering 2.9 million employees; assessed $2.25 billion in premiums; and incurred $1.35 billion in benefits (L&I, 2018g). During the first half of fiscal year 2017, the program accepted 85 percent or 93,896 of the 109,962 claims received (L&I, 2018g). In that same fiscal year, 356 employers were self-insured, covering 917,127 workers, or about 25 percent of all workers covered by L&I (L&I, 2018g).

Selected Private Disability Insurance Providers

Prudential Financial

Prudential Financial provides customers with a range of products and services, including group and individual disability insurance, life insurance, annuities, and retirement-related services (Prudential Financial, 2018a). In the disability realm, the company's mission is to "enable customers' employees to return to work as soon as possible after a disabling event" (Tugman and Kramschuster, 2016). Prudential reports the provision of disability insurance for 2,500 clients (representing 1.3 million participants) for short-term disability and 3,100 clients (representing 2.1 million participants) for long-term disability (Prudential Financial, 2018b).

Sun Life Financial

Sun Life Financial is a financial services company that helps "clients achieve lifetime financial security and live healthier lives" (Sun Life Assurance Company of Canada, 2018). It offers a broad range of products and services to individuals, businesses, and institutions, including short- and long-term disability, life, and other types of insurance. Sun Life's goal with clients who have discontinued work as a result of illness or an injury is to learn why they are unable to work and how to get them back to work as soon as possible (Hamill, 2018). Therefore, the company focuses on assessing the ability to function in a work environment and increasing education on work modifications that can be performed to keep clients at work.

DISABILITY DETERMINATION (ADJUDICATION) PROCESSES

Selected Federal Disability Benefit Programs

Veterans Benefits Administration

The Compensation program "provides tax-free monthly benefits to Veterans in recognition of the effect of disabilities caused by diseases, events, or injuries incurred or aggravated during active military service" (VA, 2017a). Monthly benefits are also provided to surviving spouses, dependent children, and dependent parents. Eligibility is granted if the individual was discharged without dishonorable conditions and is at least 10 percent disabled by an injury or disease incurred or aggravated during active duty, active duty for training, or inactive-duty training (VA, 2017a).

The VA process for assessing disability compensation claims has eight distinct steps (VA, 2018c). In the first step, the VA receives the claim. In the second step, the claim is assigned to a veterans service representative who reviews the information provided to determine whether additional evidence is needed. If additional information is not needed, the claim proceeds to the preparation for decision phase.

In the third step, the representative gathers evidence by requesting it from various sources, such as the claimant, a medical professional, or a government authority. The claimant must submit all relevant evidence, including discharge or separation records, service treatment records, medical evidence from doctors, and hospital reports, to the VA to help the agency obtain all relevant information (VA, 2018d). The Veterans Claims Assistance Act defines claimants' responsibilities in providing evidence to support a claim, as well as the VA's responsibilities in helping claimants obtain that evidence (VA, 2018d). The VA is responsible for obtaining relevant records from any federal agency, such as the military, VA medical centers (including private facilities where the VA authorized treatment), and SSA. The VA is also responsible for providing a medical examination or obtaining a medical opinion. The claimant is responsible for obtaining records held by nonfederal agencies, such as state or local governments, private doctors and hospitals, and current or former employers. The VA can assist in obtaining these records. The claimant is also responsible for providing information needed by the VA to request records.

Two types of evidence are required: medical evidence of a current physical or mental disability and evidence establishing the relationship between the current disability and an injury, disease, or event that occurred during military service. Medical records or medical opinions are required to support this relationship. During this step, the VA may ask the claimant to complete a compensation and pension exam, which helps determine whether the disability is service connected, the level of disability, and whether the current condition should receive a higher rating because it is worsening.

Disability Benefits Questionnaires (DBQs) are used for veterans to provide disability-related information during the disability evaluation process. Veterans also have the option of using their own health care provider instead of a VA facility to complete the questionnaire. DBQs help speed the processing of claims by supplying private health care providers with the information and questions about the disability needed for them to submit medical evidence that will enable the VA to evaluate a claim accurately (VA, 2018f). Providers with active medical licenses are authorized to complete DBQs.

Mental health DBQs can be completed by a board-certified or board-eligible psychiatrist; a licensed doctorate-level psychologist; a doctorate-level mental health provider under close supervision of a board-certified or board-eligible psychiatrist or licensed doctorate-level psychologist; a psychiatry resident under close supervision of a board-certified or board-eligible psychiatrist or licensed doctorate-level psychologist; a clinical or counseling psychologist who has completed a 1-year internship or residency (for purposes of a doctorate-level degree) under close supervision of a board-certified or board-eligible psychiatrist or licensed doctorate-level psychologist; or a licensed clinical social worker, nurse practitioner, clinical nurse specialist, or physician assistant under close supervision of a board-certified or board-eligible psychiatrist or licensed doctorate-level psychologist (VA, 2018e). There are more than 70 DBQs covering a full range of medical conditions, all using standardized language to streamline the process. Checkboxes rather than long narrative summaries are used to document medical conditions. The forms collect specific information relevant to the medical condition of the claimant.

In step four, all of the evidence is reviewed. The VA provides detailed information for veterans to understand the mental health examination (VA, 2018i). When evaluating mental health conditions, the VA thoroughly reviews both medical and nonmedical evidence and can rate an individual claimant for only one mental health condition. The VA reviews the evidence to gain an understanding of how symptoms of the claimant's mental health condition impact social and industrial functioning and determines which mental health condition is related to the claim, based on the *Diagnostic and*

Statistical Manual of Mental Disorders, Fourth or Fifth Edition (DSM-IV or DSM-5) and using the Mental Disorder Criteria in the VA Schedule for Rating Disabilities (VA, 2018i).

In step five, the veterans service representative recommends a decision and prepares documents detailing that decision. In step six, the recommended decision is reviewed, and a final determination regarding award approval is made. A claim decision packet containing details of the decision or award is prepared for mailing (step seven), and the VA sends the decision packet to the claimant (step eight).

The disability compensation amount depends on the claimant's degree of disability based on the evidence submitted. The VA rates disability from 0 to 100 percent in 10 percent increments. Additional amounts are paid based on whether the claimant has a very severe disability or loss of limb(s); whether the claimant has a spouse, child(ren), or dependent parent(s); or whether the claimant has a seriously disabled spouse (VA, 2018b). If the VA finds that a veteran has multiple disabilities, the Combined Ratings Table is used to calculate a combined disability rating; the disability ratings are not additive. To calculate the combined rating, disabilities are arranged in order of severity, beginning with the most severe, and combined using the left column and top row of the Combined Ratings Table to determine where the values intersect (VA, 2018b). This value is then rounded to the nearest 10 percent. If there are more than two disabilities, the combined value for the first two is determined (as previously described), and this new value is then combined with the value for the third disability.

Canada Pension Plan

The CPP disability benefit is "a taxable monthly payment that is available to people who have contributed to the CPP and who are not able to work regularly because of a disability" (Government of Canada, 2018e). To qualify as eligible for a CPP disability benefit, an individual must have a severe and prolonged disability, be under the age of 65, and meet the CPP contribution requirements (Government of Canada, 2018c). Contributions are required in 4 of the last 6 years or, for those having contributed for more than 25 years, 3 of the last 6 years (Government of Canada, 2018c).

The CPP Adjudication Framework, a comprehensive framework of the policy elements of disability adjudication, provides decision makers with the information needed to adjudicate CPP disability applications. The framework consists of five components: a "severe" criterion for the prime indicator, a "severe" criterion for "incapable regularly of pursuing any substantially gainful occupation," personal characteristics and socioeconomic factors, a "prolonged" criterion, and a reasonably satisfied standard of review for determining eligibility or continuing eligibility for CPP disability benefits (Government of Canada, 2018b).

In the first of these components, the claimant's medical condition is defined as the prime indicator in determining whether a disability is "severe" and "prolonged." Multiple factors are considered in assessing the medical condition, including the nature of the condition and whether it is progressive, functional limitations due to the condition, the impact of treatment, statements/opinions expressed by the health professional and client, the existence of multiple medical conditions, and personal characteristics (Government of Canada, 2018b). To determine whether the nature of the medical condition is "severe" and can lead to the inability to work, a medical adjudicator uses his or her medical knowledge, the CPP legislation, and medical information provided by the individual and health professionals. When determining disability, the CPP focuses only on those functional limitations that affect the capacity to work. Functional limitation is defined as "an impairment that leads to less than normal performance of an individual" (Government of Canada, 2018b). The adjudicator also determines the impact of treatments on the medical condition and the person's ability to work. For the CPP, individuals with two or more medical conditions are considered to have multiple conditions. Even if individual medical conditions are not considered to result in an inability to work, they may be considered to do so if they are among an individual's multiple conditions.

For the second component, the applicant "must demonstrate that he or she has a 'severe' and 'prolonged' physical or mental disability that prevents him or her from regularly pursuing any substantially gainful occupation" (Government of Canada, 2018b). To satisfy this disability test, three factors are considered in relation to work activity—performance, productivity, and profitability.

The third component relates to personal characteristics and socioeconomic factors. Personal characteristics include age, education, and work experience and its impact on the individual's medical condition, work capacity, and ability to perform in a substantially gainful occupation. The individual's particular characteristics are considered only with respect to their direct effect on the ability to work. Socioeconomic factors are not considered in the determination of CPP disability.

For the fourth component, after an individual meets the "severe" criterion, the "prolonged" criterion is assessed. Similar to SSA's duration requirement, the CPP's "prolonged" criterion has two components—"likely to result in death" or "likely to be long continued" and "of indefinite duration" (Government of Canada, 2018b). Only one of these two components must be met.

For the fifth component, the CPP uses a "reasonably satisfied" standard of proof when determining eligibility or continuing eligibility for disability benefits, established by considering all evidence pertinent to the medical condition, the capacity to work, and personal characteristics. Pertinent evidence refers to evidence collected from physicians; medical reports by family physicians, specialists, and/or other health care professionals; reports of diagnostic investigations; employer reports; and functional capacity assessments (Government of Canada, 2018b). Providers such as psychologists, neuropsychologists, physiotherapists, occupational therapists, and vocational rehabilitation professionals can provide evidence of work capacity. Examples of such evidence include functional capacity assessments, statements from educational institutions and employers, power-of-attorney documents, and certificates of incapacity. Since Canada has universal health insurance, it is reasonable to expect that claimants can obtain evidence of their impairment and its limitations, in contrast to the United States. The CPP compares and evaluates opinions expressed by all of these professionals and statements made by the applicant to obtain a total picture of the individual. If the overall available evidence supports a determination that the individual meets the "severe" and "prolonged" criteria, the disability benefit is granted.

Disability Support Pension (Australia)

To be eligible for DSP, applicants must be "between 16 and pension age; meet residency requirements; meet income and assets tests; and have a permanent and diagnosed disability or medical condition or get a Department of Veterans' Affairs special rate disability pension due to total and permanent incapacity" (Australian Government Department of Human Services, 2018a). The applicant also may need to show participation in a Program of Support—resources for job preparation, work training, and injury management—and demonstrate that he or she cannot work or retrain to work for at least 15 hours a week in the next 2 years (Australian Government Department of Human Services, 2018a). During disability assessment, DSP seeks to understand how an individual's disability or medical condition affects him or her. The applicant is responsible for providing current medical evidence from a treating health professional(s) about each of the medical conditions that may impact his or her ability to work. Examples of medical evidence that may be required include medical history reports; physiotherapy or audiology reports; psychologist reports, including results of IQ testing; and physical examination reports (Australian Government Department of Human Services, 2018c). More importantly, current information must be provided regarding the diagnosis, treatment, symptoms, functional impact, and prognosis of each medical condition that may impact the individual's ability to work. Specific evidence is required for some medical conditions: mental health conditions, which require a diagnosis from a psychiatrist or clinical psychologist; intellectual impairment, which requires evidence from a psychologist, who must assess intellectual function and adaptive behavior and provide information that includes an IQ score or ability to undergo testing; ear conditions affecting hearing or balance, which require evidence that an audiologist or an ear, nose, and throat specialist supports the diagnosis; and eye conditions affecting vision, which require evidence that an ophthalmologist or ophthalmic surgeon supports the diagnosis (Australian Government Department of Human Services, 2018c).

Applicants may be assessed as manifestly medically eligible if they "are permanently blind; have a terminal illness with a life expectancy of less than 2 years; have an intellectual disability with an IQ of less than 70; or need nursing home level care" (Australian Government Department of Human Services, 2018b). If applicants fail to meet the criteria described above, their disability or medical condition needs to be fully diagnosed, treated, and stabilized. DSP will then assess how the condition affects the applicant's ability to function daily.

A Job Capacity Assessment may be used to help DSP assess functioning using impairment tables. DSP uses 15 impairment tables to determine an impairment rating. These tables include functions requiring physical exertion and stamina; upper-limb function; lower-limb function; spinal function; mental health function; functioning related to alcohol, drug, and other substance use; brain function; communication function; intellectual function; digestive and reproductive function; hearing and other functions of the ear; visual function; continence function; functions of the skin; and functions of consciousness (Australian Government, 2011). The tables are "function based rather than diagnosis based; describe functional activities, abilities, symptoms and limitations; and are designed to assign ratings to determine the level of functional impact of impairment and not to assess condition" (Australian Government, 2011, p. 5).

Additional medical evidence is evaluated during a Disability Medical Assessment performed by a government-contracted doctor. An applicant is considered medically eligible for DSP benefits if his or her medical condition prevents working at least 15 hours per week in the next 2 years, and if he or she has an impairment rating of 20 points or more on a single impairment table or 20 points or more combined across more than one impairment (Australian Government Department of Human Services, 2018b).

Individuals receiving DSP benefits may be required to follow a Participation Plan to encourage return to work. If their medical condition changes after benefits have been awarded, an Employment Services Assessment is performed to reassess the claimant's medical condition and capacity to work (Australian Government Department of Human Services, 2018d,e).

Selected State Workers' Compensation Programs

Chesapeake Employers' Insurance Company

The structure at Chesapeake encompasses claims management, health services, and fraud management and investigation, along with an in-house legal department. Within the health services department, in-house doctors, nurses, and medical specialists work with claims professionals to ensure appropriate medical treatment for injured workers. The goal is to return injured workers to work as soon as it is medically possible (Fisher and Smulyan, 2018). The health services team consists of 3 doctors, 20 nurses, a pharmacist, and a physical therapist (Chesapeake Employers' Insurance, n.d.-b). Two orthopedic surgeons and a third physician with experience in workers' compensation injuries provide advice on treatment interventions, as well as cost containment options such as generic and lower-cost prescriptions. Approximately 20 nurses provide telephonic case management and ensure that the medical treatment received by injured workers is appropriate and timely (Chesapeake Employers' Insurance, n.d.-b). Nurses also coordinate with claims adjusters to ensure contact among injured workers, their employer, and medical providers (Chesapeake Employers' Insurance, n.d.-b). The pharmacist monitors drug utilization and prescriptions, while the physical therapist monitors the progress of injured workers for short- and long-term goals.

When an individual is injured, his or her employer reports the injury in an initial report to Chesapeake. The adjuster speaks with the injured worker to learn about the job situation and how it contributed to the injury (Fisher and Smulyan, 2018). A medical history and a history of previous injury and/or surgery are also collected. Once the injury has been reported, medical care is initiated at an occupational health facility, a private physician's office, or an emergency room/urgent care facility (Fisher and Smulyan, 2018). Thirty days of treatment, including physical therapy, is covered automatically to prevent delays in receiving care. Chesapeake's goal for disability management is to achieve maximal medical improvement (MMI), defined as "the point at which the condition of an injured person is stabilized. No further recovery or improvement is expected, even with additional medical intervention" (Chesapeake Employers' Insurance, 2018b). MMI status is determined through review of regular medical reports from physicians and ancillary providers. Psychological evaluations may also be used to identify behavioral health factors preventing a return to work (Fisher and Smulyan, 2018). Once MMI has been reached, a determination as to whether the individual can return to previous work activity can be made. If an individual is unable to return to previous work activity, the job analysis provided by the employer is used to identify modified duties (Fisher and Smulyan, 2018). An independent medical evaluation may be provided to determine preexisting conditions, conditions related to work injury or exposure, and current diagnosis.

Washington State Department of Labor and Industries

Under the L&I disability assessment process, an individual who is injured at work can file an accident report at a doctor's office, on the L&I website, or by phone. If the applicant works for a self-insured employer, the accident report must be filed with the employer (L&I, 2018c). Workers may file claims within 1 year of their injury date or within 2 years of receiving a diagnosis of an occupational disease (L&I, 2018c). The accident report includes information about the applicant's injury, employer, wages, diagnosis, and treatment, along with other background information (L&I, 2018c). The applicant's doctor may complete a Physicians Initial Report. If the worker's ability to work is limited, an Activity Prescription Form is completed (L&I, 2018c). This form rates the worker's current capacity to work and may be completed by a qualified attending health care provider, including a doctor currently licensed in medicine (including osteopathic), surgery (including podiatric), or dentistry or a chiropractor who is a department-approved examiner (L&I, n.d.-a). After a claim has been filed, a Claim Arrival Card is sent to the worker by mail, along with information on benefits and return-to-work resources. Correspondence requesting additional information may also be included (L&I, 2018c). A claims manager may require that an independent medical examination (IME) be performed to determine the extent of the impairment or to learn more about the condition's treatment or duration. The IME is scheduled and paid for by L&I (L&I, 2018d). L&I or the self-insured employer will approve the claim if the applicant's doctor can certify that the applicant was injured at a specific time and place at work or has an occupational disease (L&I, 2018e). The benefits provided will cover medical bills and may include wage replacement, return-to-work assistance, and disability or pensions for the severely injured (L&I, 2018e).

Selected Private Disability Insurance Providers

Prudential Financial

Prudential's group disability insurance provides both short- and long-term disability benefits, return-to-work services, absence management, and health and productivity data analytics and consulting. When a claimant applies for disability insurance, Prudential's goals include understanding all of the medical conditions impacting the employee, identifying functional capacity and restrictions and limitations, identifying prognosis and when and whether return to work is expected, and using vocational rehabilitation resources to facilitate a safe return to work (Kramschuster, 2018). The initial acquisition of clinical information encompasses a statement of disability by the employee; an attending physician's statement; and medical records, including a history of surgery, therapy, and medications (Kramschuster, 2018). Once this information has been collected, the claims manager facilitates an internal review of the information with a vocational rehabilitation specialist and clinician—typically a registered nurse (Kramschuster, 2018). During this review, the claim is discussed in detail to determine whether the information is consistent, the prognosis is understood, and there is any current capacity for return to work at any level (Kramschuster, 2018). If additional information is needed, it is requested; otherwise, the file is sent to an internal clinician for a full file review to determine capacity to return to work. If the information is sufficient, Prudential will have a clear understanding of functional capacity and can make a determination.

