

Mental health assessment in rehabilitation research

John R. McQuaid, PhD;^{1–2*} Brian P. Marx, PhD;^{3–4} Marc I. Rosen, MD;^{5–6} Lynn F. Bufka, PhD;⁷ Wendy Tenhula, PhD;⁸ Helene Cook, MA;⁹ Terence M. Keane, PhD^{3–4}

¹San Francisco Department of Veterans Affairs (VA) Medical Center, San Francisco, CA; ²University of California, San Francisco, San Francisco, CA; ³National Center for PTSD, VA Boston Healthcare System, Boston, MA; ⁴Boston University School of Medicine, Boston, MA; ⁵VA Connecticut Healthcare System, West Haven, CT; ⁶Yale School of Medicine, New Haven, CT; ⁷American Psychological Association, Washington, DC; ⁸VA Office of Mental Health Services, Washington, DC; ⁹Texas A&M University, College Station, TX

Abstract—Assessment in mental health research has evolved from focusing on symptoms and diagnosis to addressing a broad range of change, including psychosocial functioning. This is consistent with developments in the areas of psychosocial rehabilitation and the increase in recovery-oriented intervention models for mental disorders. We reviewed the status of assessment in mental health research, providing an overview of symptom and diagnostic assessment that is the cornerstone of most mental health research assessment. We then focused on measurement that can be applied across diagnostic groups and on functioning as a key mental health outcome. We reviewed the International Classification of Functioning, Disability, and Health and its implications for improvements in assessment. We provided an example of a new assessment, the Inventory of Psychosocial Functioning, which highlights key issues in the measurement of functioning. We then addressed improving research assessment, including issues of assessment in diverse populations and the need to capitalize on new data sources and new assessment technologies to advance assessment in mental health research. Finally, we reviewed and discussed areas for research and quality improvement, drawing on examples from the Department of Veterans Affairs to illustrate potential opportunities.

Key words: assessment; computer-aided assessment; Department of Veterans Affairs; diagnosis; functioning; International Classification of Functioning, Disability, and Health; Inventory of Psychosocial Functioning; mental health; recovery; rehabilitation; research.

Abbreviations: ADIS = Anxiety Disorder Interview Schedule; AUDIT = Alcohol Use Disorders Identification Test; BASIS-R = Revised Behavior and Symptom Identification Scale; BDI = Beck Depression Inventory; CAPS = Clinician-Administered PTSD Scale; CIDI = Composite International Diagnostic Interview; DAST = Drug Abuse Screening Test; DSM = Diagnostic and Statistical Manual of Mental Disorders; EMR = electronic medical record; GAF = Global Assessment of Functioning; ICD = International Classification of Diseases; ICF = International Classification of Functioning, Disability, and Health; IOM = Institute of Medicine; IPF = Inventory of Psychosocial Functioning; LIFE = Longitudinal Interval Follow-up Evaluation; MIRECC = Mental Illness Research, Education, and Clinical Center; PROMIS = Patient-Reported Outcomes Measurement Information System; PTSD = posttraumatic stress disorder; QOLI = Quality of Life Inventory; QUERI = Qualified Enhancement Research Initiative; RR&D = Rehabilitation Research and Development; SCID = Structured Clinical Interview for DSM-IV; SDS = Sheehan Disability Scale; SF-36 = 36-Item Short Form Health Survey; SF-36V = SF-36 for Veterans; SOFAS = Social and Occupational Functioning Assessment Scale; VA = Department of Veterans Affairs; WHO = World Health Organization; WHODAS-II = WHO Disability Assessment Scale-II.

*Address all correspondence to John R. McQuaid, PhD; San Francisco VA Medical Center–Mental Health (116-B) 4150 Clement St, San Francisco, CA 94121; 415-221-4810, ext 4106; fax: 415-750-6921. Email: john.mcquaid@va.gov
<http://dx.doi.org/10.1682/JRRD.2010.08.0143>

INTRODUCTION

Assessment in mental health research has progressed from imprecise, subjective conceptualizations of disorder and symptoms to objective, standardized definitions and measures. Improved assessment has derived from multiple disciplines and perspectives: better conceptual understanding of psychopathology from biopsychosocial frameworks, advances in measurement theory and assessment, and growth in the awareness of a broader range of variables that are relevant to understanding change in mental health outcomes. Since the advent of the Diagnostic and Statistical Manual of Mental Disorders (DSM), 3rd Edition [1], mental health research has focused on the presence of observable symptoms as targets for identifying patients within diagnostic categories. The atheoretical nature of the DSM and its subsequent iterations provided strong impetus for research in mental health and behavioral health more broadly. Growing research on the biological, psychological, and social aspects of psychiatric conditions over the past three decades is evidence of the success of this approach to diagnosis, assessment, and treatment [2–3].

The observable, measurable approach espoused in the DSM yielded a proliferation of disorder-specific measurement strategies. However, very few measures were developed that could effectively assess outcomes across studies of different disorders. Precise measurement of diagnostic criteria led to an exclusive focus on those disorders examined in a particular study, and cross-study and cross-disorder comparisons could rarely be achieved. These limitations became increasingly apparent as the rates of comorbidity across anxiety, mood, personality, and psychotic disorders were documented by a range of epidemiological surveys conducted in the 1980s and 1990s [4–6].

The purpose of this article is to highlight future directions in the assessment of mental health disorders in the context of rehabilitation research and when the assessment aims to understand the broad effect of an intervention. Major advances in our understanding of the complex nature of mental health conditions, coupled with portable and inexpensive technologies that may be used for assessment, encourage the field to think broadly of what might be possible. Powerful computers capable of rapid information processing in the service of data management and statistical analyses yield a new array of methods, means, and models for research in this field [7]. Both educational and mental health research now use approaches based on item response theory to minimize

the length of measures. All of these changes make the next decade a promising one in the development of new approaches to assessment in our field.

ISSUES RELATED TO DIAGNOSTIC AND SYMPTOM ASSESSMENT

The most common psychiatric diagnoses in the population served by the Department of Veterans Affairs (VA) healthcare system are depression, substance abuse or dependence (including alcohol), posttraumatic stress disorder (PTSD) and other anxiety disorders, and schizophrenia [8]. Each of these conditions is readily measured using diagnostic interviews (structured and/or semistructured), psychological tests, and questionnaires. Yet standardized assessments are often limited to funded clinical research studies in the VA, with clinicians' diagnoses predominating in the electronic medical record (EMR). Research that exclusively employs the EMR for determining diagnostic status in mental health (e.g., epidemiological or health services studies) needs to acknowledge the limitations of diagnoses determined in this manner. Additionally, other information pertinent to clinical intervention is typically not captured, such as information specific to functioning, further limiting the information captured by relying on diagnoses and the EMR as a data source. Nonetheless, diagnostic assessment remains central to mental health research, and evaluations can be conducted in multiple ways for research that involves specific inclusion and exclusion criteria related to diagnostic status. The next section provides examples of structured diagnostic interviews and psychological questionnaires widely used in research settings, while the importance of broader inclusion of functional assessment will be addressed in detail later.

Diagnostic Interviews

The need for reliable and valid diagnostic tools in research studies has led to the widespread availability of instruments such as the Structured Clinical Interview for DSM-IV (SCID) [9], the Anxiety Disorder Interview Schedule (ADIS) [10], and the Clinician-Administered PTSD Scale (CAPS) [11]. These standardized assessments have greatly improved the level of standardization in diagnostic assessment. The Composite International Diagnostic Interview (CIDI) [12] is a particularly promising instrument. While the previously mentioned measures require a clinician to conduct the assessment, the CIDI features many

aspects of the other measures and can be administered by someone without advanced clinical training. Its potential use by psychological technicians is an advantage supporting its use.