Sun Life Financial

Sun Life Financial's disability claims management process consists of a thorough assessment, starting with a comprehensive analysis of information related to the claimant's injury and absence from work (Sun Life Assurance Company of Canada, n.d.). This information is collected through a telephone conversation between the claims handler and the claimant, followed by the submission of functional assessment forms and the Attending Physician Statement (APS). The information provided by the attending physician is key to the process (Hamill, 2018). Sun Life's assessment includes engaging the claimant, with a focus on his or her functional abilities (Sun Life Assurance Company of Canada, n.d.). Ancillary information, such as medical records, pharmacy scans, information from the employer, publicly available information, and in-person interviews or surveillance, is used before a decision is made (Hamill, 2018). This ancillary information includes the types of medications the claimant uses and physicians the claimant sees.

Sun Life's goal is to ensure optimal treatment and establish expectations for recovery to help the claimant understand that other factors can affect recovery, such as perceptions regarding recovery, workplace, family, and financial issues (Hamill, 2018). When establishing expectations for recovery, the discussion between the claimant and his or her physician is focused on identifying return-to-work goals and providing the claimant with information on work accommodations to help with return-to-work planning (Sun Life Assurance Company of Canada, n.d.). When in-depth rehabilitation, treatment, or other interventions are necessary, Sun Life strives to ensure that appropriate resources are involved. Sun Life will also negotiate return-to-work plans that align the claimant's functional abilities with work demands. Predictive modeling and innovative data mining are used to improve claimant outcomes (Sun Life Assurance Company of Canada, n.d.).

COLLECTION OF FUNCTIONAL INFORMATION

Selected Federal Disability Programs

Veterans Benefits Administration

The VA uses DBQs to collect functional information related to the impairments described in Chapter 7, including back conditions, heart conditions, mental disorders, and traumatic brain injury (TBI).

During the collection of medical history in the DBQ for back (thoracolumbar spine) conditions, the veteran is asked to provide a description of flareups that impact the function of the thoracolumbar spine, as well as any functional loss or impairment to the back (VA, 2018a). The VA defines functional loss as "the inability, due to damage or infection in parts of the system, to perform normal working movements of the body with normal excursion, strength, speed, coordination and/or endurance" (VA, 2018a, p. 4). Factors that contribute to functional loss or impairment are identified using information from the medical history and a physical exam. The physical exam includes measurement of range of motion (ROM) using a goniometer. The VA requires repetitive-use testing in all joint exams. Specifically, three repetitions of ROM are performed. ROM measurements are taken for joint movements such as forward flexion, extension, right and left lateral flexion, and right and left lateral rotation. Information is collected on pain associated with ROM movements, as well as pain during weight-bearing and non-weight-bearing movement. The physical exam also includes muscle strength testing, with strength ratings such as normal strength, active movement against some resistance, active movement against gravity, and no muscle movement being assigned (VA, 2018a). In addition, functional information related to back conditions is collected from a reflex exam, sensory exam, and straight-leg raising test. Finally, the form collects information on the functional impact of the diagnosed condition(s) on the individual's ability to perform any type of occupational task.

Functional information collected on the heart conditions DBQ encompasses ischemic and nonischemic heart disease, arrhythmias, valvular disease, and cardiac surgery (VA, 2018g). This information, including heart rate, rhythm, point of maximal impact, heart sounds, peripheral pulses, peripheral edema, and blood pressure, is collected from the physical exam. For all heart conditions, the VA requires a determination of the presence of cardiac hypertrophy and dilation. The suggested order of tests for cardiac hypertrophy/dilation is electrocardiogram (EKG), then chest X-ray (posterioranterior and lateral), then echocardiogram (VA, 2018g). An echocardiogram is necessary only if the EKG and chest X-ray results are negative. Tests included on the form that may reflect the veteran's current functional status are EKG, chest X-ray, echocardiogram, holter monitor, multigated acquisition scan, coronary artery angiogram, and computed tomography angiogram (VA, 2018g). In addition, the VA requires metabolic equivalents of task testing for all heart exams—either exercise based or interview based—to determine the activity level at which symptoms develop.

The mental disorders DBQ provides options for indicating the veteran's level of occupational and social impairment with regard to all mental diagnoses (VA, 2018h). Examples of these options include "occupational and social impairment due to mild or transient symptoms which decrease work efficiency and ability to perform occupational tasks, only during periods of significant stress, or, symptoms controlled by medication" and "occupational and social impairment deficiencies in most areas, such as work, school, family relations, judgment, thinking and/or mood" (VA, 2018h, p. 2). The form also asks about multiple symptoms that may affect functioning, including "impairment of short and long term memory, for example, retention of only highly learned material, while forgetting to complete tasks"; "difficulty in understanding complex commands"; "difficulty in establishing and maintaining effective relationships"; and "intermittent inability to perform activities of daily living, including maintenance of minimal personal hygiene" (VA, 2018h, p. 4).

The Residuals of TBI DBQ is used to evaluate the residuals of TBI (VA, 2011). The section on assessment of cognitive impairment and other residuals of TBI mentions that neurological testing may be necessary to complete it accurately. The section collects information on the veteran's current level of functional status for 10 facets of TBI-related cognitive impairment and subjective symptoms: memory, attention, concentration, and executive functions; judgment; social interaction; orientation; motor activity (with intact motor and sensory system); visual spatial orientation; subjective symptoms; neurobehavioral effects; communication; and consciousness (VA, 2011). For subjective symptoms, information is collected on their interference with instrumental activities of daily living (IADLs). The form also collects information on "any subjective symptoms or any mental, physical or neurological conditions or residuals attributable to a TBI," such as motor dysfunction, hearing loss and/or tinnitus, speech disorder, and mental disorder (including emotional, behavioral, or cognitive) (VA, 2011, p. 62).

Canada Pension Plan

Applicants for CPP disability benefits are required to submit a medical report. A physician, a nurse practitioner, or—in geographically isolated communities—a registered nurse may complete the Medical Report for a Canada Pension Plan Disability Benefit (Service Canada, 2018). Information is collected on each medical condition, its impairments, and its functional limitations and treatment. Information on the condition's prognosis (improve, deteriorate, remain the same, or unknown), expected duration (less than 1 year, more than 1 year), and frequency (recurrent/episodic, continuous, unknown) is also collected. Another section collects information that can be used to assess current and future restrictions on the patient's ability to work. The form includes a list of grave medical conditions (with marked and severe functional limitations) that have a high probability of meeting the eligibility criteria for CPP disability benefits.

The form also provides examples of functional limitations by physical abilities, behaviors and emotional abilities, communication and thinking abilities, and other daily abilities. Physical abilities may include restrictions related to changing body position (kneeling or squatting), maintaining body position (remaining seated or standing), fine hand use (turning a dial or knob), using transportation (as a passenger in a taxi or on a bus or the subway), and using a computer (being able to look at a computer screen for at least 20 minutes) (Service Canada, 2018). Behaviors and emotional abilities may include restrictions related to basic interpersonal interactions (showing respect and tolerance), maintaining formal relationships (with employers or service providers), and handling stress and other psychological demands (Service Canada, 2018). Communication and thinking abilities may include restrictions related to making conversation (with known individuals or strangers), thinking (sequencing thoughts in a structured, logical manner), and making decisions (identifying and choosing among several options) (Service Canada, 2018). Other daily abilities include restrictions related to toileting, dressing, looking after one's health (taking medication as directed), acquiring goods and services, maintaining economic self-sufficiency (managing money), doing housework, preparing meals, and driving (Service Canada, 2018).

Disability Support Pension (Australia)

As mentioned earlier, DSP uses impairment tables, "designed to assign ratings to determine the level of functional impact of impairment" to assess impairment in relation to work (Australian Government, 2011, p. 5). To select the correct table to use, the following steps are followed: "identify the loss of function; then refer to the Table related to the function affected; then identify the correct impairment rating" (Australian Government, 2011, p. 9). If a single condition causes multiple impairments, each impairment should be assessed with the relevant table. If two or more conditions cause a combined impairment, a single rating is assigned under a single table. Impairment tables related to both physical and mental functioning are discussed below.

For Table 1 on functions requiring physical exertion and stamina, the introduction indicates that the table should be used "where the person has a permanent condition resulting in functional impairment when performing activities requiring physical exertion or stamina" (Australian Government, 2011, p. 12). It also states that the diagnosis must be made by an appropriately qualified medical practitioner. With respect to evidence, self-report of symptoms alone is not sufficient; corroborating evidence of the person's impairment is necessary. Examples of corroborating evidence include a report from the person's treating doctor; a report from a medical specialist confirming a diagnosis of a condition commonly associated with cardiac or respiratory impairment; and results of exercise, cardiac stress, or treadmill testing.

The next section of the table includes multiple levels of functional impact on activities requiring physical exertion or stamina, with various point ratings: none = 0 points, mild = 5 points, moderate = 10 points, severe = 20 points, and extreme = 30 points. The table also provides detailed information for each functional impact level. For no functional impact, the person is "able to undertake exercise appropriate to their age for at least 30 minutes at a time; and has no difficulty completing physically active tasks around their home and community" (Australian Government, 2011, p. 12). If there is an extreme functional impact related to physical exertion or stamina, the person is "completely unable to perform activities requiring physical exertion or stamina or experiences symptoms (e.g., shortness of breath, fatigue, cardiac pain) when performing any activities requiring physical exertion or move around inside the home without assistance" (Australian Government, 2011, p. 14).

For Table 5 on mental health function, the introduction indicates that the table should be used "where the person has a permanent condition resulting in functional impairment due to a mental health condition (including recurring episodes of mental health impairment)" (Australian Government, 2011, p. 22). It also states that the condition must be diagnosed by an appropriately qualified medical practitioner (includes a psychiatrist) with evidence from a clinical psychologist (if the diagnosis has not been made by a psychiatrist). In addition to self-report of symptoms, corroborating evidence is required, which may include a report from the person's treating doctor; supporting letters, reports, or assessments relating to the person's mental health or psychiatric illness; and interviews with the person and those providing care or support to the person. The table also includes a note stating that the signs and symptoms of mental health impairment may vary over time, and it is important not to rely solely on how the individual may present on the day of the assessment. More specifically, "for mental health conditions that are episodic or fluctuate, the rating that best reflects the person's overall functional ability must be applied, taking into account the severity, duration and frequency of the episodes or fluctuations as appropriate" (Australian Government, 2011, p. 22).

The next section of the table provides multiple levels and point ratings of functional impact on activities involving mental health function. A person with no functional impact has no difficulties with most of the following: "self care and independent living; social/recreational activities and travel; interpersonal relationships; concentration and task completion; behavior, planning and decision-making; and work/training capacity" (Australian Government, 2011, p. 23). If there is an extreme functional impact, the person has extreme difficulty with these mental health functions (Australian Government, 2011).

Selected State Workers' Compensation Programs

Chesapeake Employers' Insurance Company

Chesapeake collects functional information using a job analysis form and functional capacity evaluation. On the job analysis form, employees are asked how often—never, occasional (11–33 percent), frequent (34–66 percent), and constant (67–100 percent)—they perform various tasks, such as standing, sitting, walking, stooping, kneeling, reaching, and fine manipulations (Chesapeake Employers' Insurance, n.d.-c). They are also asked about the amount of weight that can be lifted, carried, pushed/pulled, and handled. This form utilizes the U.S. Department of Labor's classification of five degrees of work in terms of lifting requirements: sedentary work, light work, medium work, heavy work, and very heavy work. Based on the evidence collected with this form, the physician completing the form will determine whether and when an employee can return to work on regular or transitional/modified duty.

The functional capacity evaluation is performed by a physical therapy group to determine work capability (Chesapeake Employers' Insurance, n.d.-a). Information is collected on lifting tasks, including the amount of weight lifted, how often (occasional, frequent, or constant), and the type of lift (floor to waist, knee to waist, waist to shoulder, waist to overhead, carry, push, or pull). In addition, information is collected on such tasks as sitting, standing, walking, and squatting. Information on active ROM and strength deficits is also collected. Worker performance findings on the form include the categories of symptom management, worker traits, and consistency of effort. Symptom management includes the following abilities: knowledge of appropriate strategies, demonstration of appropriate strategies, response to activity/evaluation, and effectiveness of current strategies (Chesapeake Employers' Insurance, n.d.-a). Worker traits include productivity, safety, and interpersonal behavior. Consistency of effort includes observation of illness behavior, reliability of pain and disability reports, and physical effort (Chesapeake Employers' Insurance, n.d.-a). Vocational information is also collected through a summary of duties and description of physical requirements; physical requirements can be assessed based on client report, employer report, the U.S. Department of Labor's *Dictionary of Occupational Titles*/Occupational Information Network, or a job description (Chesapeake Employers' Insurance, n.d.-a).

Washington State Department of Labor and Industries

L&I collects functional information via the Activity Prescription Form (APF), the IME Doctor's Estimate of Physical Capacities, the Functional Capacity Summary, and the Job Analysis Summary. The APF collects information on work status, including when the worker was released to the job where injury occurred, whether the worker can perform modified duties, whether the worker may work limited hours, and measurable objective findings related to work ability (L&I, 2018b). The IME Doctor's Estimate of Physical Capacities consists of seven sections (L&I, 2007). The first assesses the worker's ability to sit, stand, and walk for durations at one time and during an entire 8-hour day. The second and third sections evaluate the worker's ability to lift and carry, by both weight and frequency. The fourth and fifth sections address repetitive hand and foot tasks and movements. The sixth section assesses whole-body movements by frequency. The seventh section records specific activity restrictions involving unprotected heights; being around moving machinery; exposure to marked changes in temperature and humidity; driving automotive equipment; and exposure to dust, fumes, and gases. The Functional Capacity Summary and Job Analysis Summary similarly collect information on the amount of time (never, seldom, occasional, frequent, and constant) various tasks can be performed, such as climbing stairs, bending/stooping, kneeling, reaching forward, keyboarding, and fine manipulation (L&I, 2014, 2016).

L&I also makes available a list of particular functional scales: the Patient Specific Functional and Pain Scale, Functional Activity Back Questionnaire, STarT Back Screening Tool, Tampa Scale-11, Yellow Flags Questionnaire, Bournemouth Questionnaire-Neck, Bournemouth Questionnaire-Back, Neck Disability Index, Revised Oswestry Disability Questionnaire, Roland-Morris Low Back Pain & Disability Questionnaire, QuickDASH, Shoulder

Pain & Disability Index, Upper Extremity Functional Index, Foot & Ankle Ability Measure, and Lower-Extremity Function Scale (L&I, n.d.-b).

Selected Private Disability Insurance Providers

Prudential Financial

Prudential provides multiple forms for use by physicians to collect medical information, including the Capacity Questionnaire, Mental Status Examination, Behavioral Health Capacity Questionnaire, Kurtzke Functional Systems Scores, Visual Capacity Questionnaire, and Psychiatric Work Readiness Assessment (Prudential Financial, 2018c). The Capacity Questionnaire is a full-time work capacity assessment tool (reviewing 8-hour days for 5 days per week) (Prudential Financial, 2017). Specifically, the Capacity Questionnaire asks the physician completing the form whether the patient is capable of full-time work for 8 hours per day, 5 days per week or part-time transitional work. The form also collects information on the percentage of time—never (0 percent), occasionally (1–33 percent), frequently (34–66 percent), or constantly (67–100 percent)—during a typical day that the patient can perform various tasks, such as standing, walking, sitting, reaching overhead, stooping, and kneeling/crawling. The physician must also document any other medically necessary restrictions and/or limitations and any accommodative measures that would allow the patient to increase work capacity.

The Mental Status Examination form collects information on various aspects of functioning to determine whether the patient is within normal limits with respect to appearance, attitude, behavior, speech, mood, affect, thought process, thought content, cognition, and insight/judgment (Prudential Financial, 2016). Specifically for cognition, information is collected on orientation, concentration/attention, recent memory, remote memory, calculations, and abstractions.

The Behavioral Health Capacity questionnaire asks multiple questions related to functional capacity, restricted to within the last 14 days (Prudential Financial, 2015a). Functional information collected includes description of any work the patient should not do because of psychiatric symptoms/disease, psychosocial stressors that caused the patient to leave work, whether there is cognitive disease and how it was measured, and what barriers to returning to work exist.

The Kurtzke Functional Systems Scores assessment, based on neurological examination of functions, includes such categories as pyramidal functions, brainstem functions, sensory functions, cerebellar functions, bowel and bladder functions, cerebral (or mental) functions, and visual functions (Prudential Financial, 2014a). The Visual Capacity Questionnaire asks about visual functions that include best corrected acuity, field of vision, muscle function, two-eye depth perception, color perception, intraocular pressure, and fundus exam (Prudential Financial, 2014b). The form also asks whether the patient is able to do various tasks and for what percentage of an 8-hour workday. Tasks include walk on uneven surfaces, balance, drive motorized vehicles, climb, bend, use a computer, and read instructions. The form also asks whether a low-vision assessment can help in establishing visual functional capacity and/or accommodation needs.

The Psychiatric Work Readiness assessment, which is completed by a physician, asks whether the employee has work restrictions or limitations due to a psychiatric condition (Prudential Financial, 2015b). The form defines "restriction," "limitation," and "treatment plan." Restriction would "indicate your recommendation that the employee not perform a specified activity because of risk to self or others" (Prudential Financial, 2015b, p. 1). Limitation would "indicate your opinion that the employee is not cognitively or otherwise capable of performing a specified activity" (Prudential Financial, 2015b, p. 1). The treatment plan describes "in detail treatment methods that specifically target the psychiatric impairment" (Prudential Financial, 2015b, p. 1). For each restriction or limitation, the form asks the physician to provide a specific treatment plan for restoring work readiness in that particular area.

Sun Life Financial

For both long- and short-term disability benefits, Sun Life provides questionnaires for the claimant, attending physician, and employer to complete. The Attending Physician Statement (APS) can be completed by a "family doctor, a doctor at a walk-in clinic, a specialist or nurse practitioner—any medical professional who is a doctor of medicine and has treated you for your condition" (Sun Life Financial, n.d.-c, p. 3). Information is collected on the current diagnosis; what the condition is related to, such as a workplace injury or auto accident; symptoms and frequency of symptoms; and severity. In the clinical findings and observations section of the APS, the attending physician is asked to attach any test results, investigations, or consultation reports. The form also asks whether any formal functional tests have been performed and for the claimant to provide a copy of the report. In addition, the physician is asked whether the patient exhibits difficulty (specifying "none, slight, moderate, severe") in the following abilities: memory, decision making, concentration/focus, speech, sleep, sensation, walking, and climbing (Sun Life Financial, n.d.-c). The form also collects information on cardiac conditions and asks for functional capacity as defined by the American Heart Association. Options include Class I (no limitation), Class II (slight limitation), Class III (marked limitation), and

Class IV (complete limitation). If functional capacity is Class III or IV, a copy of a stress test or cardiac echograms should be included. Information on hospitalizations and medications is collected as well.

The attending physician can also complete a form that provides information on activities of daily living (ADLs) (Sun Life Financial, n.d.-b). This information includes whether no assistance is needed, stand-by assistance is needed, or physical hands-on assistance is needed. ADLs listed on the form include bathing, continence, dressing, eating, toileting, and transferring (getting into a bed or a wheelchair). The ADL questionnaire also asks about the presence of a cognitive impairment, defined as "an individual has a deterioration or loss in intellectual capacity resulting from injury, sickness, advanced age or Alzheimer's disease and similar forms of irreversible dementia and the individual needs another person's assistance or verbal cuing for the individual's protection or for the protection of others" (Sun Life Financial, n.d.-b, p. 2). Claimants provide information describing their injury/illness, treatment, doctors seen for their injury/illness, and job duties they currently cannot complete.

The Disability Job Demands questionnaire asks for information related to the plan member's specific job duties. Questions related to the conditions of the job are asked about work in varying conditions, such as outside, in extremes of cold or heat, in a noisy environment, or around toxic fumes (Sun Life Financial, n.d.-a). The questionnaire also asks about the percentage of time the plan member lifts or carries various weights. In addition, the plan member provides information on work activities that include walking, climbing, bending/crouching, and kneeling/crawling and the percentage of time performed. Information on cognitive and nonphysical aspects of the job is collected as well.

SUMMARY

Annex Table 8-1 provides a comparison of the size of the disability programs reviewed herein, as well as their disability focus and the types of functional information they collect, which include physical abilities, mental abilities, and ADLs. Although most of the programs use generic forms to collect functional information, both the VBA and DSP report that they use an impairment-specific approach in collecting this information. The VBA provides applicants with more than 70 DBQs with which to collect disability information for a full range of medical conditions. DSP utilizes 15 impairment tables categorized by functioning to assign ratings used to determine the level of functional impact of impairments. Annex Table 8-1 also shows that the VBA, the CPP, and Sun Life Financial collect information on ADLs to assess functioning. The VBA's residuals of TBI DBQ collects information on whether subjective symptoms mildly or moderately interfere with IADLs.

The CPP's Medical Report provides examples of daily abilities that may be affected by a medical condition, including toileting, dressing, doing housework, and preparing meals. Sun Life Financial's ADL questionnaire collects information on various ADLs—bathing, dressing, toileting, transferring, continence, and eating—providing definitions of each.

FINDINGS AND CONCLUSIONS

Findings

8-1. The mission and size of different disability benefits programs vary greatly. State workers' compensation programs (e.g., Chesapeake Employers' Insurance Company, Washington State Department of Labor and Industries) and private disability insurers (e.g., Sun Life Financial, Prudential Financial) focus on and facilitate the return of individuals to work. The U.S. Veterans Benefits Administration's (VBA's) mission is to compensate veterans who become disabled as a result of their military service, independently of whether an individual is able to work. The Canadian Pension Plan (CPP) and Australia's Disability Support Pension (DSP) are federal government programs that provide disability benefits to qualified individuals who are unable to perform work at the level specified by the program. The CPP serves by far the greatest number of beneficiaries relative to the other programs examined, and Prudential the least.

8-2. The VBA uses more than 70 disability benefit questionnaires to collect functional information about specific diseases and organ systems across the entire range of medical conditions. If a veteran is found to have more than one disability, a rating table is used to calculate a combined disability rating.

8-3. The CPP Adjudication Framework consists of five components: a severity criterion for the prime medical condition, severity criteria for the inability to regularly pursue "any substantially gainful occupation," personal characteristics and socioeconomic factors, a criterion related to the length of disability ("prolongation"), and a reasonably satisfied standard of review for determining eligibility or continuing eligibility for CPP disability benefits.

8-4. The goal of the Chesapeake Employers' Insurance Company (the state of Maryland's workers' compensation program) is for workers to achieve maximal clinical and functional improvement after a work-related illness or injury, to the point at which the condition is stabilized and no further recovery or improvement is expected, even with additional medical intervention. This goal, reflected in the program's mission to return individuals to work, is similar to those of other state workers' compensation programs, as well as private disability insurers, but differs from those of the federal programs examined. This difference necessitates a different approach to disability assessment in terms of functional and prognostic evaluation instruments and rehabilitation.

8-5. Prudential Financial, a private disability insurance company, has several goals when a claimant applies for assistance, including understanding all of the medical conditions impacting the employee, identifying functional capacity and restrictions and limitations, identifying prognosis and when and whether return to work is expected, and using vocational rehabilitation resources to facilitate a safe return to work. Most of the disability assessments are performed by company staff, which may include the claims manager, a registered nurse, and a vocational rehabilitation specialist.

8-6. Although many of the benefit programs examined by the committee collect functional information on physical and mental abilities, only the VBA, the CPP, and Sun Life Financial collect information on activities of daily living.

8-7. Only two of the benefit programs (the VBA and DSP) use forms tailored to specific types of impairments to collect functional information; the remainder (the CPP, Chesapeake, Washington State, Prudential, and Sun Life) use generic questionnaires.

Conclusions

8-1. The resources of the disability benefit programs examined for developing cases and gathering functional information differ dramatically. The programs vary greatly in size and have somewhat different goals, which can affect methods used and availability of staff to collect information. The willingness of both the private and public disability programs to share their methods and assessment instruments would help strengthen the processes both nationally and internationally.

8-2. In general, programs with a mission of rehabilitation and returning individuals to work (e.g., workers' compensation programs and private disability insurers) have a different approach to disability assessment and management relative to organizations charged solely with disability compensation (e.g., the VBA, the CPP, and DSP).

REFERENCES

Australian Government. 2011. *Social Security (tables for the assessment of work-related impairment for disability support pension) determination 2011*. https://www.dss.gov.au/sites/default/files/documents/05_2012/dsp_impairment_final_tables.pdf (accessed April 4, 2019).

Australian Government Department of Human Services. 2018a. *Disability support pension eligibility*. https://www.humanservices.gov.au/individuals/services/centrelink/disability-support-pension/eligibility (accessed April 4, 2019).

Australian Government Department of Human Services. 2018b. *Disability support pension: How we assess your disability or condition*. https://www.humanservices.gov.au/individuals/services/centrelink/disability-support-pension/eligibility/how-we-assess-your-disability-or-condition (accessed April 4, 2019).

Australian Government Department of Human Services. 2018c. *Disability support pension: Medical evidence*. https://www.humanservices.gov.au/individuals/services/centrelink/disability-support-pension/claiming/medical-evidence (accessed April 4, 2019).

Australian Government Department of Human Services. 2018d. *Disability support pension: Participation requirements*. https://www.humanservices.gov.au/individuals/services/centrelink/disability-support-pension/participation-and-reporting/participation-requirements (accessed April 4, 2019).

Australian Government Department of Human Services. 2018e. *Employment services assessments*. https://www.humanservices.gov.au/individuals/enablers/employment-services-assessments/37496 (accessed April 4, 2019).

Australian Government Department of Human Services. 2018f. *Payment rates for disability support pension*. https://www.humanservices.gov.au/individuals/enablers/payment-rates-disability-support-pension/39881 (accessed April 4, 2019).

Australian Government Department of Human Services. 2018g. *2017–18 Annual report*. https://www.humanservices.gov.au/sites/default/files/2018/10/8802-1810-annual-report-web-2017-2018.pdf (accessed April 4, 2019).

Chesapeake Employers' Insurance. 2017. *2017 Annual report*. https://www.ceiwc.com/pdf/2018/2018%20Annual%20Report/Chesapeake%20Employers%202017%20Annual%20Report%20Rev.%20PDF.pdf (accessed April 4, 2019).

Chesapeake Employers' Insurance. 2018a. *About Chesapeake Employers' Insurance Company*. https://www.ceiwc.com/about-chesapeake/company-overview/index.html (accessed April 4, 2019).

Chesapeake Employers' Insurance. 2018b. *Recent Chesapeake Employers success cases at the Maryland Workers' Compensation Commission*. https://www.ceiwc.com/i-am-an-employer/forms-and-publications-forms-pdfs/Articles%20Library/2011/Recent%20Cases%20-%20December%20%202011.html (accessed April 4, 2019).

Chesapeake Employers' Insurance. n.d.-a (unpublished). *Functional capacity evaluation summary*. Towson, MD: CEIWC.

Chesapeake Employers' Insurance. n.d.-b. *Making a difference: Health & claims services*. https://www.ceiwc.com/pdf/2018/Chesapeake%20Health%20&%20Claims%20Services%20Promotional%20Brochure%20Low%20Res%201-30-18....pdf (accessed April 4, 2019).

Chesapeake Employers' Insurance. n.d.-c (unpublished). *Sample job analysis*. Towson, MD: CEIWC.

Fisher, S., and W. I. Smulyan. 2018. *The process of disability management at Chesapeake Employers' Insurance Company: Workers' compensation perspective*. Presentation to the Committee on Functional Assessment for Adults with Disabilities, Washington, DC, April 19.

Government of Canada. 2017. *Annual report of the Canada Pension Plan for fiscal year 2016 to 2017*. https://www.canada.ca/content/canadasite/en/employment-social-development/programs/pensions/reports/annual-2017.html (accessed April 4, 2019).

Government of Canada. 2018a. *Canada Pension Plan—Overview*. https://www.canada.ca/en/services/benefits/publicpensions/cpp.html (accessed April 4, 2019).

Government of Canada. 2018b. *Canada Pension Plan adjudication framework*. https://www.canada.ca/en/employment-social-development/programs/disability/benefits/framework.html (accessed April 4, 2019).

Government of Canada. 2018c. *Canada Pension Plan disability benefit—Eligibility*. https://www.canada.ca/en/services/benefits/publicpensions/cpp/cpp-disability-benefit/eligibility.html (accessed April 4, 2019).

Government of Canada. 2018d. *Canada Pension Plan disability benefit—How much could you receive*. https://www.canada.ca/en/services/benefits/publicpensions/cpp/cpp-disability-benefit/benefit-amount.html (accessed April 4, 2019).

Government of Canada. 2018e. *Canada Pension Plan disability benefit—Overview*. https://www.canada.ca/en/services/benefits/publicpensions/cpp/cpp-disability-benefit.html (accessed April 4, 2019).

Government of Canada. 2018f. *Service Canada*. https://www.canada.ca/en/employment-social-development/corporate/portfolio/service-canada.html (accessed April 4, 2019).

Hamill, T. 2018. *Functional assessment of adults applying for disability benefits*. Presentation to the Committee on Functional Assessment for Adults with Disabilities, Washington, DC, April 19.

Kramschuster, J. 2018. *Functional assessment and RTW*. Presentation to the Committee on Functional Assessment for Adults with Disabilities, Washington, DC, April 19.

L&I (Washington State Department of Labor & Industries). 2007. *IME doctor's estimate of physical capacities*. https://www.lni.wa.gov/Forms/pdf/F242-387-000.pdf (accessed April 4, 2019).

L&I. 2014. *Job analysis summary.* https://www.lni.wa.gov/FormPub/Detail.asp?DocID=2657 (accessed April 4, 2019).

L&I. 2016. *Functional capacity summary.* https://www.lni.wa.gov/FormPub/Detail.asp?DocID=2803 (accessed April 4, 2019).

L&I. 2018a. *About Labor and Industries (L&I)*. https://www.lni.wa.gov/Main/AboutLNI/?F=MainFooter (accessed April 4, 2019).

L&I. 2018b. *Activity prescription form*. https://www.lni.wa.gov/forms/pdf/F242-385-000.pdf (accessed April 4, 2019).

L&I. 2018c. *Filing your claim*. https://www.lni.wa.gov/ClaimsIns/Claims/File/FilingClaim/default.asp (accessed April 4, 2019).

L&I. 2018d. *Independent medical exams*. https://www.lni.wa.gov/ClaimsIns/Claims/File/IME/default.asp (accessed April 4, 2019).

L&I. 2018e. *Injured? What you need to know: Approved claim*. https://www.lni.wa.gov/ClaimsIns/Claims/Guide/phase3.asp (accessed April 4, 2019).

L&I. 2018f. *Intro to L&I's workers' comp insurance*. https://www.lni.wa.gov/ClaimsIns/Insurance/Learn/Intro/Default.asp (accessed April 4, 2019).

L&I. 2018g. *L&I facts and figures*. https://www.lni.wa.gov/News/files/LNIFactsAndFigures.pdf (accessed April 4, 2019).

L&I. 2018h. *Medical examiners' handbook*. http://www.lni.wa.gov/IPUB/252-001-000.pdf (accessed April 4, 2019).

L&I. n.d.-a. *Completing the activity prescription form*. https://www.lni.wa.gov/ClaimsIns/Providers/Claims/ActivityRx/Completing.asp (accessed April 4, 2019).

L&I. n.d.-b. *Patient specific functional and pain scales (PSFS)*. https://www.lni.wa.gov/ClaimsIns/Files/OMD/IICAC/2018DocFuncImprovfunctionalscales.pdf (accessed April 5, 2019).

Prudential Financial. 2014a. *Kurzke functional systems scores (FSS)*. http://www.prudential.com/media/managed/GL.2012.238.pdf (accessed April 4, 2019).

Prudential Financial. 2014b. *Prudential visual capacity questionnaire*. http://www.prudential.com/media/managed/GL.2009.345.pdf (accessed April 4, 2019).

Prudential Financial. 2015a. *Behavioral health capacity questionnaire*. http://www.prudential.com/media/managed/GL.2015.021.pdf (accessed April 4, 2019).

Prudential Financial. 2015b. *Psychiatric work readiness assessment*. http://www.prudential.com/media/managed/GL.2015.038.pdf (accessed April 4, 2019).

Prudential Financial. 2016. *Mental status examination form*. http://www.prudential.com/media/managed/GL.2011.093.pdf (accessed April 4, 2019).

Prudential Financial. 2017. *Capacity questionnaire*. http://www.prudential.com/media/managed/GL.2012.240.pdf (accessed April 4, 2019).

Prudential Financial. 2018a. *About Prudential Financial*. https://www.prudential.com/about (accessed April 4, 2019).

Prudential Financial. 2018b. *Disability insurance and absence management: The numbers tell the story*. https://www.prudential.com/employers/group-insurance/products/disability-insurance (accessed April 4, 2019).

Prudential Financial. 2018c. *Resource center for physicians*. https://www.prudential.com/links/forms/group-insurance-physicians/?pswid=Z7_K8HEHHG0L884C0AEFNFT411G23&urile=wcm%3Apath%3A%2Fwps%2Fwcm%2Fconnect%2Fportal%20site%2Fz6_000000000000000000000000a0%2Fz6_9a0i18c0l0po80am2vkmno2g20%2Fz6_9a0i18c0lgos30a6cm62v52085%2Fz6_k8hehhg0lgsge0aer96u6e3k75%2Fz6_k8hehhg0lgojc0aefq2gbm28a5%2Fz6_k8hehhg0lo1m00al0noo2c1m52 (accessed April 4, 2019).

Service Canada. 2018. *Medical report for a Canada Pension Plan disability benefit*. https://catalogue.servicecanada.gc.ca/apps/EForms/pdf/en/ISP-2519.pdf (accessed April 4, 2019).

Sun Life Assurance Company of Canada. 2018. *Who we are*. http://www.sunlife.com/Global/About+us/Who+we+are?vgnLocale=en_CA (accessed April 4, 2019).

Sun Life Assurance Company of Canada. n.d. *Absence and disability management: Supporting optimum health management*. http://www.sunlife.ca/static/canada/Sponsor/About%20Group%20Benefits/Group%20benefits%20products%20and%20services/Absence%20and%20Disability%20Management/PDF6683-E.pdf (accessed April 4, 2019).

Sun Life Financial. n.d.-a. *Disability job demands questionnaire*. https://www.sunlife.ca/static/canada/Sponsor/About%20Group%20Benefits/Forms/Package-Standard-PS-STD-LTD-E_RE.pdf (accessed April 4, 2019).

Sun Life Financial. n.d.-b. *Long term disability activities of daily living questionnaire*. http://forms.sunlife-usa.com/onlineordering/get_file.cfm?form_id=28087 (accessed April 4, 2019).

Sun Life Financial. n.d.-c. *Long term disability plan member package: Attending physician's questionnaire claim for long-term disability benefits*. https://www.sunlife.ca/static/canada/Sponsor/About%20Group%20Benefits/Forms/PDF%20static%20files/Package-Standard-PM-LTD-E.pdf (accessed April 4, 2019).

Tugman, K., and J. Kramschuster. 2016. *The use of assistive technology in return-to-work*. Presentation to the Committee on the Use of Selected Assistive Products and Technologies in Eliminating or Reducing the Effects of Impairments, Washington, DC, September 27.

VA (U.S. Department of Veterans Affairs). 2011. *DBQ initial evaluation of residuals of traumatic brain injury (I-TBI) disability*. https://www.dcms.uscg.mil/Portals/10/CG-1/PSC/PSD/docs/VBA%20-%20Initial%20evaluation%20of%20TBI%20residuals.pdf?ver=2017-03-28-105323-973 (accessed April 4, 2019).

VA. 2017a. *Compensation*. https://www.benefits.va.gov/REPORTS/abr/docs/FY17-Compensation.pdf (accessed April 4, 2019).

VA. 2017b. *Veterans Benefits Administration annual benefits report fiscal year 2017*. https://www.benefits.va.gov/REPORTS/abr/docs/FY17-Introduction-Appendix.pdf (accessed April 4, 2019).

VA. 2018a. *Back (thoracolumbar spine) conditions disability benefits questionnaire*. https://www.vba.va.gov/pubs/forms/VBA-21-0960M-14-ARE.pdf (accessed April 4, 2019).

VA. 2018b. *Compensation: Benefit rates*. https://www.benefits.va.gov/compensation/rates-index.asp (accessed April 4, 2019).

VA. 2018c. *Compensation: Claims process*. https://www.benefits.va.gov/compensation/process.asp (accessed April 4, 2019).