Each of these diagnostic tools is premised on the existing diagnostic criteria in use and directly follows the methods suggested by the DSM for making psychiatric diagnoses. The SCID has the virtue of providing coverage for all axis I and II disorders but is limited by its focus on a binary (present or absent) decision-making process imposed on the clinician. The ADIS covers anxiety disorders and the most common comorbidities associated with them. It has the advantage of measuring the symptoms of disorders in dimensional ways. Its limitation is the restricted range of disorders included in its modular form. The CAPS, the most widely used diagnostic measure for PTSD in the world, employs the same dimensional approach for each symptom as well as the associated features of PTSD. It is, however, limited to the measurement of PTSD and related problems.

For evaluating people for schizophrenia in the United States, the SCID is the most common diagnostic instrument employed in research. It has the virtue of high levels of sensitivity and specificity in the diagnosis of schizophrenia, but the limitation of requiring administration by a well-trained clinician; further, it is costly in terms of time for clinician and patient when several modules are included in an evaluation. The Positive and Negative Symptoms Scale is a dimensional measure for the central features of schizophrenia and possesses excellent reliability and validity [13]. It is a 30-item instrument that is administered by trained clinicians. Its use in treatment outcome studies is largely secondary to its capacity to detect changes in these symptoms as a function of the clinical intervention provided.

While the DSM uses a primarily categorical approach to psychiatric disorders, focusing on the dimensional aspects of psychiatric conditions is important because of the many recent studies using taxonomic analyses to assess the extent to which a disorder is dimensional or continuous in nature [14–15]. These studies, in the aggregate, find that the psychiatric disorders are indeed dimensional in nature. Moreover, dichotomous data are much less powerful in data analytic strategies than dimensional measures. Thus, the use of dimensional measurement is to be encouraged in research within the VA. The use of the CAPS, for example, can provide both a dimensional measure and a dichotomous meas-

ure of PTSD and is widely recommended for use in research on that disorder [11].

Similarly, the Hamilton Rating Scale for Depression provides dimensional measurement of the many symptoms of depression and, for more than 50 years, served as the primary outcome measure for studies of mood disorders, such as major depressive disorder [16]. Its widespread use in clinical trials is standard, and it provides an opportunity to examine treatment effects across studies, patient populations, and interventions.

Psychological Tests and Questionnaires

Self-report measures are typically used for large-scale survey-based studies. They are frequently used as screens for identifying potential participants in a research study. They are more time- and cost-efficient than structured interviews and can allow for anonymity. Furthermore, self-report questionnaires evaluate symptom severity while most interview schedules do not. A key issue in using a self-report measure is to ensure that it has been normed on the target population for optimal accuracy and efficiency [17]. Additional strengths and weaknesses of self-report and other assessment styles are addressed later in the “Sources of Mental Health Data” section. This article presents examples of widely used questionnaires next.

The Beck Depression Inventory (BDI) [18] and its revision (BDI-II) [19] are the most widely used psychological questionnaires for depression and are considered the central tools for measuring depressive symptoms dimensionally. The Patient Health Questionnaire-9 is a more recent measure that is also well-validated and provides a simpler and briefer assessment than the BDI [20]. The PTSD Checklist is the most commonly used brief measure of PTSD in the scientific literature [21]. Its 17 items make it a brief, but highly sensitive, measure of PTSD with good specificity.

The Alcohol Use Disorders Identification Test (AUDIT) is a relatively brief questionnaire that assesses alcohol use [22]. The AUDIT-Lite is the most efficient psychometrically sound alcohol use questionnaire available, consisting of 10 items and taking about 2 minutes to administer. It identifies people with heavy drinking histories. The Michigan Alcoholism Screening Test consists of 25 items, takes 8 to 10 minutes to complete, is widely used, possesses good psychometric properties, and is best at identifying alcohol use disorders rather than the upper range of drinking histories [23].

The Drug Abuse Screening Test (DAST) covers a range of potential substances of abuse [24]. It has two forms (10 items and 28 items) that take 5 and 10 minutes, respectively, to complete. The focal point of this assessment instrument is on the negative consequences of drug use. Primarily used as a screen to determine who might need further evaluation, the DAST is widely used in primary care, medical, and psychiatric settings to assist in identifying those for whom intervention might be warranted. Available data on internal reliability and validity suggest that this measure deserves its reputation as the screening measure of choice for drug use over the prior year.

Symptom measurement for schizophrenia has, until recently, relied on clinician ratings of severity. This is, of course, expensive in terms of time and professional staffing. The Revised Behavior and Symptom Identification Scale (BASIS-R) was designed to overcome this fundamental problem in evaluating patients with serious mental disorders [25]. Initial studies are promising because the BASIS-R appears to discriminate patients with psychotic conditions from patients without psychotic conditions, and efforts to computerize the instrument appear to be largely successful [26]. The BASIS-32, a shortened form of the original instrument, appears to be a valuable addition to the tools available for investigators studying general psychiatric populations, including those with schizophrenia and other psychotic conditions.

The number of instruments available to evaluate specific conditions is growing exponentially. Earlier, this article outlined many of the most widely used and well-validated instruments for investigators to consider. A variety of other instruments in each domain are available for assessment, with good reliability and validity. Some are designed for subgroups (e.g., the Geriatric Depression Scale [27]), while others represent somewhat different theoretical constructs. These can characterize patients for inclusion and/or exclusion in research studies and also measure change in symptoms and diagnostic status as a function of interventions. The key point is that while room still exists for improvement in basic diagnostic and symptom assessment of specific disorders, this area of assessment is much more established than those reviewed next.

Implications for Directions in Measurement Development

Among the limitations of the specific measurement approach in VA settings is that a large proportion of patients present to clinics with multiple comorbidities in the psychiatric and physical health domains [5]. While improvement

in the symptoms related to the disorder is clearly an important goal, researchers in rehabilitation are interested in outcomes that extend beyond symptoms. As a result, the field needs measurements that will permit examining a range of outcomes across patient groupings. For example, assessing whether participants in compensated work therapy improve in employment and interpersonal functioning is critical regardless of whether the participant is experiencing PTSD or schizophrenia. Similarly, individual factors such as readiness to change or different levels of self-efficacy may be important moderators of functional outcomes [28]. Each of these constructs, together or separately, would constitute unique ways to measure the effect of an intervention on patients who carry diverse diagnostic labels.

The focus of assessment needs to increasingly turn toward functional abilities and limitations in the interest of burden reduction. Information on specific conditions will always be needed for the development of treatment plans, and an individual's functional status cannot be known solely on the basis of diagnostic status. However, tracking and monitoring patient outcomes by measuring functional capacity and quality of life are ultimately the targets of choice for determining the extent to which interventions are needed and successful. The next section of this article focuses on current efforts to develop measures that can be widely adapted to use in research within the VA on patients with psychological disorders and physical diseases. This discussion focuses on new developments in conceptualizing functioning that can provide a framework for the next stage of measurement development for mental health researchers, much as the DSM served as a framework for mental health symptom and diagnosis measurement developments.

ASSESSING PSYCHIATRIC-RELATED FUNCTIONAL STATUS

Our ability to adequately and competently assess functioning and functional impairment across domains related to psychiatric conditions has great importance [29]. First, given that the DSM requires an individual to fulfill the clinical significance criterion (i.e., symptoms result in either clinically significant distress or impairment in social, occupational, or other important areas of functioning) for most disorders to meet diagnostic criteria, the means and methods by which we assess psychiatric-related functional impairment affect our understanding of the prevalence and incidence of psychiatric disorders. Second, the means and

methods by which we assess psychiatric-related functional impairment are also important for determining the extent to which various therapies may be considered beneficial. For example, two recent large-scale VA Cooperative Studies examining the effects of group therapy [30] and individual cognitive-behavior therapy [31] for PTSD showed that while PTSD symptoms improved significantly, neither study found improvements on the 36-Item Short Form Health Survey (SF-36) for Veterans (SF-36V) [32]. These results suggest that the treatment did not improve participant functioning or quality of life. Alternatively, these results might suggest that the measures used in those studies were not able to adequately assess the changes in functioning and quality associated with the changes in PTSD symptom improvement. However, more research is needed to determine the extent to which this possibility is correct.