VA. 2018d. *Compensation: Evidence requirements*. https://www.benefits.va.gov/compensation/evidence.asp (accessed April 4, 2019).

VA. 2018e. *Compensation: Frequently asked questions*. https://www.benefits.va.gov/COMPENSATION/dbq_FAQS.asp (accessed April 4, 2019).

VA. 2018f. *Compensation: Overview*. https://www.benefits.va.gov/compensation/dbq_overview.asp (accessed April 4, 2019).

VA. 2018g. *Heart conditions disability benefits questionnaire*. https://www.vba.va.gov/pubs/forms/VBA-21-0960A-4-ARE.pdf (accessed April 4, 2019).

VA. 2018h. *Mental disorders disability benefits questionnaire*. https://www.vba.va.gov/pubs/forms/VBA-21-0960P-2-ARE.pdf (accessed April 4, 2019).

VA. 2018i. *VA compensation and pension: Understanding mental health examinations*. https://www.benefits.va.gov/COMPENSATION/docs/claimexam-mentalhealth-exams.pdf# (accessed April 4, 2019).

CHAPTER 8 ANNEX TABLE BEGINS ON THE NEXT PAGE

ANNEX TABLE 8-1
Comparison of Selected Disability Benefit Programs

	Federal Disability Benefit Programs			
	Veterans Benefits Administration	Canada Pension Plan	Disability Support Pension (Australia)	
Size of Program	2017 Compensation program recipients: 4,964,209	2016–2017 5.6 million beneficiaries	2017–2018 104,000 claims processed and 31,000 claims granted	
Disability Focus	Personal (compensation for service-related disability)	Vocational (incapable regularly of pursuing any substantial gainful occupation because of a medical condition)	Vocational (disability or medical condition prevents working at least 15 hours per week in the next 2 years)	
Assessment Questionnaires and Forms				
Consider Physical Abilities	• Back Conditions Disability Benefits Questionnaire • Residuals of TBI Disability Benefits Questionnaire	• Medical report	• Impairment tables	
Consider Mental Abilities	• Mental Disorders Disability Benefits Questionnaire • Residuals of TBI Disability Benefits Questionnaire	• Medical report	• Impairment tables	
Consider Activities of Daily Living	• Residuals of TBI Disability Benefits Questionnaire	• Medical report		

	State Workers' Compensation Programs		Private Disability Insurance	
	Chesapeake Employers' Insurance Company	**Washington State Department of Labor and Industries**	**Prudential Financial**	**Sun Life Financial**
	2017 20,889 injured workers covered	2017 93,896 claims accepted	2015 short-term disability insurance: 2,500 clients long-term disability insurance: 3,100 clients	Unknown
	Vocational/return to work (inability to perform the duties of one's occupation because of injury)	Vocational/return to work (inability to perform the duties of one's occupation because of injury)	Vocational/return to work (inability to perform the duties of one's occupation because of injury)	Vocational/return to work (inability to perform the duties of one's occupation because of injury)
	• Functional capacity evaluation • Job analysis	• Activity Prescription Form • Independent medical examination (IME) doctor's estimate of physical capacities • Functional capacity summary • Job analysis summary	• Capacity Questionnaire • Visual Capacity Questionnaire • Kurtzke Functional Systems scores	• Attending Physician Statement • Job Demands Questionnaire
	• Functional capacity evaluation		• Mental Status Examination Form • Behavioral Health Capacity Questionnaire • Psychiatric Work Readiness Assessment • Kurtzke Functional Systems scores	• Attending Physician Statement • Activities of Daily Living Questionnaire • Job Demands Questionnaire
				• Activities of Daily Living Questionnaire

continued

ANNEX TABLE 8-1
Continued

	Federal Disability Benefit Programs			
	Veterans Benefits Administration	Canada Pension Plan	Disability Support Pension (Australia)	
Impairment Specific	• Back Conditions Disability Benefits Questionnaire • Heart Conditions Disability Benefits Questionnaire • Mental Disorders Disability Benefits Questionnaire • Residuals of TBI Disability Benefits Questionnaire		• Impairment tables	

NOTE: TBI = traumatic brain injury.

	State Workers' Compensation Programs		Private Disability Insurance	
	Chesapeake Employers' Insurance Company	Washington State Department of Labor and Industries	Prudential Financial	Sun Life Financial

9

Overall Conclusions[1]

This chapter presents overall conclusions derived from the chapter-specific findings and conclusions detailed throughout the report.

OVERALL CONCLUSIONS

The committee's chapter-specific findings and conclusions (some of which are highlighted in the next section) served as the basis for the five overall conclusions (see Box 9-1) presented in the following subsections.

Relationship of Functional Abilities to Work Participation

Current models of disability, such as the *International Classification of Functioning, Disability and Health* (ICF) model, consider disability to involve the effects (limitations) an individual's health condition places on his or her ability to function and participate fully in society. In keeping with these models, assessment of individuals' functional abilities relevant to work requirements is an important part of determining whether they are able to meet workplace demands and sustain work performance on a regular and continuing basis.

Numerous validated performance-based and self-report instruments are available to assess physical and mental functions and can be used to inform disability determination. However, as illustrated by the committee's

[1]This chapter does not include references. Citations to support the text and conclusions herein are provided in previous chapters of the report.

BOX 9-1
Overall Conclusions

1. Individuals' assessed functional abilities relevant to work requirements when assessed outside of actual work settings may be insufficient to establish their capacity to perform full-time work on a regular and continuing basis.
2. The validity of the results of work-related functional assessments is enhanced by a comprehensive approach that includes test results and other information about an individual's physical and mental functional abilities from multiple sources, as well as relevant social and environmental factors and the full scope of tasks involved in a job and sustained gainful employment.
3. Assessments that integrate information about impairments and abilities, including multiple tests of different types, repeated over time, provide the most useful information about work-related function.
4. Numerous challenges complicate accurate assessment of an individual's ability to work.
5. A number of factors, including age, gender, lower socioeconomic status, race, ethnicity, cultural group, and geographic location, may limit the quality and quantity of functional information available for a disability applicant.

conceptual framework (see Figure 2-3 in Chapter 2), it is a challenge to extrapolate from individuals' ability to perform specific activities and tasks to their ability to perform and sustain full-time work on a regular and continuing basis. Certain physical demands of jobs, such as sitting, standing, walking, lifting, and climbing, may relate more directly than mental/cognitive demands to activities that are amenable to functional assessment. For example, assessment of individuals' functional abilities with respect to adaptability and work-related personal interactions is more complicated than assessment of whether and how long an individual can sit, stand, or walk. Moving from assessment of individual functional abilities to the ability to perform tasks and meta-tasks as required for work participation creates challenges. Evaluation of the ability to perform a single work activity needs to reflect the context and practical relevance of an individual's being able to hold a job, taking account of personal and contextual (organizational and environmental) factors that influence individuals' capability to perform and sustain work. These include factors associated with an individual's health condition and its treatment that limit the ability to perform sustained work activities on a regular and continuing basis.

For these reasons, the committee drew the following conclusion:

1. **Individuals' assessed functional abilities relevant to work requirements when assessed outside of actual work settings may be**

insufficient to establish their capacity to perform full-time work on a regular and continuing basis.

- Assessment of functional abilities does not necessarily address an individual's capacity to perform tasks required for work participation. Although an individual may be capable of performing each activity separately, he or she may not be able to coordinate and sequence them effectively.
- While an individual may be able to perform work tasks successfully during a single assessment, he or she may be unable to perform required work tasks on a sustained or consistent (day-to-day) basis because of one or more underlying physical and/or mental health conditions.
- It is important to consider that testing is typically administered in a controlled, quiet environment without extraneous noise, social demands, and other factors that typically occur on a job, which, depending on the individual, can adversely affect the ability to perform work tasks.
- Factors associated with an individual's health condition (e.g., treatment demands, side effects) may limit the ability to participate in work on a regular and continuing basis even if the person is able to perform each of the tasks associated with a job.
- Similarly, environmental factors (e.g., physical [built and natural], social, and organizational) may limit an individual's ability to participate in work on a regular and continuing basis even if the person is able to perform the relevant work requirements.
- An individual's capacity to perform work requirements successfully in one specific work environment does not necessarily indicate the ability to perform the same work in a different setting.

Multiple Sources of Work-Related Functional Information

There are a variety of methods for collecting functional information (e.g., diagnostic testing, performance-based measures, self- or proxy-report measures), each of which has strengths and weaknesses, and the results of one are often used to validate those of another. Each method can yield instruments with satisfactory psychometric properties that allow their implementation in disability decision making. Numerous evidence-based self-report and performance-based measures of physical and mental function are available, although they may be limited by a number of factors, including an individual's underlying physical condition and cognitive status; the experience of pain, depression, or anxiety; and respondent bias or the person's level of effort. The use of validated instruments or test batteries that include validity measures can help testers determine the validity of the results obtained.

Another potential threat to the validity of assessments of functional abilities is use of measures in populations in whom they have not been validated. Self-report and performance measures of physical function provide complementary information, and together can be used to assess an individual's overall functional status, providing a more complete picture of whether or how well the individual will be able to perform everyday activities, including work, on a sustained basis than can be obtained with either type of measure alone. Third-party sources (e.g., friends and family members, health care and social service professionals, workplace colleagues and employers) who are sufficiently familiar with the applicant's activities, health, and functional status can be particularly helpful for providing ancillary information on health and behavioral matters, physical and mental functioning, and workplace performance, although such reports are at times influenced by such factors as self-interest, mixed motives, and partial or inaccurate observations. Combining and evaluating the convergence of information from different sources (e.g., self-reports, quantitative measures, medical records, consultative examinations) increases confidence in the validity of the information available for evaluating an individual's ability to work.

For these reasons, the committee drew the following conclusion:

2. **The validity of the results of work-related functional assessments is enhanced by a comprehensive approach that includes test results and other information about an individual's physical and mental functional abilities from multiple sources, as well as relevant social and environmental factors and the full scope of tasks involved in a job and sustained gainful employment.**
   - No single source of information is likely to provide all of the information needed to evaluate an individual's ability to work.
   - Professionals in multiple disciplines administer and interpret results of assessments for physical and mental function. Those with responsibility for repeated assessments may render more detailed and accurate evaluations of an individual's physical and/or mental functioning over time relative to medical specialists who have less frequent interactions with the person and less time per encounter during the same observation period.
   - Convergence of information from multiple sources increases confidence in its validity. It is important to combine and evaluate the consistency of information from different sources (e.g., self-reports, quantitative measures, medical records, consultative examinations) when evaluating an individual's ability to work.
   - Standardized self-report questionnaires are an important source of information regarding the nature and severity of an applicant's

functional limitations, especially when used in conjunction with other assessments.

- Qualitative data provided by applicants, family members, and other key sources who are sufficiently familiar with the applicant's activities, health, and functional status, in combination with review of medical evidence, complement quantitative information that serves as the basis for disability decisions.
- The use of measures based on item response theory that can be administered using computer adaptive testing can decrease respondent burden by reducing survey length and administration time while minimizing measurement error.

Integrated Assessment of Work-Related Functional Ability

Given that measuring function is complex and that work participation is a multidimensional construct, a single physical or mental assessment instrument, by itself, cannot provide a complete assessment of function. While specific assessment instruments measure physical and mental functional abilities at the impairment, body part, or organ system level, "integrated" assessment instruments that provide information regarding the integrated effect of individuals' impairments on general daily life and participation can capture the additive and sometimes multiplicative effects of multiple impairments and comorbid conditions on individuals' functional abilities. Several evidence-based instruments and instrument sets are available that provide integrated information about individuals' overall functional capabilities and limitations and could provide helpful information for determinations of work disability. The most informative evaluations of function may include integrated assessments in addition to specific assessments of body structures and systems.

The Work Disability Functional Assessment Battery (WD-FAB) is a new instrument developed to assess physical and mental functional abilities relevant to work requirements. It may be most useful for understanding self-reported physical function. The Patient-Reported Outcomes Measurement Information System (PROMIS), Quality of Life in Neurological Disorders (Neuro-QoL), and the National Institutes of Health (NIH) Toolbox also may be useful in understanding the functioning of an applicant. Currently, there is no evidence to support drawing direct inferences from the scores of these instruments with respect to employability.

Professionals with responsibility for repeated assessments using standardized assessment tools and procedures may render more detailed and accurate evaluations of an individual's physical and/or mental functioning over time relative to medical specialists who have less frequent interactions with the person and less time per encounter during the same observation period. Understanding the relationship between chronic illness and functioning is important because some major illnesses are episodic in nature, with severity of symptoms and functional impairments varying over time, and with periods of greater severity ranging from weeks to months.

For these reasons, the committee drew the following conclusion:

3. **Assessments that integrate information about impairments and abilities, including multiple tests of different types, repeated over time, provide the most useful information about work-related function.**
   - Numerous validated tests are available for measuring physical and mental functional abilities at the impairment and body or organ system level.
   - No single tool, by itself, can reliably and consistently determine the inability or ability to work.
   - Available instruments, whether based on performance, self-report, or third-party sources, are useful individually, but their value may be increased when different types of instruments are combined to provide a fuller picture of an individual's ability or inability to sustain work on a regular and continuing basis, especially when they can be repeated over time.
   - Integrated assessment measures are useful for capturing the additive and sometimes multiplicative effects of multiple impairments and comorbid conditions on an individual's functional ability to meet work requirements.

Challenges for Assessment of Work-Related Functional Abilities

The committee's conceptual framework for assessing work capacity (see Figure 2-3 in Chapter 2) demonstrates the complexity and challenges of functional assessments, especially the use of instruments that assess only body and structure function or impairment, in moving from individuals' ability to perform specific activities and tasks to their capacity to perform and sustain full-time work on a regular and continuing basis. In addition, there are a number of threats to the validity of assessments of functional abilities, including testing of maximal versus typical performance, assessment of episodic activity versus sustained task performance, absence of standardized testing conditions, mixed-motive incentives, compromised test integrity owing to prior use of the test in low-stakes testing applications, and diverse test populations on whom tests may not have been validated. Symptoms associated with depression (e.g., fatigue, difficulty concentrating, and slowed response speed) can impair functioning and frequently compound work-related functional limitation in the context of other primary impairments. It is important to collect information about the nature and original purpose of an assessment instrument as well as the conditions and context in which it was administered to help in understanding the results with respect to potential limitations to their generalizability.

For these reasons, the committee drew the following conclusion:

4. **Numerous challenges complicate accurate assessment of an individual's ability to work, including the following:**
   - Measures of physiological, morphological, psychological, or cognitive severity (e.g., laboratory findings, signs, or symptoms of impairments) may not correlate with the severity of functional limitations (i.e., the effect of a condition on an individual's ability to work or conduct daily life).
   - It is simpler to demonstrate inability or limitation to perform a specific activity (e.g., reaching overhead, climbing a ladder) than to demonstrate an individual's ability to perform the combination of activities required for different occupations. Tests of functional abilities often do not measure whether an individual is able to combine functions to perform tasks as needed for work.
   - Successful work performance is more than the sum of the specific tasks and skills required, and the overall limitation to successful work for an individual is often more than the sum of single impairments.
   - Threats to the validity of assessments of functional abilities include testing of maximal versus typical performance, assessment of episodic activity versus sustained task performance, absence of standardized testing conditions, mixed-motive incentives, compromised test integrity owing to prior use of the test in low-stakes testing applications, and diverse test populations on whom tests may not have been validated.
   - Symptoms associated with psychological conditions such as depression and anxiety can affect a person's ability to manage one or more limitations in a work setting. Therefore, it is necessary to consider them when assessing an individual's ability to sustain work on a regular and continuing basis because a person's capacity to work may be overestimated if a psychological comorbidity is present.

Factors Limiting the Quality and Quantity of Information on Functional Ability for an Applicant

When evaluating the utility of a functional assessment instrument for informing disability determinations, it is important to consider the instrument's performance across multiple subgroups (e.g., age, gender, socioeconomic status, race, ethnicity, cultural group) as a principle of ethical practice. Numerous instruments are available for assessing physical and mental functions, but not all account for the range of cultural, linguistic, or literacy factors among the population being assessed. Differences in gender, race, ethnicity, and culture can affect individuals' perceptions of illness and their reporting of health information. Development and validation of patient-reported symptom measures and clinician/observer-rendered assessments vary in the extent to which they have been tested or adapted across diverse racial/ethnic and cultural populations. Cross-cultural adaptations and validations of assessments in different cultural contexts and languages are predicated on the notion that such efforts take into account distinct groups' experiences and meanings of health, behaviors, illness, symptoms, disability, and help-seeking behaviors. Assessment instruments developed for use in research applications may not account for cultural, linguistic, or literacy factors, such as limited English proficiency or low literacy, that limit access to such assessments. Consequently, few or no assessments are available that can capture valid and reliable administration and scoring information for these populations.

In addition, the extent and types of medical evidence in an applicant's file likely will be affected by the availability and cost of tests. Health care data relevant to disability determinations, such as the results of specific, expensive tests (e.g., certain cardiovascular tests and psychological test batteries) that are valid and potentially useful, may not be readily available because an individual may be uninsured or underinsured, or the tests may be denied by an insurance plan because they are not deemed medically necessary. Health disparities can have a significant effect on the availability of health information to inform disability determinations. Disability applicants who are uninsured or underinsured are less likely to have a well-developed body of health data, including the results of expensive, specialized tests, to demonstrate evidence of disability. Disparities in access to care and consequently health outcomes can affect not only the quantity of tests conducted in the context of disability determinations but also the quality of the tests and resulting information. Access to health care professionals, including those with expertise in providing information relevant to disability determination, often is limited by lower socioeconomic status and/or geographic location. Acquisition of an applicant's clinical records may be difficult for several reasons: providers' fear of sharing confidential information, the limited capacity of a provider's organization to gather and transmit records, and high administrative costs for record transfer.

For these reasons, the committee drew the following conclusion:

5. **A number of factors, including age, gender, lower socioeconomic status, race, ethnicity, cultural group, and geographical location, may limit the quality and quantity of functional information available for a disability applicant.**
   - Functional assessment instruments vary in the extent to which they have been tested or adapted across diverse populations, making it important to consider an instrument's performance across multiple subgroups.
   - Assessment instruments developed for research applications may not account for cultural, linguistic, or literacy factors that influence access to such assessments (i.e., no assessments available for people with limited English proficiency or those with low literacy).
   - Lower socioeconomic status is associated with less access to high-quality care and health care providers with expertise in providing information relevant to disability determination.
   - A number of additional factors limit available information, including cost and administrative challenges (e.g., costs of tests, assembling and sharing medical records).

SUPPORTING EVIDENCE FOR THE COMMITTEE'S OVERALL CONCLUSIONS

Box 9-2 shows the links between the overall conclusions presented above and some of the most relevant chapter-specific findings and conclusions that support them.[2]

[2]Not all of the committee's chapter-specific findings and conclusions are included in Box 9-2. Those that are included are numbered according to the chapter in which they appear.

BOX 9-2
Overall Conclusions and Supporting Evidence

1. **Individuals' assessed functional abilities relevant to work requirements when assessed outside of actual work settings may be insufficient to establish their capacity to perform full-time work on a regular and continuing basis.**

Findings

2-4. Although the worker abilities in the Occupational Information Network (O*NET) and the physical and proposed cognitive demands collected in the Occupational Requirements Survey (ORS) may be affected by physical or mental impairments and are to some extent amenable to functional assessment, many instruments used to assess function do not necessarily correlate with individuals' ability to perform work-related activities. In addition, certain physical demands of jobs, such as sitting, standing, walking, lifting, and climbing, may correlate more directly than mental/cognitive demands with activities that are amendable to functional assessment.

2-5. Assessment of individuals' functional abilities with respect to adaptability and work-related personal interactions is more complicated than assessment of whether and how long an individual can sit, stand, or walk.

2-6. Extrapolation from assessment of functional abilities ("activities" in ICF parlance) to the ability to perform tasks or meta-tasks as required for work participation is a challenge.

2-8. The committee's conceptual framework includes "interrupters," factors associated with an individual's health condition and its treatment that limit the ability to perform sustained work activities on a regular and continuing basis.

2-9. Assessment of the capacity of an individual to work and to sustain full-time work on a regular and continuing basis encompasses many factors that often go beyond whether the person can complete specific individual physical and mental activities or tasks.

Conclusions

2-1. In keeping with current models of disability, assessment of individuals' functional abilities relevant to work requirements is an important part of determining whether they are able to meet workplace demands and sustain work performance on a regular and continuing basis.

2-3. The committee's conceptual framework for assessing work capacity demonstrates the complexity and challenges of functional assessments, especially the use of instruments that assess only body and structure function or impairment, in extrapolating from individuals' ability to perform specific activities and tasks to their capacity to perform work and to sustain full-time work on a regular and continuing basis.

4-6. Stronger evidence is needed to link activities of daily living (ADLs) and instrumental activities of daily living (IADLs) performance to work capacity, perhaps by comparing ADL and IADL performance among applicants who are awarded Social Security Disability Income benefits versus those who are denied.

5-1. Given the complexity of measuring physical function and the multidimensional nature of work participation, no single instrument has yet been demonstrated to provide a comprehensive assessment of an individual's physical functional abilities relevant to work.

2. The validity of the results of work-related functional assessments is enhanced by a comprehensive approach that includes test results and other information about an individual's physical and mental functional abilities from multiple sources, as well as relevant social and environmental factors and the full scope of tasks involved in a job and sustained gainful employment.

Findings

2-7. To capture the context of work, the committee's conceptual framework for functional assessment for work adds the hierarchy of job and task analyses between function and work and takes account of personal and contextual (organizational and environmental) factors that influence individuals' capacity to perform sustained work activities.

2-9. Assessment of the capacity of an individual to work and to sustain full-time work on a regular and continuing basis encompasses many factors that often go beyond whether the person can complete specific individual physical and mental activities or tasks.

3-3. The validity of functional assessment tests is enhanced when the test users administer them for the purpose and in the context for which they were designed (e.g., target population).

3-7. The use of instruments or test batteries that include validity measures can help testers determine the validity of the results obtained.

3-8. Third-party sources (e.g., friends and family members, health care and social service professionals, workplace colleagues and employers) who are suitably familiar with the applicant's activities, health, and functional status can be particularly helpful in providing ancillary information on health and behavioral matters, physical and mental functioning, and workplace performance, sometimes supported by written documents. Such reports are at times influenced by such factors as self-interest, mixed motives, or inaccurate observations. Tests assessing beliefs, attitudes, moods, and other internal states are not suitable for proxy respondents.

4-1. Specific assessment instruments measure physical and mental functional abilities at the impairment, body part, or organ system level. Integrated assessments can capture the additive and sometimes multiplicative effects of multiple impairments and comorbid conditions on individuals' functional abilities.

5-1. Self-report and performance-based measures provide different perspectives on physical functional ability.

continued

BOX 9-2 Continued

5-7. There are multiple functional capacity evaluation (FCE) instruments with varying degrees of reliability and validity. No single FCE instrument has proven superior for determining an individual's functional ability. The reliability and validity of FCEs can reflect a variety of confounders, including assessors' training; nonstandard testing environments; and examinees' effort, cooperation, and interest in returning to work. Assessors' estimate of the examinee's level of effort can enhance the accuracy of test results.

Conclusions

4-1. The most informative evaluations of function may include both specific assessments of body structures and systems and integrated assessments that describe the effects of multiple impairments and comorbid conditions.

4-7. The utility of information about ADLs and IADLs in the context of disability determination may be enhanced by asking additional questions about context; environmental factors, including use of assistive technologies; required assistance; and the effect of performing ADLs and IADLs on pain, fatigue, confusion, concentration, and other physical or cognitive factors that can interfere with work performance.

4-8. Evidence-based instruments and sets of instrument that provide integrated information about individuals' overall functional capabilities and limitations could provide helpful information for determinations of work disability.

5-1. Given the complexity of measuring physical function and the multidimensional nature of work participation, no single instrument has yet been demonstrated to provide a comprehensive assessment of an individual's physical functional abilities relevant to work.

6-3. There is no single measure that captures all important aspects of mental abilities needed for work, although the WD-FAB, as a self-report battery of relevant questions, shows promise. More development work is needed for the WD-FAB to fulfill its promise for use in disability determination.

7-1. Consideration of age and comorbidities is critically important in both the evaluation and the likely trajectory of common conditions and their effects on work-related impairment.

3. Assessments that integrate information about impairments and abilities, including multiple tests of different types, repeated over time, provide the most useful information about work-related function.

Findings

3-5. Direct performance testing of physical and neurocognitive functional abilities is well developed and typically is used to assess common disease-specific deficits and monitor functional increments or decrements over time. Such testing may be useful for tracking the progress of those diseases, but they are not necessarily generalizable to other disabling conditions.

3-8. Third-party sources (e.g., friends and family members, health care and social service professionals, workplace colleagues and employers) who are suitably familiar with the applicant's activities, health, and functional status can be particularly helpful in providing ancillary information on health and behavioral matters, physical and mental functioning, and workplace performance, sometimes supported by written documents. Such reports are at times influenced by such factors as self-interest, mixed motives, or inaccurate observations. Tests assessing beliefs, attitudes, moods, and other internal states are not suitable for proxy respondents.

4-1. Specific assessment instruments measure physical and mental functional abilities at the impairment, body part, or organ system level. Integrated assessments can capture the additive and sometimes multiplicative effects of multiple impairments and comorbid conditions on individuals' functional abilities.

5-7. There are multiple FCE instruments with varying degrees of reliability and validity. No single FCE instrument has proven superior for determining an individual's functional ability. The reliability and validity of FCEs can reflect a variety of confounders, including assessors' training; nonstandard testing environments; and examinees' effort, cooperation, and interest in returning to work. Assessors' estimate of the examinee's level of effort can enhance the accuracy of test results.

6-1. It is important to assess the persistence of impairment due to mental disorders, given the possibility of episodic or persistent symptoms.

Conclusions

3-2. Professionals with responsibility for repeated assessments using standardized assessment tools and procedures may render more detailed and accurate evaluations of an individual's physical and/or mental functioning over time relative to medical specialists who have less frequent interactions with the person and less time per encounter during the same observation period.

3-5. When evaluating the utility of a functional assessment instrument for informing disability determinations, it is important to consider the instrument's performance across multiple subgroups (e.g., age, gender, socioeconomic status, race, ethnicity, cultural group) as a principle of good practice.

4-1. The most informative evaluations of function may include both specific assessments of body structures and systems and integrated assessments that describe the effects of multiple impairments and comorbid conditions.

4-8. Evidence-based instruments and sets of instruments that provide integrated information about individuals' overall functional capabilities and limitations could provide helpful information for determinations of work disability.

continued

BOX 9-2 Continued

6-2. Understanding the relationship between mental illness and functioning is important because some major mental illnesses are episodic in nature, with severity of symptoms and functional impairments varying over time, and with periods of greater severity ranging from weeks to months.

4. Numerous challenges complicate accurate assessment of an individual's ability to work.

Findings

` 2-5. Assessment of individuals' functional abilities with respect to adaptability and work-related personal interactions is more complicated than assessment of whether and how long an individual can sit, stand, or walk.

2-6. Extrapolation from assessment of functional abilities ("activities" in ICF parlance) to the ability to perform tasks or to meta-task as required for work participation is a challenge.

2-7. To capture the context of work, the committee's conceptual framework for functional assessment for work adds the hierarchy of job and task analyses between function and work and takes account of personal and contextual (organizational and environmental) factors that influence individuals' capacity to perform sustained work activities.

2-8. The committee's conceptual framework includes "interrupters," factors associated with an individual's health condition and its treatment that limit the ability to perform sustained work activities on a regular and continuing basis.

3-4. Assessment instruments developed for use in research and training settings may not account for cultural, linguistic, or literacy factors, such as limited English proficiency or low literacy, that can limit access to such assessments.

3-6. The accuracy of self-reported information can be affected, intentionally or unintentionally, by the respondent, who may either under- or overestimate his or her ability to perform different tasks.

3-9. Threats to the validity of assessments of functional abilities include testing of maximal versus typical performance, assessment of episodic activity versus sustained task performance, absence of standardized testing conditions, mixed-motive incentives, compromised test integrity owing to prior use of the test in low-stakes testing applications, and diverse test populations in whom tests may not have been validated.

4-5. Depression can limit performance of ADLs or IADLs irrespective of physical or cognitive impairments or age.

4-8. Research is limited on the relationship between assessments of ADL and IADL performance and an individual's ability to return to work.

6-13. Symptoms associated with depression, including fatigue, difficulty concentrating, and slowed response speed, can impair work functioning.

Conclusions

2-3. The committee's conceptual framework for assessing work capacity demonstrates the complexity and challenges of functional assessments, especially the use of instruments that assess only body and structure function or impairment, in extrapolating from individuals' ability to perform specific activities and tasks to their capacity to perform work and to sustain full-time work on a regular and continuing basis.

3-4. It is important to collect information about the nature and original purpose of an assessment instrument, as well as the conditions and context in which it was administered, to help in understanding the results with respect to potential limitations on their generalizability.

5. **A number of factors, including age, gender, lower socioeconomic status, race, ethnicity, cultural group, and geographic location, may limit the quality and quantity of functional information available for a disability applicant.**

Findings

3-4. Assessment instruments developed for use in research and training settings may not account for cultural, linguistic, or literacy factors, such as limited English proficiency or low literacy, that can limit access to such assessments.

3-14. Lower socioeconomic status is associated with less access to high-quality care and health care professionals, including those with expertise in providing information relevant to disability determination.

3-15. Patient-reported symptom measures and clinician/observer-rendered assessments vary in the degree to which they have been tested or adapted across diverse racial, ethnic, and cultural populations.

Conclusions

3-5. When evaluating the utility of a functional assessment instrument for informing disability determinations, it is important to consider the instrument's performance across multiple subgroups (e.g., age, gender, socioeconomic status, race, ethnicity, cultural group) as a principle of good practice.

3-6. Disparities in access to care and health outcomes can affect not only the quantity of assessments conducted in the context of disability determinations but also the quality of the assessments that are conducted and the resulting information.

Appendix A

Public Session Agendas

MEETING 1: PUBLIC SESSION

Hosted by the Committee on Functional Assessment for Adults with Disabilities

December 7, 2017

Keck Center of the National Academies
Room 208
500 Fifth Street, NW
Washington, DC 20001

Agenda

10:15 a.m. **Welcome and Introductions**
Paul Volberding, M.D., Committee Chair
Gina Clemons, Associate Commissioner, Office of Disability Policy, Social Security Administration (SSA)
Melissa Spencer, Deputy Associate Commissioner, Office of Disability Policy, SSA

10:30 a.m. **Social Security Administration Presentations Relevant to the Committee's Task**

- **Background and basis for the task order**
  Mary Beth Rochowiak, Policy Analyst, Office of Vocational Policy, SSA
- **Adult sequential evaluation process and functional assessment**
  Mary Beth Rochowiak, Policy Analyst, Office of Vocational Policy, SSA
  Megan Butson, Policy Analyst, Office of Vocational Policy, SSA
  Joanna Firmin, Supervisor, Office of Medical Policy, SSA

11:15 a.m. **Discussion of Tasks #1 and #3 from the Statement of Task**
Committee Members and SSA Staff

12:15 p.m. **Break for Lunch**

1:00 p.m. **Social Security Administration Presentations Relevant to the Committee's Task**

- **Occupational Information Network (O*NET), Job Zone 1 Occupations, SSA's Occupational Information System Project, and the Bureau of Labor Statistics' Occupational Requirements Survey**
  Deborah Harkin, Senior Policy Advisor, SSA

1:45 p.m. **Discussion of Task #2 from the Statement of Task**
Committee Members and SSA Staff

2:25 p.m. **Summary and Closing Remarks**
Paul Volberding, M.D., Committee Chair

2:30 p.m. **Adjourn**

MEETING 2: PUBLIC SESSION

Hosted by the Committee on Functional Assessment for Adults with Disabilities

February 26, 2018

Keck Center of the National Academies
Room 106
500 Fifth Street, NW
Washington, DC 20001

Agenda

9:30 a.m. **Welcome and Introductions**
Paul Volberding, M.D., Committee Chair

9:40 a.m. **Functional Assessment of Physical Abilities Relevant to Work Requirements**
Susan J. Isernhagen, PT, Co-Director, DSI Limited LLC

10:25 a.m. **Break**

10:40 a.m.	**Functional Assessment of Noncognitive Mental Abilities Relevant to Work Requirements** *Howard H. Goldman, M.D., Ph.D., Professor, Department of Psychiatry, University of Maryland School of Medicine*
11:25 a.m.	**Functional Assessment of Cognitive Mental Abilities Relevant to Work Requirements** *Philip D. Harvey, Ph.D., Professor of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine*
12:00 p.m.	**Discussion**
12:30 p.m.	**Break for Lunch**
1:15 p.m.	**Patient-Reported Outcomes Measurement Information System (PROMIS) and NIH Toolbox** *Richard C. Gershon, Ph.D., Vice Chair for Research in Medical Social Sciences and Professor of Medical Social Sciences and Preventive Medicine-Health and Biomedical Informatics, Northwestern University Feinberg School of Medicine*
2:30 p.m.	**Break**
2:40 p.m.	**Work Disability Functional Assessment Battery (WD-FAB)** *Leighton Chan, M.D., M.P.H., Chief, Rehabilitation Medicine Department, National Iinstitutes of Health Clinical Center*
3:40 p.m.	**Discussion with Stakeholder Representatives** *Marty Ford, J.D., Senior Executive Officer, Public Policy, The Arc* *Kate Lang, J.D., Senior Staff Attorney, Justice in Aging* *Kevin Liebkemann, J.D., Chief Section Counsel at Legal Services of New Jersey* *Barbara Silverstone, J.D., Executive Director, National Organization of Social Security Claimants' Representatives (NOSSCR)*
4:55 p.m.	**Closing Remarks** *Paul Volberding, Committee Chair*

5:00 p.m. **Adjourn**

MEETING 3: PUBLIC SESSION

Hosted by the Committee on Functional Assessment for Adults with Disabilities

April 19, 2018

National Academy of Sciences Building
Board Room
2101 Constitution Avenue, NW
Washington, DC 20418

Agenda

9:00 a.m. **Welcome and Introductions**
Paul Volberding, M.D., Committee Chair

9:10 a.m. **Functional Assessment of Depression and Anxiety Disorders Relevant to Work Requirements**
Paul S. Appelbaum, M.D., Elizabeth Dollard Professor of Psychiatry, Medicine, and Law, Columbia University

10:00 a.m. **Break**

10:15 a.m. **Veterans Benefits Administration**
Robert (Mike) Carr, Deputy Director, Benefits Assistance Service, Veterans Benefits Administration
Jacqueline Imboden, Policy Lead Analyst, Compensation Service, Veterans Benefits Administration

11:15 a.m. **Prudential Financial**
John Kramschuster, Director, Vocational Services, Prudential Financial

12:15 p.m. **Break for Lunch**

1:15 p.m. **Chesapeake Employers' Insurance Company**
Stephen N. Fisher, M.D., Ph.D., Director of Health Services, Chesapeake Employers' Insurance Company
William I. Smulyan, M.D., FAAOS, Medical Director, Chesapeake Employers' Insurance Company

2:15 p.m.	**Break**
2:30 p.m.	**Sun Life Financial** *Tracy Hamill, M.D., Assistant Vice President and Medical Director, Clinical Claims, Sun Life Financial*
3:30 p.m.	**Closing Remarks** *Paul Volberding, M.D., Committee Chair*
3:35 p.m.	**Adjourn**

TELECONFERENCE WITH REPRESENTATIVES FROM THE NATIONAL ASSOCIATION OF DISABILITY EXAMINERS (NADE)

Hosted by the Committee on Functional Assessment for Adults with Disabilities

May 23, 2018

Agenda

1:00 p.m.	**Opening Remarks** *Paul Volberding, M.D., Committee Chair*
	Discussion with NADE representatives *Jennifer Pounds, NADE President* *Sharon Bland-Brady, NADE President-elect* *Jeff Price, NADE Legislative Director*
2:00 p.m.	**Adjourn**

Appendix B

Glossary

Abilities: "Enduring attributes of the individual that influence performance" (O*NET, 2019a).

Activities: Actions or tasks performed by an individual, such as walking, lifting, keyboarding, or problem solving.

Activities of daily living: Basic tasks of daily life that typically include personal care and hygiene, dressing, feeding, continence management, and mobility.

Adaptability:
(1) *Occupational Requirements Survey* definition: "'Adaptability' measures characteristics of an occupation that cause a worker to adjust to changes in work routines," including work tasks, work schedule, and work location (DOL, 2017, p. 61).
(2) *O*NET* definition: "Adaptability/Flexibility— job requires being open to change (positive or negative) and to considerable variety in the workplace" (O*NET, 2019b).
(3) U.S. Social Security Administration (SSA) definition: Adapt or manage oneself—"This area of mental functioning refers to the abilities to regulate emotions, control behavior, and maintain well-being in a work setting. Examples include: responding to demands; adapting to changes; managing your psychologically based symptoms; distinguishing between acceptable and unacceptable work performance; setting realistic goals; making plans for yourself independently of others; maintaining personal hygiene and

attire appropriate to a work setting; and being aware of normal hazards and taking appropriate precautions" (Mental Disorders Listings Paragraph B Criteria, Paragraph B4 [SSA, n.d.-b]).