Additionally, the means and methods by which we assess psychiatric-related functional impairment have implications for compensation and pension procedures and decisions for those contending that they are experiencing impairments related to service-related psychiatric conditions. Related to this point are the recent findings of a committee convened by the Institute of Medicine (IOM) to address ongoing concerns about the current procedures used to assess PTSD among veterans in compensation and pension examinations. One of the committee's tasks was reviewing the use of the Global Assessment of Functioning (GAF) in evaluating impairment associated with PTSD [33]. The GAF is a clinician-rated global index of illness severity that ranges from 0 to 100, with higher scores indicating better functioning. In line with the requirement that VA clinicians provide a GAF rating for each of their mental health patients every 90 days, these ratings are used by VA clinicians to track clinical progress and treatment planning. The GAF instructions provided by the DSM-IV, Text Revision [34] state that the GAF rating should be based on an individual's worst functioning within either occupational, social, or psychological domains. As a result, the GAF has been previously criticized for combining psychiatric symptomatology and social-occupational functioning into one score [35] despite the fact that these constructs are distinct. Past research has found that GAF scores are most significantly associated with symptom ratings, rather than social or occupational standing, and that its reliability and validity are dubious [36–38]. The IOM committee further determined that the GAF score has limited utility for assessing disability associated with mental disorders, such as PTSD among veterans, because the GAF scale content reflects its intended emphasis on mood disorder and schizophrenia symptoms.

In an attempt to rectify some of these concerns, a recent effort was undertaken to develop a new version of the GAF, one in which the three domains were rated separately and anchor points for each of these domains were well defined [39]. This new version of the GAF, termed the Mental Illness Research, Education, and Clinical Center (MIRECC) GAF, corrects the conflation of symptom and functional assessment of a single GAF score by generating three GAF scores, which demonstrate high reliability and convergent and predictive validity. However, the social subscale of the MIRECC GAF was associated more with symptom impairment than social functioning, and most patients included in the sample displayed severe levels of symptomatology or impairment. Thus, whether the MIRECC GAF is suited for use with less impaired individuals is unclear. Given the existing limitations of the GAF and MIRECC GAF, as well as recognition by the IOM committee that the VA uses the GAF as an important determinant of disability status, the IOM report recommended that the VA ultimately identify and implement an appropriate replacement for the GAF.

International Classification of Functioning, Disability, and Health as Framework for Functional Assessment in Mental Health Research

A recent advancement in defining functioning that can provide a framework for addressing the question of functional assessment is the World Health Organization (WHO) International Classification of Functioning, Disability, and Health (ICF) [40]. The ICF is based on a biopsychosocial approach that allows users to document the effect of health conditions (as captured by the WHO International Classification of Diseases [ICD] [41]) on human functioning from biological, individual, and societal perspectives. The ICF groups functioning into two basic components: (1) body functions and structures and (2) activities and participation. Categories within body functions and structures cover aspects of the body and what the body structure can do (e.g., structure of the hand and mobility [function] of a joint) while activities and participation capture functions of the whole person (e.g., movement) or the person within the environment (e.g., working). Additionally, the ICF also enables users to classify environmental factors that may interact with any of the domains to affect functioning. In addition to classifying the environmental influences that affect the individual's context, the ICF structure recognizes personal factors (the internal influences on functioning and disability) as important but has not fully articulated this aspect of the classification. Elements of every component interact with elements of the other components, providing a

model that better reflects the reality of human functioning and the fact that interventions targeting one element affect other aspects of functioning. Each of the categories within the components is identified with a neutral code, such as control of thought, followed by a brief description. It is up to the user to determine whether the particular code is applicable for the case at hand and to assign one or more qualifier ratings to indicate any impairment or other change in functioning. As described by Reed et al., the “ICF allows for the comprehensive description of functioning, recognizing human functioning as multifaceted and involving more than just the biological aspects of health. By explicitly delineating real-world performance and environmental factors, the ICF underlines the role of the environment in human functioning. The concepts and assumptions underlying the ICF reflect the values espoused in the practice of rehabilitation psychology: the dignity and worth of all people, inclusion of people with disabilities in society to the fullest extent possible, and the need for advocacy to provide people with disabilities the best opportunity to maximize their independent functioning” [42]. Several authors have emphasized the importance of these values within rehabilitation psychology [43–45].

The ICF makes no assumptions about etiology and is concerned simply with the classification of functioning, not the causes of functional difficulties. As Reed et al. state, “Most interventions provided by psychologists—in rehabilitation, medical, and mental health settings—are aimed at the improvement of functioning in the context of chronic disease or the prevention, delay, or amelioration of the severity and course of illness rather than at the elimination of an underlying disease process. Therefore, functional status is often a better indicator of service needs and treatment outcomes than diagnosis alone, and diagnosis should not be used as the basis for limiting the eligibility of psychologists to provide services aimed at improving functioning” [42]. (For additional discussion, see Bruyère and Peterson [46].)

Individuals who could benefit from psychological and behavioral interventions, regardless of type of health condition, may have better access to those services in systems in which functional status information is routinely incorporated into decision making. Present healthcare systems typically limit psychologists to intervene with individuals with mental health diagnoses. The advent of the health and behavior codes reflects the recognition that psychologists provide needed behavioral services to individuals with primary physical health problems. However, some healthcare systems are not set up to allow easy

access to such services. And for others, regardless of health condition, functional needs may be the primary focus of intervention, and measures designed to gauge changes in functional status, rather than diagnostic status, may be more relevant.

The ICF provides a conceptual framework for considering the elements of human functioning as well as a classification system to capture the wide array of capacities of the human person. The ICF can be used to capture broad functional concepts or, alternatively, the use of specific codes and items can capture unique elements of individual functioning. However, the ICF is not a measurement tool itself. ICF users need to understand the conceptual framework and identify the key concepts appropriate for particular applications. These key concepts are sometimes already captured in existing measures and assessments, but other times, specific assessment strategies need to be developed. Some measures have been linked to the ICF and rules for linking measures have been developed [47–48].

Until ICF use is mandated in U.S. healthcare systems, along with the already mandated ICD, Ninth Revision, Clinical Modification [49], users will need to determine the appropriate application of the ICF for their settings. For many users, this conceptual framework is the logical implementation point to elevate consideration of functioning and the consequences of health and mental health conditions in the planning and evaluation of interventions. By studying the framework, users can select the domains of functioning pertinent to their settings and research and clinical questions and then identify or develop measures and assessment strategies to capture these domains. Written guidance for adopting and implementing the ICF is limited, but some case examples provide useful strategies that can be adapted to specific settings. Rentsch et al. describe the process of implementation of the ICF for a multidisciplinary team in a neurorehabilitation setting [50]. The guidance provided by Reed et al. [42], as well as the multidisciplinary “Procedural manual and guide for a standardized application of the ICF: A manual for health professionals sample and prototype”^{*} spearheaded by the American Psychological Association and WHO, are additional sources of implementation information.

^{*}American Psychological Association. Procedural manual and guide for a standardized application of the ICF: A manual for health professionals sample and prototype. Unpublished manuscript. Washington (DC); 2003.