Body functions: "The physiological functions of body systems, including psychological functions. 'Body' refers to the human organism as a whole, and thus includes the brain. Hence, mental (or psychological) functions are subsumed under body functions. The standard for these functions is considered to be the statistical norm for humans" (WHO, 2001, p. 213).

Body structures: "The structural or anatomical parts of the body such as organs, limbs and their components classified according to body systems. The standard for these structures is considered to be the statistical norm for humans" (WHO, 2001, p. 213).

Capability: "The quality or state of being capable"; see also "Abilities" (*Merriam-Webster*, 2019a).

Capacity: "An individual's ability to execute a task or an action" (WHO, 2001, p. 123).

Cognitive test: "Standardized measure of task performance used to assess cognitive functioning (e.g., intellectual capacity, attention and concentration, processing speed, language and communication, visual-spatial abilities, memory)" (IOM, 2015, p. 223).

Disability:
(1) *International Classification of Functioning, Disability and Health* (ICF) definition: "An umbrella term for impairments, activity limitations and participation restrictions. It denotes the negative aspects of the interaction between an individual (with a health condition) and that individual's contextual factors (environmental and personal factors)" (WHO, 2001, p. 213).
(2) SSA definition: In adults, "the inability to engage in any substantial gainful activity … by reason of any medically determinable physical or mental impairment(s) which can be expected to result in death or which has lasted or can be expected to last for a continuous period of not less than 12 months" (SSA, n.d-a).

Executive function: "How well a person can plan, prioritize, organize, sequence, initiate, and execute multi-step procedures" (OIDAP, 2009, p. C-22).

Functional limitation: "A loss or restriction of an individual's ability to perform a specific physical or mental function or activity, such as walking, speaking, memory, and the like" (IOM, 2015, p. 224).

Functional severity: "The impact of [a] disorder on an individual's ability to perform age-appropriate activities, irrespective of illness type and under a broad range of circumstances.... Functional severity reflects the effect of a condition on a final common pathway—ability to conduct daily life" (Stein et al., 1987).

Functioning: "An umbrella term encompassing all body functions, activities, and participation" (WHO, 2001, p. 3).

Impairment: "A loss or abnormality in body structure or physiological function (including mental functions). Abnormality here is used strictly to refer to a significant variation from established statistical norms (i.e., as a deviation from a population mean within measured standard norms) and should be used only in this sense" (WHO, 2001, p. 213).

Instrumental activities of daily living: Tasks that are considered to be more complex than "activities of daily living" and relate to independent living in the community, such as navigating transportation options and shopping, preparing meals, managing one's household, managing finances and medications, communicating with others, and providing companionship and mental support.

Medically determinable impairment: "A medically determinable physical or mental impairment is an impairment that results from anatomical, physiological, or psychological abnormalities that can be shown by medically acceptable clinical and laboratory diagnostic techniques. The medical evidence must establish that an individual has a physical or mental impairment; a statement about the individual's symptoms is not enough" (SSA, n.d.-a).

Neuropsychological tests: Performance-based tests used to measure various aspects of an individual's cognitive functioning, including "memory, attention, processing speed, reasoning, judgment, and problem-solving, spatial, and language functions" (Harvey, 2012, p. 91).

Noncognitive measure: "Standardized self-report measure that assesses noncognitive psychological complaints" (IOM, 2015, p. 224).

Participation: "A person's involvement in a life situation. It represents the societal perspective of functioning" (WHO, 2001, p. 213).

Performance: "The execution of an action" (*Merriam-Webster*, 2019c).

Performance validity: The validity of actual ability task performance; often referred to as effort in the literature (Larrabee, 2012, 2014).

Performance validity test: "Stand-alone or embedded/derived measures used to assess whether an examinee is performing at a level consistent with his/her actual abilities" (IOM, 2015, p. 225; adapted from Larrabee, 2014).

Performance-based measure: Requires that the individual being assessed perform a set of functional tasks so that his or her ability to execute them can be ascertained. Examples of such measures include assessments of gait, balance, and lifting in the physical realm and cognition in the mental realm.

Psychological assessment: "The comprehensive integration of information from a variety of sources—including formal psychological tests, informal tests and surveys, structured clinical interviews, interviews with others, school and/or medical records, and observational data—to make inferences regarding the mental or behavioral characteristics of an individual or to predict behavior" (IOM, 2015, p. 225; adapted from Hubley and Zumbo, 2013, p. 3).

Psychological testing: "The use of formal, standardized procedures for sampling behavior that ensure objective evaluation of the test-taker regardless of who administers the test (Furr and Bacharach, 2013; Hubley and Zumbo, 2013). Major categories of psychological tests include (1) intelligence tests, (2) neuropsychological tests, (3) personality tests, (4) clinical or diagnostic tests (e.g., depression, anxiety), (5) achievement tests, (6) aptitude tests, and (7) occupational or interests tests" (IOM, 2015, p. 225).

Psychometrics: "The scientific study, including the development, interpretation, and evaluation, of psychological tests and measures used to assess variability in behavior and link such variability to psychological phenomena" (IOM, 2015, p. 225; adapted from Furr and Bacharach, 2013, pp. 9–10; Hubley and Zumbo, 2013, p. 3).

Rehabilitation: "The physical restoration of a sick or disabled person by therapeutic measures and reeducation to participation in the activities of a normal life within the limitations of the person's physical disability"; "the process of restoring an individual to a useful and constructive place in society especially through some form of vocational, correctional, or therapeutic retraining" (*Merriam-Webster*, 2019b).

Reliability: "The consistency of scores across replications of a measurement procedure" (Brennan, 2006, p. 3).

Residual functional capacity: "The most [an applicant] can still do despite [his or her impairment-related] limitations" or restrictions on "a regular and continuing basis," currently defined as 5 days per week, 8 hours per day, or an equivalent work schedule (20 CFR 404.1545; 20 CFR 416.945; SSA, 2017).

Response bias: "Misrepresentation of abilities in any neuropsychological domain of ability (memory, sensorimotor, language, etc.) through performance, or self-report regarding performance capabilities" (Heilbronner et al., 2009, p. 1100).

Self-report measure: "Standardized instruments that rely on self-report with population-based normative data that allow the examiner to compare an individual's reported behaviors or symptoms with an appropriate comparison group" (IOM, 2015, p. 225).

Self-report of symptoms: An individual's "own description of [his or her] physical or mental impairment" (20 CFR § 404.1528).

Sensory processing: "The way the nervous system receives messages from the senses and turns them into appropriate motor and behavioral responses" (STAR Institute, 2018).

Substantial gainful activity (SGA): "Work that—(a) involves doing significant and productive physical or mental duties; and (b) is done (or intended) for pay or profit" (20 CFR 404.1510). The monthly SGA amount for non-blind individuals in 2019 is $1,220 after deducting impairment-related work expenses (SSA, 2019).

Symptom validity: "The accuracy of symptomatic complaint on self-report measures" (Larrabee, 2012, p. 2; see also Larrabee, 2014).

Symptom validity test: "Embedded or stand-alone measures used to assess whether an examinee is providing an accurate report of his or her actual symptom experience on non-cognitive psychological measures (e.g., emotional, behavioral, and personality measures)" (IOM, 2015, p. 226; adapted from Larrabee, 2014).

Task: A set of mental and physical activities in which an individual engages to accomplish a specific goal at or by a specific time.

Validity: "The degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (AERA et al., 2014, p. 11).

REFERENCES

AERA, APA, and NCME (American Educational Research Association, American Psychology Association, and National Council on Measurement in Education). 2014. *Standards for educational and psychological testing*. Washington, DC: AERA.

Brennan, R. L. 2006. Perspectives on the evolution and future of educational measurement. In *Educational measurement*, 4th ed., edited by R. L. Brennan. Westport, CT: Praeger Publishers. Pp. 1-16.

DOL (U.S. Department of Labor). 2017. *Occupational Requirements Survey (ORS): Collection manual*, 3rd edition. Washington, DC: Bureau of Labor Statistics. https://www.bls.gov/ors/about/information-for-survey-participants/pdf/occupational-requirements-survey-collection-manual-072017.pdf (April 10, 2019).

Furr, R. M., and V. R. Bacharach. 2014. *Psychometrics: An introduction*, 2nd ed. Thousand Oaks, CA: Sage Publications, Inc.

Harvey, P. D. 2012. Clinical applications of neuropsychological assessment. *Dialogues in Clinical Neuroscience* 14(1):91–99.

Heilbronner, R. L., J. J. Sweet, J. E. Morgan, G. J. Larrabee, S. R. Millis, and Conference participants. 2009. American Academy of Clinical Neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *Clinical Neuropsychologis*t 23(7):1093–1129.

Hubley, A. M., and B. D. Zumbo. 2013. Psychometric characteristics of assessment procedures: An overview. In *APA handbook of testing and assessment in psychology: Vol. 1 Test theory and testing and assessment in industrial and organizational psychology*, edited by K. F. Geisinger (editor-in-chief) and B. A. Bracken, J. F. Carlson, J. C. Hansen, N. R. Kuncel, S. P. Reise, and M. C. Rodriguez (associate editors). Washington, DC: American Psychological Association. Pp. 3–19.

IOM (Institute of Medicine). 2015. *Psychological testing in the service of disability determination*. Washington, DC: The National Academies Press.

Larrabee, G. J. 2012. Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society* 18(4):625–630.

Larrabee, G. J. 2014. *Performance and Symptom Validity*. Presentation to the Committee on Psychological Testing, Including Validity Testing, for Social Security Administration Disability Determinations, Washington, DC, June 25.

Merriam-Webster. 2019a. *Capability*. https://www.merriam-webster.com/dictionary/capability (accessed April 10, 2019).

Merriam-Webster. 2019b. *Medical definition of rehabilitation*. https://www.merriam-webster.com/dictionary/rehabilitation (accessed April 10, 2019).

Merriam-Webster. 2019c. *Performance*. https://www.merriam-webster.com/dictionary/performance (accessed April 10, 2019).

OIDAP (Occupational Information Development Advisory Panel). 2009. *Mental cognitive subcommittee: Content model and classification recommendations*. https://www.ssa.gov/oidap/Documents/AppendixC.pdf (accessed April 10, 2019).

O*NET. 2019a. *The O*NET Content Model*. https://www.onetcenter.org/content.html (accessed April 25, 2019).

O*NET. 2019b. *Work styles*. https://www.onetonline.org/find/descriptor/browse/Work_Styles (accessed April 10, 2019).

SSA (U.S. Social Security Administration). 2012. *DI 00115.015 Definitions of disability*. https://secure.ssa.gov/poms.nsf/lnx/0400115015 (accessed April 10, 2019).

SSA. 2017. *Adult functional assessment in Social Security disability claims*. Presentation to the Committee on Functional Assessment for Adults with Disabilities, Washington, DC, December 7.

SSA. 2019. *Substantial gainful activity*. https://www.ssa.gov/oact/cola/sga.html (accessed March 7, 2019).

SSA. n.d.-a. *Disability evaluation under Social Security; Part I—General information*. http://www.ssa.gov/disability/professionals/bluebook/general-info.htm (accessed April 10, 2019).

SSA. n.d.-b. *Disability evaluation under Social Security: 12.00 Mental disorders adult*. https://www.ssa.gov/disability/professionals/bluebook/12.00-MentalDisorders-Adult.htm (accessed April 10, 2019).

STAR Institute. 2018. *About SPD*. https://www.spdstar.org/basic/about-spd (accessed April 10, 2019).

Stein, R. E. K., S. L. Gortmaker, E. C. Perrin, J. M. Perrin, I. B. Pless, D. K. Walker, and M. Weitzman. 1987. Severity of illness: Concepts and measurements. *Lancet* 2(8574):1506–1509.

WHO (World Health Organization). 2001. *International Classification of Functioning, Disability and Health*. Geneva, Switzerland: WHO.

Appendix C

Literature Search Strategies

FUNCTIONAL ASSESSMENT FOR ADULTS WITH DISABILITIES:

SEARCH STRATEGY #1

Requested by: Study staff
Conducted by: Rebecca Morgan
Date: November 9, 2017

Search Parameters:

Date: 1980–Present
Age Group: Adults (18+)
Country: U.S. and International
Language: English

Document Types: Peer-reviewed articles, grey literature reports, conference proceedings, reviews, and National Academies reports (exclude editorials)

Databases and Websites:
Medline (Ovid)
PubMed
Web of Science
Scopus
nap.edu

Social Security Administration website
Occupational Information Development Advisory Panel website
U.S. Department of Labor website
Bureau of Labor Statistics website

Search Syntax:

Medline (Ovid):

Search No.	Search Syntax	Results
1	("functional capacity" or "functional assessment"). ti,ab.	17,951
2	"functional capacity evaluation".ti,ab.	186
3	"functional assessment tool".ti,ab.	64
4	Work Capacity Evaluation/	5,941
5	"self reported function".ti,ab.	226
6	"self reported functional limitation".ti,ab.	26
7	"Outcome Assessment (Health Care)"/	66,687
8	Disability Evaluation/	45,125
9	"disability evaluation".ti,ab.	454
10	"functional status".ti,ab.	20,418
11	"disability assessment".ti,ab.	1,252
12	"work capacity evaluation".ti,ab.	119
13	"functional outcome measure".ti,ab.	95
14	"functional outcome".ti,ab.	18,115
15	"outcome measures".ti,ab.	129,880
16	Health Status Indicators/	23,476
17	"Task Performance and Analysis"/	30,271
18	"functional evaluation".ti,ab.	3,524
19	"work instability scale".ti,ab.	46
20	"severity of illness scale".ti,ab.	180
21	"PROMIS Physical Function".ti,ab.	58
22	"Work Disability Functional Assessment Battery".ti,ab.	3
23	"WD-FAB".ti,ab.	2
24	"Pfeffer Questionnaire".ti,ab.	5
25	or/1–24	334,477
26	Disabled Persons/	39,129

Search No.	Search Syntax	Results
27	“work disability”.ti,ab.	1,650
28	“functional impairment”.ti,ab.	12,350
29	“functional limitation”.ti,ab.	1,614
30	“functional ability”.ti,ab.	3,976
31	“functional abilities”.ti,ab.	1,677
32	“physical impairment”.ti,ab.	1,226
33	“mental impairment”.ti,ab.	918
34	“Recovery of Function”/	45,564
35	“physical function”.ti,ab.	9,053
36	“mental function”.ti,ab.	966
37	“social function”.ti,ab.	2,101
38	“work related functioning”.ti,ab.	12
39	“impairment progression”.ti,ab.	23
40	Work/	20,296
41	Employment/	43,851
42	employment.ti,ab.	43,842
43	Occupations/	22,889
44	occupation.ti,ab.	23,373
45	“occupational requirement”.ti,ab.	13
46	“job requirement”.ti,ab.	17
47	“work requirement”.ti,ab.	43
48	“job zone 1”.ti,ab.	0
49	“occupational information network”.ti,ab.	37
50	“O Net”.ti,ab.	12
51	“occupational requirements survey”.ti,ab.	0
52	Return to Work/	1,562
53	“return to work”.ti,ab.	7,017
54	Rehabilitation, Vocational/	9,457
55	“vocational rehabilitation”.ti,ab.	1,977
56	“occupational data”.ti,ab.	241
57	“work activity”.ti,ab.	672
58	“work performance”.ti,ab.	1,847
59	Absenteeism/	8,731

Search No.	Search Syntax	Results
60	absenteeism.ti,ab.	4,516
61	Sick Leave/	5,263
62	"sick leave".ti,ab.	3,985
63	Presenteeism/	122
64	presenteeism.ti,ab.	678
65	or/40–64	158,907
66	or/26–39	116,023
67	25 and 65 and 66	2,141
68	Adult/	4,812,310
69	Middle Aged/	4,134,177
70	Aged/	2,909,737
71	Humans/	17,833,607
72	or/68–70	6,751,937
73	67 and 72 and 71	1,512
74	73	1,512
75	limit 74 to (English language and yr="1980–Current")	1,339
76	limit 75 to (comment or editorial or letter)	3
77	75 not 76	1,336
78	limit 77 to (meta analysis or "review" or systematic reviews)	68
79	77 not 78	1,268

Results:

Peer-reviewed articles: 1,268

Reviews: 68

Scopus:

TITLE-ABS-KEY("functional assessment" OR "functional capacity" OR {functional assessment tool} OR {self reported function} OR {self reported functional limitation} OR "outcomes assessment" OR "disability evaluation" OR "functional status" OR "disability assessment" OR {work capacity evaluation} OR {functional outcome measure} OR "functional outcome" OR "outcome measures" OR {health status indicators} OR "task performance" OR "functional evaluation" OR {work instability scale} OR {severity of illness scale} OR {PROMIS Physical Function} OR {work disability functional assessment battery} OR {WD-FAB} OR {pfeffer questionnaire}) AND TITLE-ABS-KEY(disability OR "work disability" OR "disabled persons" OR "functional impairment" OR "functional limitation" OR "functional abilities" OR "physical impairment" OR "mental impairment" OR {recovery of function} OR {change in function} OR "physical function" OR "mental function" OR "social function" OR {work related functioning} OR "impairment progression") AND TITLE-ABS-KEY(employment OR occupation OR "occupational requirement" OR job OR "job requirement" OR "work requirement" OR "job zone 1" OR "O NET" OR {occupational information system} OR {occupational requirements survey} OR {return to work} OR "vocational rehabilitation" OR "occupational data" OR "work activity" OR "work performance" OR absenteeism OR "sick leave" OR presenteeism) AND NOT INDEX(medline)

Limit: English

Date: 1980–Present

Results: 957

Web of Science:

TS=("functional assessment" OR "functional capacity" OR "functional assessment tool" OR "self reported function" OR "self reported functional limitation" OR "outcomes assessment" OR "disability evaluation" OR "functional status" OR "disability assessment" OR "work capacity evaluation" OR "functional outcome measure" OR "functional outcome" OR "outcome measures" OR "health status indicators" OR "task performance" OR "functional evaluation" OR "work instability scale" OR "severity of illness scale" OR "PROMIS Physical Function" OR "work disability functional assessment battery" OR "WD-FAB" OR "pfeffer questionnaire") AND TS=(disability OR "work disability" OR "disabled persons" OR "functional impairment" OR "functional limitation" OR "functional abilities" OR "physical impairment" OR "mental impairment" OR "recovery of function" OR "change in function" OR "physical function" OR "mental function" OR "social function" OR "work related functioning" OR "impairment progression") AND TS=(employment OR occupation OR "occupational requirement" OR job OR "job requirement" OR "work requirement" OR "job zone 1" OR "O NET" OR "occupational information system" OR "occupational requirements survey" OR "return to work" OR "vocational rehabilitation" OR "occupational data" OR "work activity" OR "work performance" OR absenteeism OR "sick leave" OR presenteeism)