LIMITATIONS OF CURRENT FUNCTIONAL IMPAIRMENT MEASURES

One of the key points emphasized in the development of the ICF as a complement to the ICD and DSM is the need for distinct measures of functioning and disease, respectively. A number of efforts have been made to assess psychiatric-related functional impairment that would be more compatible with the ICF and better at separating function from disorder than the current GAF. An early measure that assessed multiple domains was the Personal Adjustment and Role Skills Scale, which covers seven areas, including both psychiatric symptoms (agitation/depression, confusion, alcohol and drug use) and functional skills and behaviors (interpersonal involvement, household activity, outside social activity, and work) within the same scale [51]. A more recent measure is the SF-36 [52]. The SF-36 is probably one of the most widely used measures for assessing a range of functional variables. It provides eight domain scores indexing physical functioning, physical role, bodily pain, general health, vitality, social functioning, emotional role, and mental health; in addition, summary physical and mental health scores may be computed. The reliability and validity of the SF-36 are well documented [53].

The WHO Disability Assessment Scale-II (WHODAS-II) [54–55] is another alternative that assesses a wide range of impairment and disability dimensions using multi-item scales: understanding and communicating, mobility, self-care, getting along with others, life activities (includes work, education, and household responsibilities), and participation in society. The WHODAS-II is used across countries and population groups, has high test-retest reliability, and correlates with other measures of functioning, such as the SF-36. It is becoming widely used in investigations of functional disability across many diverse populations, including those with physical illness (i.e., rheumatology, pulmonary, primary care cohorts) and severe mental illness (i.e., schizophrenia cohorts).

However, the task of identifying alternatives to the GAF for assessing psychiatric-related functional impairment may not be simple or easy. The ICF provides a much more comprehensive framework for classifying functioning than the rather gross estimate derived by the GAF, but the ICF is not a measurement tool itself. The available alternatives have their own limitations. For example, a number of the measures that might replace the GAF to assess functional impairment are resource intense and/or require exten-

sive training before use (e.g., WHODAS-II, Social and Occupational Functioning Assessment Scale [SOFAS] [56], Patient-Reported Outcomes Measurement Information System [PROMIS] [57], Longitudinal Interval Follow-up Evaluation [LIFE] [58], Person-In-Environment System [59]). Other measures, such as the SF-36 or WHODAS-II, may be difficult to score or require the purchase of a license or authorization before use. Many possible alternatives to the GAF, such as the Work Limitations Questionnaire [60], Work Productivity and Activity Impairment instrument [61], Social Adjustment Scale [62], or Sheehan Disability Scale (SDS) [63], may be too brief or narrow to capture all domains of interest. In contrast, other measures may be too lengthy or impractical to be included in many healthcare or research settings (e.g., Life Stressors and Social Resources [64]). Some measures focus predominantly on physical health-related impairment (e.g., SF-36, SOFAS, WHODAS-II, PROMIS). Other available measures of functioning require the respondent to make an attribution about the etiology of the impairment in question (e.g., LIFE, SDS, WHODAS-II, SF-36), which may bias results as most individuals are incapable of accurately making such attributions. This state of affairs suggests that the development of a new measure of psychiatric-related functional impairment is warranted. Importantly, any new measure would assess all the pertinent domains of functioning with sufficient breadth and depth without requiring respondents to make attributions regarding the cause of the impairments.

INVENTORY OF PSYCHOSOCIAL FUNCTIONING

As an alternative to the GAF, the Inventory of Psychosocial Functioning (IPF) is a newly developed 80-item self-report measure designed to assess functional impairment across multiple domains experienced by veterans and Active Duty servicemembers [65]. Unlike the other instruments discussed here, the IPF is easy to use and score, is not disorder specific but has relevant content for impairment associated with psychiatric disorders, distinguishes between symptoms and impairment, does not require attributions regarding the cause of the impairments, and is usable in both research and clinical contexts.

The IPF yields a total score for each of seven subscales: romantic relationships with a spouse or partner, family relationships, work, friendships and socializing, parenting, education, and self-care. These scales were created as a result of work with a series of focus groups,

which established the domains of functioning that were particularly relevant to veterans experiencing mental health problems, such as PTSD. Scale total scores are computed by taking the mean of the responses to each item within that subscale, with greater scores indicating greater impairment. Because functioning over the past 30 days is assessed, respondents may skip sections of the instrument that do not apply to them.

Currently, the psychometric properties of the IPF are being tested with a sample of 300 male and female veterans at the VA Boston Healthcare System, with representation from the Vietnam, Persian Gulf, and Operation Iraqi Freedom/Operation Enduring Freedom conflicts. Thus far, based on data collected from 236 veterans, the IPF demonstrates excellent psychometric properties. The IPF demonstrates strong internal consistency, with Cronbach alphas ranging from 0.79 to 0.90. The IPF correlates significantly with a number of other self-report measures of impairment and quality of life, such as the SDS, WHODAS-II, SF-36V, Quality of Life Inventory (QOLI) [66], and GAF. Specifically, scores on the social and interpersonal IPF subscales (i.e., romantic relationships, family, friendships and socializing, parenting) correlate significantly with associated scales in other measures, with correlations ranging from $r = -0.20$ to $r = -0.67$ (all $p < 0.05$) (i.e., QOLI Love, Friends, Children, and Relatives scales; SF-36V Social Functioning scale; WHODAS-II Getting Along with People domain; SDS Social Life and Family Life/Home Responsibilities scales; GAF). Scores on the occupation and education IPF subscales correlate significantly with other occupational scales in other functioning measures, with correlations ranging from $r = -0.26$ to $r = 0.69$ (all $p < 0.05$) (i.e., SF-36V Role-Emotional scale; WHODAS-II Life Activities domain; SDS Work/School scale; GAF). Lastly, scores on the IPF Self-Care subscale correlate significantly with several similar scales in other measures (i.e., QOLI Health, Money, Play, and Creativity scales; SF-36V Physical Health and Mental Health component summary measures, and also significantly correlated with each of the SF-36V scales; WHODAS-II Getting Around, Self Care, and Life Activities domains; GAF) with correlations ranging from $r = -0.30$ to $r = 0.63$ (all $p < 0.01$).

IPF scores are significantly associated in the expected direction with PTSD and depression symptom severity. IPF scores are also positively correlated with number of days in the past month participants reported experiencing emotional problems as well as in the expected direction with being a current recipient of pen-

sion and/or compensation for a psychiatric disability and receiving medication treatment for psychological and/or emotional problems in the past month.

These preliminary findings suggest that the IPF is already a viable option for clinicians and researchers who need to assess psychiatric-related functional impairment. However, more research with a representative sample of patients with other disorders and difficulties is needed to determine whether it can serve as a comprehensive alternative to the GAF. In particular, additional evaluation with patients experiencing comorbid conditions is needed. However, the areas assessed by the IPF are relevant across domains and should be generalizable to patients with a variety of psychiatric conditions.

ACCOUNTING FOR CULTURE IN MENTAL HEALTH RESEARCH AND TREATMENT IN VA

The VA healthcare system treats a remarkably diverse patient population. It is, therefore, of utmost importance to meet the diverse cultural needs of individuals being served within this system. Mental health professionals and all clinicians are tasked with considering, integrating, and applying culturally competent and specific approaches to assessment and treatment to optimize the care provided for patients from diverse racial and ethnic backgrounds. Delivering innovative assessment methods, interventions, treatment regimens, prevention strategies, and social services within the sociocultural context of the patient is a prime objective of VA-based healthcare of all types.

Race, ethnicity, religion, sexual orientation, gender, and a host of other factors greatly affect the response of patients to prescribed pharmacological and behavioral interventions. Understanding the implications of the patients' background for assessment and treatment is required to reduce barriers and increase access to any healthcare to be provided. Recently, VA has seen an influx in the number of women veterans coming for healthcare of all types. The current conflicts in Iraq and Afghanistan have dramatically increased the proportion of female military servicemembers, such that 15 percent of Active Duty, 17 percent of National Guard/Reserves, and 20 percent of new military recruits are women [67]. This change, no doubt, has already and will continue to have an effect on the services provided by VA clinicians. This shift in demographics will ultimately result in sex-sensitive research and

care in both the mental and physical health arenas, given the specific needs of women veterans [68].

In attempting to develop measures of diagnosis, severity of condition, and functional impairment, researchers are encouraged to include sufficient numbers of patients who are representative of these various minority groups to ensure that assessment methods operate in the same way for all groups in the VA healthcare system. Many examples of this broad-based approach to instrument development are present in the literature on psychological assessment. Providing information on sensitivity, specificity, and overall utility of a measure for each group included in a study is a worthwhile study goal for all large-scale VA research.

SOURCES OF MENTAL HEALTH DATA

In addition to the range of domains that might be assessed in mental health outcomes research (e.g., diagnosis, symptoms, functioning), the data collection methods

and sources of data are worth considering. Assessment tools differ not only in their content, but also in the structure and process of the assessment. This consideration is particularly relevant and challenging in the realm of mental health rehabilitation research since cognitive impairment and temporal fluctuations in mental status and symptoms, as well as other psychosocial factors, can greatly influence assessment results.

The major methods and sources of data typically used in mental health research are participant self-report, participant interviews, gathering collateral information from family members or clinicians who know the participant, clinician ratings, observational ratings of behavior, and performance-based assessments. The **Table** provides an overview of the key advantages and disadvantages of each type of data source, although note that this **Table** does not represent an exhaustive review and that quite a bit of variability exists among measures within each category based on the characteristics described.

Table.

Sources of mental health assessment data.

Data Source	Advantage	Disadvantage	Example
Self-report Questionnaires	<ol style="list-style-type: none"> 1. Standardization. 2. Normative data available. 3. Efficient. 4. Scored objectively. 5. Can gather information on internal experiences. 	<ol style="list-style-type: none"> 1. Recall biases. 2. Demand characteristics. 3. Under-/over-reporting. 4. Based on respondent's subjective interpretation. 5. Influenced by level of respondent insight. 	Beck Depression Inventory, Minnesota Multiphasic Personality Inventory
Structured Diagnostic Interviews	<ol style="list-style-type: none"> 1. Standardization. 2. Capacity to individualize. 	<ol style="list-style-type: none"> 1. Requires trained interviewers. 2. Subjective/selective reporting. 3. Time intensive. 	Structured Clinical Interview for DSM Disorders, Clinician-Administered PTSD Scale
Collateral Reports	<ol style="list-style-type: none"> 1. Objective/independent data source (relative to self-report). 2. Can be used with self-report. 3. Different perspective from self-report. 	<ol style="list-style-type: none"> 1. Incomplete information. 2. Recall biases. 3. Subjective/selective reporting. 	Independence Living Skills Survey, Child Behavior Checklist
Clinical Rating/Observations	<ol style="list-style-type: none"> 1. Objective/independent data source (relative to self-report). 2. Can be used with self-report. 3. Different perspective from self-report. 	<ol style="list-style-type: none"> 1. Incomplete information. 2. Time intensive. 	Hamilton Rating Scale for Depression, Brief Psychiatric Rating Scale
Observation and Performance-based Assessment	<ol style="list-style-type: none"> 1. Objective. 2. Ecological validity. 3. Scored objectively. 	<ol style="list-style-type: none"> 1. Time intensive. 2. Require trained raters. 3. Demand characteristics if patient aware of being observed. 	Work Behavior Inventory, Behavioral Approach Test, Maryland Assessment of Social Competence

DSM = Diagnostic and Statistical Manual of Mental Disorders, PTSD = posttraumatic stress disorder.

Decisions about data sources and data collection methods in mental health rehabilitation research depend on a number of factors, including the characteristics of the setting, purposes of data collection, accessibility of alternate data sources, and availability of assessment resources and training. Ideally, assessment data are not only valid and reliable for the purposes of research but are also usable for enhancing clinical care and outcomes. Particularly, if data are collected across a period of rehabilitation and recovery, it can be useful to provide feedback to the patient, his or her health and mental health care providers, and/or family members. Providing feedback on assessment data collected in the context of research also gives the research team an opportunity to collaborate with mental health providers and other clinicians about the specific research project being undertaken, provide them with information that is not likely to be available to clinicians, and inform them of the issues and hypotheses contained within the scientific study.

It is fairly common in rehabilitation settings to gather data from multiple sources. When multiple data sources are used, an important second step is to synthesize these data to arrive at a more complete understanding of the discrepancies as well as the consistencies in the data. For example, it is not unusual for family members to have a very different perception of their loved one's symptoms or functioning than the individual has of him or herself. In some situations, merely documenting these differences may be adequate (and quite useful), while in other situations it may be important to have procedures in place for clarifying, intervening, or providing feedback when discrepancies are noted (e.g., suicidal ideation or behavior). Optimizing the accuracy and usability of assessment data requires careful consideration of the methods and sources of data collection and course of action when discrepancies between data sources exist.

LOOKING TOWARD THE FUTURE: ROLE OF TECHNOLOGY IN ASSESSMENT

The technological revolution in terms of computerization and digital technology has only recently begun to be incorporated into the area of mental health assessment. While several examples of computerized assessment and measurement exist [69–70], the potential of these technologies for improving scientific inquiry are not yet used to maximum benefit. We view this as one of the critical direc-

tions for development of assessment in mental health research. In this section, we briefly review the current use of technology in the assessment of mental health variables, identify potential areas of development, discuss obstacles that interfere with adopting technological options, and recommend strategies for further development.

Current Use of Technology

As noted previously, core strategies for the assessment of most variables in mental health research include questionnaires, interviews (either of participants or collaterals), observation, and performance-based assessments. The benefits of computer-administered instruments include improved accuracy of assessment and improved ease of assessment (which can increase utilization in clinical as well as research settings). Recent studies demonstrate that such computerized self-report measures are reliable and valid, including versions administered via the Internet [71–73] and using voice-recognition technology [74]. Development of Internet-based assessment can reach thousands of participants who otherwise would not have access to involvement in research (e.g., Muñoz et al. [75]). Several diagnostic interviews have also been designed for computer-aided administration [69]. Comparisons with standard measures generally find equivalent outcomes, and in some cases, the computerized systems lead to identification of more diagnoses.

One of the advantages of computerized assessment is increased portability and therefore ecological validity. Through the use of telemetric devices (both passive, gathering information through physiological monitoring, and interactive, using cellular telephones and other personal digital assistants), several research groups are using telemetry to gather data in mental health research [70]. Participants can be prompted, for example, to fill out a mood scale at different times of the day, while observational data can be gathered through physiological or motion sensors. Such portability allows for more ecologically valid assessment of variables and reduction on dependence of retrospective reporting.

Potential Areas of Development and Expansion

The range of areas in which technology can improve assessment is nearly limitless. However, several specific domains are particularly relevant. At a most basic level, converting paper and pencil forms to computerized assessments is a critical direction of development. Direct data entry in a computerized assessment can greatly reduce assessment

time and cost while increasing accuracy. However, beyond computerizing individual measures, improving the technological infrastructure in institutions to support computerized assessment is necessary. While on a large scale the advantages of computerized assessment are clear in savings in data management expenses, it may be less true for smaller studies. For individual researchers, setting up separate databases for study management can be costly up front, and for institutions, having hundreds or thousands of separate data systems can increase vulnerability for data loss. At an institutional level, developing a database infrastructure that can be shared by researchers for managing data collected using computerized assessments can potentially reduce costs, greatly improve the efficiency of studies, and potentially improve data security. In particular, large institutions such as the VA and major academic institutions are well positioned to develop such database resources.

Obstacles and Recommendations

While significant strides have been made in using computer technology in mental health research assessment, a range of obstacles to widespread adoption of these methods exists. Concerns about such approaches include cost, information security, acceptability of computerized assessments to participants (particularly subgroups that may be less comfortable with computers), and potential validity questions. The extensive development of resources for assessing, validating, and disseminating computerized assessments is critical to effective adoption.