Limit: English

Date: 1980–Present

Results: 1,333

PubMed:
Note: The PubMed search was modified to ensure relevant literature

("Work Capacity Evaluation"[Mesh] OR "Disability Evaluation"[Mesh] OR "functional assessment" OR "functional capacity" OR "work instability scale" OR "severity of illness scale" OR "PROMIS Physical Function" OR "Work Disability Functional Assessment Battery" OR "WD FAB" OR "Pfeffer Questionnaire") AND ("Disabled Persons"[Mesh] OR Disability OR "Recovery of Function"[Mesh]) AND ("Work"[Mesh] OR "Return to Work"[Mesh] OR "Employment"[Mesh] OR "Occupations"[Mesh] OR "Sick Leave"[Mesh] OR "Absenteeism"[Mesh] OR "Presenteeism"[Mesh] OR "job zone 1" OR "occupational information network" OR "occupational requirements survey") AND ("Adult"[Mesh] OR "Aged"[Mesh] OR "Middle Aged"[Mesh]) NOT ("Comment" [Publication Type] OR "Editorial" [Publication Type] OR "Letter" [Publication Type])

Limit: 1980–Present

Language: English

Population: Human

Results: 1,594

FUNCTIONAL ASSESSMENT FOR ADULTS WITH DISABILITIES:

SEARCH STRATEGY #2

Requested by: Study staff
Conducted by: Rebecca Morgan
Date: December 11, 2017

Search Parameters:

Date: 1998–Present
Country: U.S. and International
Language: English

Document Types: Peer-reviewed articles, reviews

Databases and Websites:
Medline (Ovid)
PubMed
Web of Science
Scopus

Search Syntax:

Medline (Ovid):

Search No.	Search Syntax	Results
1	("functional capacity" or "functional assessment").ti,ab.	18,850
2	"functional assessment tool".ti,ab.	67
3	"self reported function".ti,ab.	231
4	"self reported functional limitation".ti,ab.	27
5	Disability Evaluation/	46,575
6	"disability evaluation".ti,ab.	460
7	"functional status".ti,ab.	21,257
8	"disability assessment".ti,ab.	1,316
9	"work capacity evaluation".ti,ab.	120
10	Work Capacity Evaluation/	6,156
11	"functional evaluation".ti,ab.	3,667
12	"work instability scale".ti,ab.	53
13	"PROMIS Physical Function".ti,ab.	62

Search No.	Search Syntax	Results
14	"Work Disability Functional Assessment Battery".ti,ab.	4
15	"WD-FAB".ti,ab.	3
16	"Pfeffer Questionnaire".ti,ab.	5
17	or/1–16	92,643
18	disability.ti,ab.	117,901
19	Disabled Persons/	39,920
20	("work disability" or "work ability").ti,ab.	2,951
21	"functional impairment".ti,ab.	12,920
22	"functional limitation".ti,ab.	1,673
23	"functional ability".ti,ab.	4,149
24	"functional abilities".ti,ab.	1,740
25	"physical impairment".ti,ab.	1,283
26	"mental impairment".ti,ab.	945
27	"work related functioning".ti,ab.	13
28	"Activities of Daily Living"/	63,298
29	"activities of daily living".ti,ab.	20,106
30	or/18–29	219,381
31	Social Security/	7,540
32	"social security".ti,ab.	7,823
33	"International Classification of Functioning Disability and Health".ti,ab.	2,075
34	"disability benefits".ti,ab.	642
35	Employment/	44,835
36	employment.ti,ab.	45,226
37	Occupations/ or Work/	43,058
38	occupation.ti,ab.	23,993
39	"occupational requirement".ti,ab.	15
40	"job requirement".ti,ab.	18
41	"work requirement".ti,ab.	43
42	"job zone 1".ti,ab.	0
43	"occupational information network".ti,ab.	39
44	"O Net".ti,ab.	12
45	"occupational requirements survey".ti,ab.	0

Search No.	Search Syntax	Results
46	Return to Work/	1,739
47	"return to work".ti,ab.	7,491
48	"vocational rehabilitation".ti,ab.	2,054
49	"work activity".ti,ab.	704
50	"work performance".ti,ab.	1,896
51	Absenteeism/	8,945
52	absenteeism.ti,ab.	4,690
53	Sick Leave/	5,563
54	"sick leave".ti,ab.	4,192
55	Presenteeism/	134
56	presenteeism.ti,ab.	728
57	or/31–56	169,845
58	17 and 30 and 57	5,472
59	Humans/	18,287,738
60	human.ti,ab.	2,192,739
61	or/59–60	18,534,428
62	58 and 61	5,402
63	limit 62 to (English language and yr="1998–Current")	3,864
64	limit 63 to (comment or editorial or letter)	42
65	63 not 64	3,822
66	limit 65 to (meta analysis or "review" or systematic reviews)	547
67	65 not 66	3,275

Results:

Peer-reviewed articles: 3,275

Reviews: 547

Scopus:

TITLE-ABS-KEY(("functional assessment" OR "functional capacity" OR {functional assessment tool} OR {self reported function} OR {self reported functional limitation} OR "disability evaluation" OR "functional status" OR "disability assessment" OR {work capacity evaluation} OR "functional evaluation" OR {work instability scale} OR {PROMIS Physical Function} OR {work disability functional assessment battery} OR {WD-FAB} OR {pfeffer questionnaire}) AND (disability OR "work disability" OR "work ability" OR "disabled persons" OR "functional impairment" OR "functional limitation" OR "functional abilities" OR "physical impairment" OR "mental impairment" OR {work related functioning} OR {activities of daily living}) AND ("social security" OR {International Classification of Functioning Disability and Health} OR "disability benefits" OR employment OR occupation OR (work w/10 participation) OR "occupational requirement" OR "job requirement" OR "work requirement" OR "job zone 1" OR "O NET" OR {occupational information network} OR {occupational requirements survey} OR {return to work} OR "vocational rehabilitation" OR "occupational data" OR "work activity" OR "work performance" OR absenteeism OR "sick leave" OR presenteeism)) AND NOT INDEX(medline)

Limit: English

Date: 1998–Present

Results:
Article, Article in Press, Book, Book Chapter: 382
Reviews: 50

Web of Science:

TS=(("functional assessment" OR "functional capacity" OR "functional assessment tool" OR "self reported function" OR "self reported functional limitation" OR "disability evaluation" OR "functional status" OR "disability assessment" OR "work capacity evaluation" OR "functional evaluation" OR "work instability scale" OR "PROMIS Physical Function" OR "work disability functional assessment battery" OR "WD-FAB" OR "pfeffer questionnaire") AND (disability OR "work disability" OR "work ability" OR "disabled persons" OR "functional impairment" OR "functional limitation" OR "functional abilities" OR "physical impairment" OR "mental impairment" OR "work related functioning" OR "activities of daily living") AND ("social security" OR "International Classification of Functioning Disability and Health" OR "disability benefits" OR employment OR occupation OR (work NEAR/10 participation) OR "occupational requirement" OR "job requirement" OR "work requirement" OR "job zone 1" OR "O NET" OR "occupational information system" OR "occupational requirements survey" OR "return to work" OR "vocational rehabilitation" OR "occupational data" OR "work activity" OR "work performance" OR absenteeism OR "sick leave" OR presenteeism))

Limit: English

Date: 1998–Present

Results:
Article, Book: 747
Reviews: 73

PubMed:

(("functional assessment" OR "functional capacity" OR "functional assessment tool" OR "self reported function" OR "disability evaluation" OR "functional status" OR "disability assessment" OR "work capacity evaluation" OR "functional evaluation" OR "work instability scale" OR "PROMIS Physical Function" OR "Work Disability Functional Assessment Battery" OR "WD FAB" OR "Pfeffer Questionnaire" OR "Disability Evaluation"[Mesh] OR "Work Capacity Evaluation"[Mesh]) AND (disability OR "work disability" OR "work ability" OR "functional impairment" OR "functional limitation" OR "functional abilities" OR "functional ability" OR "physical impairment" OR "mental impairment" OR "work related functioning" OR "Disabled Persons"[Mesh] OR "Activities of Daily Living"[Mesh]) AND ("International Classification of Functioning Disability and Health" OR "disability benefits" OR employment OR occupation OR "work participation" OR "occupational requirement" OR "job requirement" OR "work requirement" OR "job zone 1" OR "O NET" OR "occupational information network" OR "occupational requirements survey" OR "return to work" OR "vocational rehabilitation" OR "occupational data" OR "work activity" OR "work performance" OR absenteeism OR presenteeism OR "sick leave" OR "Social Security"[Mesh] OR "Employment"[Mesh] OR "Work"[Mesh] OR "Occupations"[Mesh] OR "Return to Work"[Mesh] OR "Absenteeism"[Mesh] OR "Presenteeism"[Mesh] OR "Sick Leave"[Mesh])) NOT ("Comment" [Publication Type] OR "Editorial" [Publication Type] OR "Letter" [Publication Type])

Limit: 1998–Present

Language: English

Population: Human

Results:
Journal Articles: 3,736
Reviews: 404

FUNCTIONAL ASSESSMENT FOR ADULTS WITH DISABILITIES:

SEARCH STRATEGY #3

Requested by: Study staff
Conducted by: Rebecca Morgan
Date: May 8, 2018

Search Parameters:

Date: 1980–Present
Country: U.S. and International
Language: English

Document Types: Peer-reviewed articles, reviews

Databases and Websites:
Medline (Ovid)
PubMed
Web of Science
Scopus

Search Syntax:

Medline (Ovid):

Search No.	Syntax	Results
1	"Disabilities of the Arm Shoulder and Hand Questionnaire".ti,ab.	338
2	"QuickDASH".ti,ab.	359
3	"dash questionnaire".ti,ab.	475
4	"roland disability questionnaire".ti,ab.	64
5	"Oswestry Disability Questionnaire".ti,ab.	170
6	"Quebec Back Pain Disability Questionnaire".ti,ab.	4
7	("Functional capacity evaluation*" or "functional capacity assessment*").ti,ab.	253
8	(blankenship or "hanoun medical").ti,ab.	24

Search No.	Syntax	Results
9	("key method" or "west-epic").ti,ab.	195
10	ergos.ti,ab.	13
11	ARCON.ti,ab.	76
12	"assess ability".ti,ab.	62
13	or/8–12	369
14	7 and 13	9
15	"ergos work simulator".ti,ab.	7
16	"ergo kit".ti,ab.	9
17	"isernhagen work system*".ti,ab.	19
18	"physical work performance evaluation".ti,ab.	9
19	ergoscience.ti,ab.	0
20	"Neck Disability Index".ti,ab.	1,104
21	"Lower Extremity Functional Scale".ti,ab.	174
22	"Patient rated wrist hand evaluation".ti,ab.	16
23	PRWHE.ti,ab.	17
24	"Patient-Rated Elbow Evaluation".ti,ab.	32
25	"functional gait assessment".ti,ab.	47
26	"Michigan Hand Outcome Questionnaire".ti,ab.	29
27	"Mayo Elbow Performance Score".ti,ab.	454
28	"Oxford Elbow Score".ti,ab.	55
29	"Constant Shoulder Score".ti,ab.	87
30	"Oxford Shoulder Score".ti,ab.	166
31	"ASES Shoulder Score".ti,ab.	13
32	"Oxford instability Shoulder Score".ti,ab.	5
33	"Rowe Score for Instability".ti,ab.	4
34	"Harris Hip Score".ti,ab.	2,977
35	"Hip Disability and Osteoarthritis Outcome Score".ti,ab.	104
36	"Knee Society Score".ti,ab.	968
37	"Knee Injury and Osteoarthritis Score".ti,ab.	8
38	"International Knee Documentation Committee Evaluation Form".ti,ab.	9
39	"Back Pain Index".ti,ab.	13
40	"Foot and Ankle Questionnaire".ti,ab.	19

Search No.	Syntax	Results
41	"Foot and Ankle Disability Index".ti,ab.	49
42	"Manchester Foot Pain and Disability Index".ti,ab.	30
43	MFPDI.ti,ab.	14
44	"foot function index".ti,ab.	223
45	"foot health assessment instrument".ti,ab.	3
46	"leeds foot impact scale".ti,ab.	11
47	"nordic musculoskeletal questionnaire".ti,ab.	135
48	or/1–6	1,350
49	or/15–47	6,679
50	14 or 48 or 49	7,955
51	"Reproducibility of Results"/	354,560
52	Psychometrics/	66,879
53	reliability.ti,ab.	117,591
54	validity.ti,ab.	125,181
55	or/51–54	518,286
56	50 and 55	698
57	56	698
58	limit 57 to (English language and yr="1980–Current")	667

Scopus:
TITLE-ABS-KEY(({Disabilities of the Arm Shoulder and Hand Questionnaire} OR {Quick DASH} OR {Roland Disability Questionnaire} OR {Oswestry Disability Questionnaire} OR {Quebec Back Pain Disability Questionnaire} OR {Neck Disability Index} OR {Lower Extremity Functional Scale} OR {Patient rated wrist hand evaluation} OR {Patient-Rated Elbow Evaluation} OR {Functional Gait Assessment} OR {Michigan Hand Outcome Questionnaire} OR {Mayo Elbow Performance Score} OR {Oxford Elbow Score} OR {Constant Shoulder Score} OR {Oxford Shoulder Score} OR {ASES Shoulder Score} OR {Oxford instability Shoulder Score} OR {The Rowe Score for Instability} OR {Harris Hip Score} OR {Hip Disability and Osteoarthritis Outcome Score} OR {Knee Society Score} OR {Knee Injury and Osteoarthritis Score} OR {International Knee Documentation Committee Evaluation Form} OR {Back Pain Index} OR {Foot and Ankle Questionnaire} OR {Foot and Ankle Disability Index} OR {Ergos Work Simulator} OR {Ergo-Kit} OR {Isernhagen Work System} OR {Hanoun Medical} OR {Physical Work Performance Evaluation} OR {WEST-EPIC}

OR {Manchester Foot Pain and Disability Index} OR {Foot and Ankle Ability Measure} OR {Foot Function Index} OR {Foot Health Assessment Instrument} OR {Leeds Foot Impact Scale} OR {Nordic Musculoskeletal Questionnaire}) AND ({Reproducibility of Results} OR Psychometrics OR validity OR reliability)) AND PUBYEAR AFT 1979

Language: English
Results: 846

TITLE-ABS-KEY((({functional capacity evaluations} OR {functional capacity assessments}) AND (Blankenship OR KEY OR ergos OR ARCON OR "assess ability") AND ({Reproducibility of Results} OR Psychometrics OR validity OR reliability)) AND PUBYEAR AFT 1979

Language: English
Results: 12

Web of Science:
TS=(("Disabilities of the Arm Shoulder and Hand Questionnaire" OR "Quick DASH" OR "Roland Disability Questionnaire" OR "Oswestry Disability Questionnaire" OR "Quebec Back Pain Disability Questionnaire" OR "Neck Disability Index" OR "Lower Extremity Functional Scale" OR "Patient rated wrist hand evaluation" OR "Patient-Rated Elbow Evaluation" OR "Functional Gait Assessment" OR "Michigan Hand Outcome Questionnaire" OR "Mayo Elbow Performance Score" OR "Oxford Elbow Score" OR "Constant Shoulder Score" OR "Oxford Shoulder Score" OR "ASES Shoulder Score" OR "Oxford instability Shoulder Score" OR "The Rowe Score for Instability" OR "Harris Hip Score" OR "Hip Disability and Osteoarthritis Outcome Score" OR "Knee Society Score" OR "Knee Injury and Osteoarthritis Score" OR "International Knee Documentation Committee Evaluation Form" OR "Back Pain Index" OR "Foot and Ankle Questionnaire" OR "Foot and Ankle Disability Index" OR "Ergos Work Simulator" OR "Ergo-Kit" OR "Isernhagen Work System" OR "Hanoun Medical" OR "Physical Work Performance Evaluation" OR "WEST-EPIC" OR "Manchester Foot Pain and Disability Index" OR "Foot and Ankle Ability Measure" OR "Foot Function Index" OR "Foot Health Assessment Instrument" OR "Leeds Foot Impact Scale" OR "Nordic Musculoskeletal Questionnaire") AND ("Reproducibility of Results" OR Psychometrics OR validity OR reliability))

Date: 1980–Present
Language: English
Results: 1,166

TS=((“functional capacity evaluation” OR “functional capacity assessment”) AND (Blankenship OR KEY OR ergos OR ARCON OR “assess ability”) AND (“Reproducibility of Results” OR Psychometrics OR validity OR reliability))

Date: 1980–Present
Language: English
Results: 12

PubMed:
((“Disabilities of the Arm Shoulder and Hand Questionnaire” OR “Quick DASH” OR “Roland Disability Questionnaire” OR “Oswestry Disability Questionnaire” OR “Quebec Back Pain Disability Questionnaire” OR “Neck Disability Index” OR “Lower Extremity Functional Scale” OR “Patient rated wrist hand evaluation” OR “Patient-Rated Elbow Evaluation” OR “Functional Gait Assessment” OR “Michigan Hand Outcome Questionnaire” OR “Mayo Elbow Performance Score” OR “Oxford Elbow Score” OR “Constant Shoulder Score” OR “Oxford Shoulder Score” OR “ASES Shoulder Score” OR “Oxford instability Shoulder Score” OR “The Rowe Score for Instability” OR “Harris Hip Score” OR “Hip Disability and Osteoarthritis Outcome Score” OR “Knee Society Score” OR “Knee Injury and Osteoarthritis Score” OR “International Knee Documentation Committee Evaluation Form” OR “Back Pain Index” OR “Foot and Ankle Questionnaire” OR “Foot and Ankle Disability Index” OR “Ergos Work Simulator” OR “Ergo-Kit” OR “Isernhagen Work System” OR “Hanoun Medical” OR “Physical Work Performance Evaluation” OR “WEST-EPIC” OR “Manchester Foot Pain and Disability Index” OR “Foot and Ankle Ability Measure” OR “Foot Function Index” OR “Foot Health Assessment Instrument” OR “Leeds Foot Impact Scale” OR “Nordic Musculoskeletal Questionnaire”) AND (“Reproducibility of Results” OR Psychometrics OR validity OR reliability))

Date: 1980–Present
Language: English
Results: 1,775

((“functional capacity evaluation” OR “functional capacity assessment”) AND (Blankenship OR KEY OR ergos OR ARCON OR “assess ability”) AND (“Reproducibility of Results” OR Psychometrics OR validity OR reliability))

Date: 1980–Present
Language: English
Results: 10

Appendix D

Biographical Sketches of Committee Members

Paul A. Volberding, M.D. (*Chair*), is a professor of medicine at the University of California, San Francisco (UCSF); the director of the AIDS Research Institute; and the co-director of the University of California, San Francisco-Gladstone Institute of Virology and Immunology Center for AIDS Research. He received his undergraduate and medical degrees at the University of Chicago and the University of Minnesota, respectively. He completed his fellowship in medical oncology at UCSF. For 20 years, Dr. Volberding's professional activities centered at San Francisco General Hospital, where he established a model program of AIDS care, research, and professional education. His research career began with investigations of HIV-related malignancies but shifted to clinical trials of antiretroviral drugs. He helped lead early studies in asymptomatic infection that led to the concept of HIV disease as the target of treatment. He more recently served as the chief of medicine at the San Francisco Veterans Affairs Medical Center. Dr. Volberding has written many research and review articles. He is the co-editor-in-chief of the *Journal of Acquired Immune Deficiency Syndrome*. He has written several textbooks including *Sande's HIV/AIDS Medicine* and the companion text, *Global Care*, specifically for use in resource-limited settings. He is the founder and chair of the board of the International Antiviral Society-USA. He was the president of the HIV Medical Association of the Infectious Diseases Society of America. He was elected a member of the National Academy of Medicine in 1999. Dr. Volberding currently serves on the National Academies of Sciences, Engineering, and Medicine Standing Committee of Medical and Vocational Experts for the U.S. Social Security

Administration's Disability Programs and previously chaired the Committee on Social Security HIV Disability Criteria. In 2014, he was elected as a Master of the American College of Physicians.