At a very basic level is the need to validate instruments when they are transported to computer-based strategies. Interaction with a computerized assessment can potentially change responses by either increasing validity (when answering about sensitive or stigmatizing information) or by decreasing it (participants misunderstanding questions and not having someone available to clarify them). Continued efforts to complete such validation analyses are necessary.

At a research level, most mental health researchers do not have technical knowledge in computer programming or other information technology skills. Developing additional expertise is needed to improve the use of computerized assessments. To address this, funding agencies can encourage studies to incorporate technological innovation, provide support for the staffing to include technology experts in studies, and include technology experts in funding review boards. Much as the incorporation of biostatisticians has been an important strategy for improving

the research design and data analysis of studies, the incorporation of technology experts at both the study and review level will be critical to moving mental health assessment forward.

At an institutional level, an obstacle in the use of computer technology in mental health research assessment is concern about information security and privacy. For example, there has been extensive growth in the use of handheld devices for both gathering information and providing interventions [76], but often institutions have significant restrictions on the use of wireless technology for safety, security, and/or privacy reasons. These valid concerns need to be addressed in a manner that does not preclude widespread adoption of technology such as portable assessment devices. We advocate that institutions invest in the development of uniform guidelines for use of wireless technology that facilitate research activities while addressing security and privacy concerns.

Currently, in our experience, many current computer networks in academic and VA settings are challenging to navigate. For example, the VA currently has a national integrated database and EMR. It has a broad range of computer-based assessments that are common research measures incorporated into the medical record. Unfortunately, in our experience these resources are underutilized for research because of lack of dissemination to staff regarding availability, lack of knowledge regarding how to access and use data, and administrative obstacles to use caused by privacy and information security concerns. Increased communication between clinical and research administrations in institutions such as the VA could greatly improve the ability to benefit from technology in conducting research while protecting patient rights.

Funding for information technology resources in research should be encouraged and emphasized to reduce overall costs. Support for secure networks to support multiple research projects and standardization of data collection, storage, and management would potentially provide a dramatic increase in efficiency. Theoretically, funding organizations such as the National Institutes of Health or VA's Office of Research and Development could develop secure database infrastructures and encourage researchers to make use of these resources. Such an infrastructure could improve the ability to share and manage data. Some projects may not be compatible with such a resource, and there are scientific and ethical reasons why some research studies would not be appropriate to include. However, the development of such an infrastructure could greatly enhance the efficiency of

mental health research on a national and potentially international scale.

OPPORTUNITIES FOR RESEARCH AND QUALITY IMPROVEMENT

We would like to offer two broad directions for further work in developing state-of-the-art outcome measures in rehabilitation research. The directions are based on organizational strengths of integrated health delivery systems with an EMR and represent both opportunities for systems and reasons for increased development of coordinated EMRs nationally. They are (1) understanding the relationship between subjective mental health measures and biological measures and (2) using measures that are compatible with those needed at other organizations such as the Veterans Benefits Administration and the Social Security Administration.

Understanding Relationship Between Subjective Mental Health Measures and Objective Biological Measures

Healthcare systems collect a host of biological data that are potentially of great interest as they relate to psychological constructs. Foremost on this list are genetic data. As it becomes more feasible to collect detailed genotyping as a routine part of VA treatment, the opportunity to access this database and relate the genetic data to psychological variables of interest exists. Other biological data are relevant to mental health functioning in rehabilitating veterans—electrocardiograms and heart-rate variability as a measure of stress, brain computed tomography and magnetic resonance imagery scans, thyroid function tests, and even cardiovascular measures such as cardiac catheterizations and exercise tolerance tests.

The accelerating collection of genetic data has implications for the type and quality of data collected. One implication is that broad categorical measures are likely to be much more useful when they are augmented by continuous measures of discrete domains. For example, a clinical diagnosis of Alzheimer Type Dementia (which is a dichotomous yes/no categorical measure) will be more useful when analyzed in conjunction with the Montreal Cognitive Assessment [77]. Still more valuable would be data that include responses to individual items on this test. Some constructs might best be evaluated with brief neuropsychological tests that examine specific variables

(e.g., short-term memory, impulsive responding). Some researchers have argued that because psychiatric disorders are not governed by simple Mendelian inheritance, understanding the role of particular alleles on psychiatric conditions will be most fruitful by relating particular allelic variants to small changes in specific neurological or psychiatric functions.

Biological variables, self-report, and interview-based data can be more efficiently and effectively managed by harnessing the power of integrated EMRs. We envision a reciprocal, iterative benefit. As healthcare systems develop better computer databases, research can better assess the relationship between these variables. Such data can then guide clinical service.

Establishing Measures That Are Compatible with Those Needed at Other Organizations

While much of research focuses on clinical treatment and outcomes, we need to develop new assessments that also address disability and compensation needs. For example, the VA system conducts thousands of compensation and pension examinations per year of veterans who are applying for service-connected disability for a psychiatric disorder. To date, the assessment procedures and measures used in these examinations have not been subjected to rigorous empirical research. In the absence of data, the mental health claims process has been controversial. Award rates vary widely by state, and this and other data have raised questions about the reliability of mental health examinations to determine service connection. The Veterans Health Administration is implicated in the criticism of these controversial examinations. Standardization of these examinations across the country would go a great distance toward enhancing the quality, accuracy, and fairness of this assessment process.

VA staff are trained to write reports based on template reports on a specialized worksheet (Automated Medical Information Exchange). This worksheet is arguably the most important functional assessment of mental health functioning that a veteran will undergo. An opportunity exists to use more standardized assessments in the compensation examination and for measures to be used in clinical care that document the disabilities that veterans are evaluated for on compensation and pension examinations. An innovative, ongoing VA Qualified Enhancement Research Initiative (QUERI) is on the forefront of this effort. In this QUERI study, veterans presenting for compensation and pension examinations for service-connected disability are being randomly

assigned to their usual examination or to an examination incorporating the CAPS and the WHODAS-II. Standardizing these examinations nationwide is an important first step in scientifically improving the quality of these disability examinations.

CONCLUSIONS

In the interest of expanding rehabilitation research, the purpose of this article was to highlight the existing methods for assessing psychiatric conditions and to offer a range of alternative ways to measure outcome, collect information, and synthesize data. As the VA's efforts to expand rehabilitation increase, a distinct need exists to improve the types of measures and the measurement strategies that are employed. This will enhance the quantity and the quality of the research completed; it will also enhance the contribution of the VA to the development of new knowledge in this field, as well as serve veterans better [78]. Moving in these directions will expand the reputation of the VA as the world leader in mental health care and research.

Our interest was also to recognize the inherent limitations of extant measurement approaches in populations that are burdened by high rates of comorbidity in both the psychiatric and physical health arenas. Inclusion in studies of global measures of functioning, self-efficacy, motivation, etc., will permit us to draw cross study comparisons among individuals with a specific disorder (e.g., PTSD) or across different populations (i.e., patients with bipolar disorder vs those with schizophrenia). Utilizing the ICF as a framework may provide a basis for identifying appropriate functional targets for intervention when making decisions about new measures. Agreement as to which of these measures ought to be included remains a question to address in future work; getting researchers to agree on the concept is a good first step to take.

The presence of improved technological resources will drive many of the future changes in mental health assessment in the field. Using item response theory may help us minimize the burden placed on participants; this would not be even a possibility were modern computers unavailable. Handheld devices are now widely used in assessment, and as a result, the ecological validity of data procured through these devices is greatly enhanced. Communication systems allow transport of data from the patient's environment to the research laboratory on an ongoing basis in a secure and confidential manner. These advances are remarkable.

Finally, the use of computers now makes multivariate data analyses possible on unthinkably large samples of patients. Statistical packages are accommodating these design parameters readily and easily.

The future of research on rehabilitation is now; we are living in very exciting times. Designing research studies to include the earlier suggestions will further our comprehensive understanding of what are complex disorders and recovery processes. More thoroughly and accurately assessing symptoms and functioning in mental health will lead to better options for the people receiving care. Let's use this opportunity to optimize the outcomes for patients, their families, and their communities.

ACKNOWLEDGMENTS

Author Contributions:

Concept and design: J. R. McQuaid.

Drafting of manuscript: J. R. McQuaid, B. P. Marx, M. I. Rosen, L. F. Bufka, W. Tenhula, H. Cook, T. M. Keane.

Critical revision of manuscript for important intellectual content: J. R. McQuaid, B. P. Marx, M. I. Rosen, L. F. Bufka, W. Tenhula, H. Cook, T. M. Keane.

Administrative support: H. Cook.

Project supervision: T. M. Keane.

Financial Disclosures: The authors have declared that no competing interests exist.

Funding/Support: This material was based on work supported by the VA Rehabilitation Research and Development (RR&D) Service (grant F6441R to Dr. McQuaid) and U.S. Army Medical Research and Materiel Command (grants W81XWH-08-2-0028 and W81XWH-08-2-0100 to Dr. Marx and Dr. Keane).

Additional Contributions: The authors would like to acknowledge the VA RR&D Service for convening the Working Group on Community Reintegration State-of-the-Art conference for VA Rehabilitation Outcomes, which led to this article.

REFERENCES

1. Diagnostic and statistical manual of mental disorders. 3rd ed. Washington (DC): American Psychiatric Association; 1980.
2. Hunsley J, Mash EJ. A guide to assessments that work. New York (NY): Oxford University Press; 2008.
3. Nathan PE, Gorman JM. A guide to treatments that work. 3rd ed. Oxford (UK): Oxford University Press; 2007.
4. Breslau N, Davis GC, Andreski P, Peterson E. Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Arch Gen Psychiatry*. 1991;48(3): 216–22. [PMID: 1996917]

5. Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB. Posttraumatic stress disorder in the National Comorbidity Survey. *Arch Gen Psychiatry*. 1995;52(12):1048–60. [\[PMID: 7492257\]](#)
6. Kulka R, Schlenger WE, Fairbank JA, Hough RL, Jordan BK, Marmar CR, Weiss DS, Grady DA, Cranston A. Trauma and the Vietnam war generation: Report of findings from the National Vietnam Veterans Readjustment Study. New York (NY): Brunner/Mazel; 1990.
7. Keane TM. Improving methods models and measures: Contributions of CITRM to the field of psychological trauma. *J Trauma Stress*. 2009;22(6):622–23. [\[PMID: 19960545\]](#)
8. Ilgen MA, Bohnert AS, Ignacio RV, McCarthy JF, Valenstein MM, Kim HM, Blow FC. Psychiatric diagnoses and risk of suicide in veterans. *Arch Gen Psychiatry*. 2010; 67(11):152–58. [\[PMID: 21041616\]](#)
<http://dx.doi.org/10.1001/archgenpsychiatry.2010.129>
9. First M, Spitzer R, Williams J, Gibbons M. Structured clinical interview for DSM-IV. In: Rush AJ, editor. Handbook of psychiatric measures. Washington (DC): American Psychiatric Press; 2000. p. 49–53.
10. Brown TA, DiNardo PA, Barlow DH. Anxiety disorders interview schedule for DSM-IV (ADIS-IV) Clinician's manual. Albany (NY): Graywind Publications; 1994.
11. Weathers FW, Keane TM, Davidson JR. Clinician-Administered PTSD Scale: A review of the first ten years of research. *Depress Anxiety*. 2001;13(3):132–56. [\[PMID: 11387733\]](#)
<http://dx.doi.org/10.1002/da.1029>
12. Kessler RC, Üstün TB. The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res*. 2004;13(2):93–121. [\[PMID: 15297906\]](#)
<http://dx.doi.org/10.1002/mpr.168>
13. Kay SR, Opler LA, Fiszbein A. Significance of positive and negative syndromes in chronic schizophrenia. *Br J Psychiatry*. 1986;149:439–48. [\[PMID: 3814927\]](#)
<http://dx.doi.org/10.1192/bjp.149.4.439>
14. Ruscio J, Ruscio AM. Informing the continuity controversy: A taxometric analysis of depression. *J Abnorm Psychol*. 2000;109(3):473–87. [\[PMID: 11016117\]](#)
<http://dx.doi.org/10.1037/0021-843X.109.3.473>
15. Ruscio AM, Ruscio J, Keane TM. The latent structure of posttraumatic stress disorder: A taxonomic investigation of reactions to extreme stress. *J Abnorm Psychol*. 2002;111(2): 290–301. [\[PMID: 12003450\]](#)
<http://dx.doi.org/10.1037/0021-843X.111.2.290>
16. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56–62. [\[PMID: 14399272\]](#)
<http://dx.doi.org/10.1136/jnnp.23.1.56>
17. Keane TM, Barlow DH. Posttraumatic stress disorder. In: Barlow DH, editor. *Anxiety and its disorders*. 2nd ed. New York (NY): Guilford Press; 2002.
18. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961;4:561–71. [\[PMID: 13688369\]](#)
19. Beck AT, Steer RA, Brown GK. BDI-II, Beck Depression Inventory: Manual. San Antonio (TX): Psychological Corp; 1996.
20. Kroenke K, Spitzer RL. The PHQ-9: A new depression diagnostic and severity measure. *Psychiatr Ann*. 2002; 32(9):509–21.
21. Weathers FW, Litz BT, Herman DS, Huska JA, Keane TM. The PTSD Checklist PCL: Reliability validity and diagnostic utility. Annual Meeting of the International Society Traumatic Stress Studies; 1993; San Antonio, TX.
22. Babor TF, Biddle-Higgins JC, Saunders JB, Montiero MG. AUDIT, the Alcohol Use Disorders Identification Test: Guidelines for use in primary care. Geneva (Switzerland): World Health Organization; 2001.
23. Selzer ML. The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *Am J Psychiatry*. 1971;127(12):1653–58. [\[PMID: 5565851\]](#)
24. Skinner HA. The drug abuse screening test. *Addict Behav*. 1982;7(4):363–71. [\[PMID: 7183189\]](#)
[http://dx.doi.org/10.1016/0306-4603\(82\)90005-3](http://dx.doi.org/10.1016/0306-4603(82)90005-3)
25. Eisen SV, Normand SL, Belanger AJ, Spiro A 3rd, Esch D. The Revised Behavior and Symptom Identification Scale (BASIS-R): Reliability and validity. *Med Care*. 2004; 42(12):1230–41. [\[PMID: 15550803\]](#)
<http://dx.doi.org/10.1097/00005650-200412000-00010>
26. Chinman M, Young AS, Schell T, Hassell J, Mintz J. Computer assisted self-assessment in persons with severe mental illness. *J Clin Psychiatry*. 2004;65(10):1343–51. [\[PMID: 15491237\]](#)
<http://dx.doi.org/10.4088/JCP.v65n1008>
27. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiatr Res*. 1983;17(1):37–49. [\[PMID: 7183759\]](#)
[http://dx.doi.org/10.1016/0022-3956\(82\)90033-4](http://dx.doi.org/10.1016/0022-3956(82)90033-4)
28. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: Applications to addictive behaviors. *Am Psychol*. 1992;47(9):1102–14. [\[PMID: 1329589\]](#)
<http://dx.doi.org/10.1037/0003-066X.47.9.1102>
29. Anthony WA, Cohen MR, Vitalo R. The measurement of rehabilitation outcome. *Schizophr Bull*. 1978;4(3):365–83.
30. Schnurr PP, Friedman MJ, Foy DW, Shea MT, Hsieh FY, Lavori PW, Glynn SM, Wattenberg M, Bernardy NC. Randomized trial of trauma-focused group therapy for posttraumatic stress disorder: Results from a Department of Veterans Affairs Cooperative Study. *Arch Gen Psychiatry*.