María P. Aranda, Ph.D., M.S.W., M.P.A., LCSW, is an associate professor at the University of Southern California (USC) Suzanne Dworak-Peck School of Social Work, with a joint appointment at the USC Leonard Davis School of Gerontology. She is the executive director of the USC Edward R. Roybal Institute on Aging, and director of the USC Alzheimer Disease Research Center Outreach, Recruitment and Engagement Core. Her research addresses the psychosocial care of adult and late-life psychiatric and neurocognitive disorders and comorbid medical conditions. She specializes in the role of racial and ethnic diversity in health care and community-based services, and sociocultural adaptations to evidence-based interventions for people with disabilities and their care partners. Dr. Aranda has served on several National Academies of Sciences, Engineering, and Medicine committees, including the Committee to Evaluate the Social Security Administration's Capability Determination Process for Adult Beneficiaries. She received an M.S.W. and a Ph.D. from USC's Suzanne Dworak-Peck School of Social Work, and an M.P.A. from USC's School of Public Policy and Development.

Jack T. Dennerlein, Ph.D., is a professor in the Department of Physical Therapy, Movement and Rehabilitation Sciences at Northeastern University's Bouvé College of Health Sciences. In addition, he is an adjunct professor of Ergonomics and Safety at the Harvard T.H. Chan School of Public Health as well as the associate director of the Harvard T.H. Chan School's Center for Work, Health, and Well-being. Dr. Dennerlein's research examines how design of work impacts worker safety, health, and well-being with a focus on prevention of musculoskeletal disorders, injury, and work disability. He is a fellow of the Human Factors and Ergonomics Society. He received a B.S. in mechanical engineering from the State University of New York at Buffalo, an M.S. in mechanical engineering from the Massachusetts Institute of Technology, and a Ph.D. from the University of California, Berkeley.

Lisa Dixon, M.D., M.P.H., is the Edna L. Edison Professor of Psychiatry at the Columbia University Vagelos College of Physicians and Surgeons where she directs the Division of Behavioral Health Services and Policy Research and the Center for Practice Innovations (CPI) at the New York State Psychiatric Institute. Dr. Dixon is an internationally recognized health services researcher with more than 25 years of continuous research funding from the National Institute of Mental Health, the U.S. Department of Veterans Affairs, and foundations. As CPI director, she oversees activities for the New York State Office of Mental Health in implementing evidence-based practices in behavioral health programs throughout the state. She leads the innovative program, OnTrackNY, a statewide initiative designed to improve outcomes and reduce disability for the population of individuals experiencing their first episode of psychosis. Dr. Dixon's grants have focused on improving the quality of care for individuals with serious mental disorders with a particular emphasis on services that include families, reducing the negative impact of co-occurring addictions and medical problems, and improving treatment engagement and adherence. Dr. Dixon's work has joined individuals engaged in self-help, outpatient psychiatric care, as well as clinicians and policy makers in collaborative research endeavors. Dr. Dixon assumed the role of editor-in-chief of the journal *Psychiatric Services* in January 2017. She has published more than 250 articles in peer-reviewed journals and has received numerous awards including the 2009 American Psychiatric Association Health Services Senior Scholar Award and the Wayne Fenton Award for Exceptional Clinical Care. In 2014, the National Alliance on Mental Illness (NAMI) Metro NYC recognized her with the Adele Anshien Volunteer of the Year Award, and NAMI national recognized her with its annual Scientific Research Award. In 2016, the Mental Health Section of the American Public Health Association recognized her work with the Carl A. Taube Award. She received a B.A. in economics from Harvard College, an M.D. from Cornell University Medical College, and an M.P.H. from the Johns Hopkins University School of Hygiene and Public Health.

Judith Green-McKenzie, M.D., M.P.H., FACP, FACOEM, FACPM, is a professor, the division chief, and the residency program director in the Department Emergency Medicine, Division of Occupational and Environmental Medicine, at the University of Pennsylvania Perelman School of Medicine where she is active in clinical practice, research, education, and administration. She is also a senior fellow of the Leonard Davis Institute for Health Economics and the Graduate Program in Public Health. Dr. Green-McKenzie received her A.B. from Princeton University where she was awarded the Frederick Douglass Prize for leadership and scholarship, her M.D. from Yale University School of Medicine where she was a Commonwealth Fellow, and her M.P.H. from Johns Hopkins University School of Hygiene and Public Health where she also completed her Occupational Medicine Fellowship and the Epidemiology Research Track. She completed her Internal Medicine training at York University/Bellevue Hospital. Dr. Green-McKenzie is a diplomate of the American Board of Preventive Medicine, Occupational Medicine, and became a diplomate of the American Board of Internal Medicine in 1993. She was honored by American College of Occupational and Environmental Medicine (ACOEM), when it bestowed on her its 2015 International Kehoe Lifetime

Award for Excellence in Education and/or Research in Occupational and Environmental Medicine, recognized in particular for her leadership of the innovative, nationally recognized Train-in-Place Occupational and Environmental Medicine Residency, the first and only such program in the nation. She serves as chair of the National Academies of Sciences, Engineering, and Medicine Committee on Disabling Medical Conditions that Might Improve with Treatment and a member of the Standing Committee of Medical and Vocational Experts for the Social Security Administration's Disability Programs. She is a former member of the Committee on Health Care Utilization and Adults with Disabilities and the Committee on VA Examinations for Traumatic Brain Injury. Dr. Green-McKenzie is a member of the *Journal of Occupational and Environmental Medicine* Editorial Board and the Accreditation Council for Graduate Medical Education Residency Review Committee for Preventive Medicine; she also serves on the National Institute for Occupational Safety and Health (NIOSH). She served on the American Board of Preventive Medicine Examination Committee and as a permanent NIOSH study section member. She is a fellow of ACOEM, American College of Preventive Medicine, and the American College of Physicians. Dr. Green-McKenzie's clinical work centers on disability management, injury care, wellness and prevention, and environmental exposures. Author of 100 scientific publications, and principal investigator on two training grants, her research focuses on occupational and environmental medicine outcomes, especially in the areas of bloodborne pathogen exposures, work-related disability, graduate medical education, work as a social determinant of health, and employee wellness. She is listed as one of America's top physicians.

Allen W. Heinemann, Ph.D., is the director of the Center for Rehabilitation Outcomes Research at the Shirley Ryan AbilityLab and a professor of physical medicine and rehabilitation at Northwestern University's Feinberg School of Medicine. His research interests focus on health services research, psychosocial aspects of rehabilitation including substance abuse, and measurement issues in rehabilitation. He is the author of more than 300 articles in peer-reviewed publications and is the editor of *Substance Abuse and Physical Disability* published by Haworth Press. Dr. Heinemann is a diplomate in Rehabilitation Psychology (ABPP), and a fellow of the American Congress of Rehabilitation Medicine (ACRM) and the American Psychological Association (APA Division 22). During 2004–2005, he served as president of ACRM and the Rehabilitation Psychology division of the American Psychological Association. He serves as co-editor-in-chief for the *Archives of Physical Medicine and Rehabilitation*, and on the editorial boards of several journals including the *Journal of Head Trauma Rehabilitation and Rehabilitation Psychology*. He is the recipient of the APA Division 22 Roger Barker Distinguished Career Award. He serves on the National Academies of Sciences, Engineering, and Medicine Standing Committee of Medical and Vocational Experts for the Social Security Administration's Disability Programs and previously served on the Committee on Improving the Disability Decision Process: SSA's Listing of Impairments and Agency Access to Medical Expertise. He received a Ph.D. in psychology from the University of Kansas.

Andrew J. Houtenville, Ph.D., is an associate professor of economics and the research director at the Institute on Disability at the University of New Hampshire. His research focuses on the design of survey questions to identify people with disabilities; analysis of time trends and geographic dispersion in disability and the employment of people with disabilities; and identification of economic, social, programmatic, and workplace barriers and facilitators to the participation of people with disabilities in the labor market. He is currently the principal investigator of the Rehabilitation Research and Training Center (RRTC) on Disability Statistics and Demographics and the RRTC on Employment Policy and Measurement, both funded by the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR). He received an M.A. and a Ph.D. in economics from the University of New Hampshire.

Kurt L. Johnson, Ph.D., CRC, is a professor in the Department of Rehabilitation Medicine and the head of the Division of Rehabilitation Counseling and director of the University of Washington Center for Technology and Disability Studies. His research interests are focused on maximizing participation for people with disabilities in community and employment. He focuses on implementation of civil rights, uses of technology and accommodations, and how to measure outcomes. He received an M.Ed. in rehabilitation and mental health counseling from the University of Washington and a Ph.D. in rehabilitation psychology from the University of Wisconsin–Madison.

Barbara L. Kornblau, J.D., OTR/L, FAOTA, DASPE, CCM, CDMS, CPE, is the executive director of the Coalition for Disability Health Equity and on the faculty in the Division of Occupational Therapy in the School of Allied Health at Florida A&M University. She also serves as a consultant to the American Association on Health and Disability and the United Spinal Association. Ms. Kornblau is past president of the American Occupational Therapy Association, a former Robert Wood Johnson Health Policy Fellow in the Offices of Senators Harkin and Rockefeller, an attorney, a Certified Case Manager, a Certified Disability Management Specialist, a Certified Pain Educator, and a person with a disability. She is recognized as an expert in disability policy, return-to-work issues, assistive technology, and reasonable accommodations under the Americans with Disabilities Act and the Rehabilitation Act. She received a J.D. from the University of Miami and her occupational therapy degree from the University of Wisconsin–Madison.

Philip Jordan Marion, M.D., M.S., M.P.H., is a board-certified physical medicine and a rehabilitation/pain management specialist. He is a clinical professor of Medicine at the George Washington University School of Medicine & Health Sciences. Dr. Marion established the Rehabilitation Medicine Unit and is an attending physician at the George Washington University Hospital. Dr. Marion is also currently the medical director for the Polytrauma Amputation Network Site at the Washington, DC, VA Medical Center. His clinical interests include physical functional assessment, disability evaluation, and pain management. Dr. Marion completed his medical degree at the New York University School of Medicine. His clinical training was completed at Bellevue General Hospital and the Rusk Institute of Rehabilitation Medicine. While completing his residency training, Dr. Marion simultaneously obtained a master's degree in finance and health policy at the Wagner School of Public Service at New York University. Dr. Marion began his clinical practice at the National Rehabilitation Hospital and during that time completed the Master of Public Health degree program at George Washington University. Also during this time, he established the Howard University Medical Student Program at the National Rehabilitation Hospital. Dr. Marion was selected as a Robert Wood Johnson Foundation Health Policy Fellow and worked on health care reform on Capitol Hill.

Susan McGurk, Ph.D., is a professor of occupational therapy, with a secondary appointment in the Department of Psychological and Brain Sciences, and is a member of the Center of Psychiatric Rehabilitation at Boston University. Dr. McGurk is a neuropsychologist, with expertise in serious mental illnesses, and vocational rehabilitation, and community-based implementation of cognitive remediation programming. She directs a multifaceted research program addressing methods and mechanisms in cognitive remediation, cognitive self-management strategies, and the role of cognitive impairments in difficulties with employment, academic pursuits, and independent living in persons with serious psychiatric illnesses, and in other conditions affecting cognition and community functioning. Current research projects address the use of physical exercise to enhance cognitive remediation-related neuroplastic processes; tablet-based home practice of computerized cognitive exercises in people with schizophrenia seeking work; a multisite dismantling study of the specific components of a cognitive enhancement program developed by McGurk and Colleagues, The

Thinking Skills for Work Program (TSW), that are essential to helping people with psychiatric illness achieve their employment goals; and the development of a scaled up training model for practitioners of TSW. Dr. McGurk received her Ph.D. from the University of California, Los Angeles.

Juan I. Sanchez, Ph.D., is professor and the Knight-Ridder Byron Harless Eminent Chair in the Department of Global Leadership and Management, Florida International University. His areas of expertise are competency modeling and job analysis, performance management, human resource management, and international human resources management. Dr. Sanchez has extensive experience in the development and validation of personnel selection systems, job and task analysis, the design of commercial tests and test batteries, the development of criterion-related validity studies, and the design of training evaluation systems. He served as a member of the Occupational Information Development Advisory Panel (OIDAP), a federal advisory committee to the U.S. Social Security Administration (SSA), which provided independent advice and recommendations to SSA on the creation of an occupational information system (OIS). Dr. Sanchez has authored more than 100 articles and 20 book chapters on topics including the consensus of competency ratings, comparison of job analysis methodologies, and the evaluation of work analysis. His work has been cited approximately 9,000 times according to Google Scholar. He is a Fellow of the American Psychological Association and The Society for Industrial and Organizational Psychology. His editorial positions include serving as associate editor of the *Journal of Occupational and Organizational Psychology*; consulting editor of the *Journal of Applied Psychology*; and editorial board member of *Personnel Psychology*, the *International Journal of Selection and Assessment, Group and Organization Management*, and *The Journal of International Business Studies*. He has served on five National Academies of Sciences, Engineering, and Medicine committees, including the Panel to Review the Occupational Information Network (O*NET) and the Workshop on Assessment of 21st Century Skills. Dr. Sanchez received a Ph.D. in industrial/organizational psychology with a minor in management from the University of South Florida.

Paul Shattuck, Ph.D., is an associate professor at Drexel University's A.J. Drexel Autism Institute and the leader of the Institute's Research Program Area on Life Course Outcomes. He has a secondary faculty appointment at Drexel's Dornsife School of Public Health. Most of his current research is aimed at improving services and related outcomes among youth with autism as they leave high school and transition to young adulthood. Dr. Shattuck's work has been funded by the National Institute of Mental Health, the Health Resources and Services Administration, the National

Science Foundation, the Institute for Education Sciences, Autism Speaks, and the Organization for Autism Research. His research publications have appeared in high-impact scientific journals, including *Pediatrics*, *Psychiatric Services*, the *Archives of Pediatrics and Adolescent Medicine*, the *American Journal of Public Health*, and the *Journal of the American Academy of Child and Adolescent Psychiatry*. Prior to joining the A.J. Drexel Autism Institute, Dr. Shattuck served as a faculty member at the George Warren Brown School of Social Work at Washington University in St. Louis. Dr. Shattuck's professional background includes nonprofit fundraising and program development. His education includes a Ph.D. in social welfare from the University of Wisconsin–Madison and postdoctoral training in epidemiology.

Lynne Warner Stevenson, M.D., is a professor of Medicine and the Lisa M. Jacobson Chair in Cardiovascular Medicine at Vanderbilt University Medical School and a senior physician and director of Cardiomyopathy and of Advanced Heart Failure Training at the Vanderbilt Heart and Vascular Center, after 24 years as director of Heart Failure at Brigham and Women's Hospital in Boston. Her research has helped to elucidate principles for therapy of patients with heart failure. Initially focusing on the relief of congestion in this population, her current research includes the impact of outpatient therapy guided by ambulatory cardiac pressures, progression and regression of right ventricular dysfunction, optimal distribution of the limited hearts for cardiac transplantation, use of patient-reported functional outcomes to alter therapy, triage for advanced therapies, and palliative care for end-stage heart disease. Dr. Stevenson has been on the writing committees for more than 30 national guidelines in heart failure, cardiac arrhythmia devices, cardiac transplantation, and patient decision making. She is a founding member of the Interagency Registry of Mechanically Assisted Circulatory Support and a member of the Medicare Evidence Development & Coverage Advisory Committee. She has served on the U.S. Food and Drug Administration cardio-renal advisory panel and as an advisor for the Joint Commission on Accreditation of Healthcare Organizations, and previously on the committee of the National Academies of Sciences, Engineering, and Medicine to address Social Security Cardiovascular Disability Criteria. She received her M.D. from the Stanford University School of Medicine and is certified in internal medicine, with subcertifications in cardiovascular disease and advanced heart failure/transplantation.

Robert B. Wallace, M.D., is the Irene Ensminger Stecher Professor of Epidemiology and Internal Medicine at the University of Iowa's College of Public Health. Dr. Wallace's research interests are in clinical and population epidemiology and focus on the causes and prevention of disabling conditions among older people. He has had substantial experience in the conduct of both observational cohort studies of older people and clinical trials, including preventive interventions related to fracture, cancer, coronary disease, and women's health. He received a B.S. in medicine from Northwestern University, an M.D. from Northwestern University Medical School, and an M.Sc. in epidemiology from the State University of New York at Buffalo. He is board certified in preventive medicine and a Fellow of the American College of Preventive Medicine. Dr. Wallace is an elected member of the National Academy of Medicine.