- 2003;60(5):481–89. [PMID: 12742869]
<http://dx.doi.org/10.1001/archpsyc.60.5.481>
31. Schnurr PP, Friedman MJ, Engel CC, Foa EB, Shea MT, Chow BK, Resick PA, Thurston V, Orsillo SM, Haug R, Turner C, Bernardy N. Cognitive behavioral therapy for posttraumatic stress disorder in women: A randomized controlled trial. *JAMA*. 2007;297(8):820–30. [PMID: 17327524]
<http://dx.doi.org/10.1001/jama.297.8.820>
 32. Kazis LE, Miller DR, Clark JA, Skinner KM, Lee A, Ren XS, Spiro A 3rd, Rogers WH, Ware JE Jr. Improving the response choices on the Veterans SF-36 Health Survey role functioning scales: Results from the Veterans Health Survey. *J Ambulatory Care Manage*. 2004;27(3):263–80. [PMID: 15287216]
 33. Endicott J, Spitzer RL, Fleiss JL, Cohen J. The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry*. 1976;33(6):766–71. [PMID: 938196]
 34. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Text revision. Washington (DC): American Psychiatric Association; 2000.
 35. Goldman HH. Do you walk to school or do you carry your lunch? *Psychiatr Serv*. 2005;56:419. <http://dx.doi.org/10.1176/appi.ps.56.4.419>
 36. Moos RH, McCoy L, Moos BS. Global Assessment of Functioning (GAF) ratings: Determinants and role as predictors of one-year treatment outcomes. *J Clin Psychol*. 2000;56(4):449–61. [PMID: 10775040]
[http://dx.doi.org/10.1002/\(SICI\)1097-4679\(200004\)56:4<449::AID-JCLP1>3.0.CO;2-8](http://dx.doi.org/10.1002/(SICI)1097-4679(200004)56:4<449::AID-JCLP1>3.0.CO;2-8)
 37. Moos RH, Nichol AC, Moos BS. Global Assessment of Functioning ratings and the allocation and outcomes of mental health services. *Psychiatr Serv*. 2002;53:730–37. [PMID: 12045311]
<http://dx.doi.org/10.1176/appi.ps.53.6.730>
 38. Skodol AE, Link BG, Shrout PE, Horwath E. The revision of axis V in DSM-III-R: Should symptoms have been included? *Am J Psychiatry*. 1988;145(7):825–29. [PMID: 3381926]
 39. Niv N, Cohen AN, Sullivan G, Young AS. The MIRECC version of the Global Assessment of Functioning scale: Reliability and validity. *Psychiatr Serv*. 2007;58(4):529–35. [PMID: 17412856]
<http://dx.doi.org/10.1176/appi.ps.58.4.529>
 40. International Classification of Functioning, Disability and Health (ICF). Geneva (Switzerland): World Health Organization; 2001.
 41. International Classification of Diseases, 10th edition (ICD-10). Geneva (Switzerland): World Health Organization; 2007.
 42. Reed GM, Lux JB, Bufka LF, Trask C, Peterson DB, Stark S, Threats TT, Jacobson J, Hawley J. Operationalizing the International Classification of Functioning, Disability and Health (ICF) in clinical settings. *Rehabil Psychol*. 2005;50:122–31. <http://dx.doi.org/10.1037/0090-5550.50.2.122>
 43. Frank RG, Elliott TR. Rehabilitation psychology: Hope for a psychology of chronic conditions. In: Frank RG, Elliott TR, editors. *Handbook of rehabilitation psychology*. Washington (DC): American Psychological Association; 2000. p. 3–8. <http://dx.doi.org/10.1037/10361-000>
 44. Riggat TF, Maki DR, editors. *The handbook of rehabilitation counseling: Issues and methods*. New York (NY): Springer; 2003.
 45. Scherer MJ, Blair KL, Banks ME, Brucker B, Corrigan J, Wegener JH. Rehabilitation psychology. In: Craighead WE, Nemeroff CB, editors. *The concise Corsini encyclopedia of psychology and behavioral science*. 3rd ed. Hoboken (NJ): John Wiley & Sons; 2004. p. 801–2.
 46. Bruyère, SM, Peterson DB. Introduction to the special section on the International Classification of Functioning, Disability and Health: Implications for rehabilitation psychology. *Rehabil Psychol*. 2005;50(2):103–4. <http://dx.doi.org/10.1037/0090-5550.50.2.103>
 47. Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, Ustün TB, Stucki G. Linking health-status measurements to the International Classification of Functioning, Disability and Health. *J Rehabil Med*. 2002;34(5):205–10. [PMID: 12392234]
<http://dx.doi.org/10.1080/165019702760279189>
 48. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustün B, Stucki G. ICF linking rules: An update based on lessons learned. *J Rehabil Med*. 2005;37(4):212–18. [PMID: 16024476]
<http://dx.doi.org/10.1080/16501970510040263>
 49. International Classification of Diseases ninth revision clinical modification: Physician (ICD-9-CM). Chicago (IL): AMA Press; 2002.
 50. Rentsch HP, Bucher P, Dommen Nyffeler I, Wolf C, Hefti H, Fluri E, Wenger U, Wälti C, Boyer I. The implementation of the ‘International Classification of Functioning, Disability and Health’ (ICF) in daily practice of neurorehabilitation: An interdisciplinary project at the Kantonsspital of Lucerne, Switzerland. *Disabil Rehabil*. 2003;25(8):411–21. [PMID: 12745951]
<http://dx.doi.org/10.1080/0963828031000069717>
 51. Ellsworth RB. Consumer feedback in measuring the effectiveness of mental health programs. In: Struening EL, Guttentag M, editors. *Handbook of evaluation research*. Beverly Hills (CA): Sage Publications; 1975.
 52. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473–83.

- [\[PMID: 1593914\]](#)
<http://dx.doi.org/10.1097/00005650-199206000-00002>
53. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of data quality scaling assumptions and reliability across diverse patient groups. *Med Care*. 1994;32(1):40–66. [\[PMID: 8277801\]](#)
<http://dx.doi.org/10.1097/00005650-199401000-00004>
54. World Health Organization [Internet]. Disability assessment schedule WHODAS II: Phase 2 field trials—Health services research: 12-item interviewer administered version. Geneva (Switzerland): World Health Organization; [updated 2000]. Available from: <http://www.who.int/classifications/icf/12int.pdf>
55. WHO Psychiatric Disability Assessment Schedule (WHO/DAS). Albany (NY): World Health Organization; 1988.
56. Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: A review of measures of social functioning. *Am J Psychiatry*. 1992;149(9):1148–56. [\[PMID: 1386964\]](#)
57. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, Ader D, Fries JF, Bruce B, Rose M; PROMIS Cooperative Group. The patient-reported outcomes measurement information system (PROMIS): Progress of an NIH roadmap cooperative group during its first two years. *Med Care*. 2007;45(5 Suppl 1):S3–S11. [\[PMID: 17443116\]](#)
58. Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott P, Andreasen NC. The Longitudinal Interval Follow-up Evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry*. 1987;44(6):540–48. [\[PMID: 3579500\]](#)
59. Williams JB, Karls JM, Wandrei K. The Person-In-Environment (PIE) system for describing problems in social functioning. *Hosp Community Psychiatry*. 1989;40(11):1125–27. [\[PMID: 2807221\]](#)
60. Lerner D, Amick BC 3rd, Rogers WH, Malspeis S, Bungay K, Cynn D. The Work Limitations Questionnaire. *Med Care*. 2001;39(1):72–85. [\[PMID: 11176545\]](#)
<http://dx.doi.org/10.1097/00005650-200101000-00009>
61. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993;45(5):353–65. [\[PMID: 10146874\]](#)
<http://dx.doi.org/10.2165/00019053-199304050-00006>
62. Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry*. 1976;33(9):1111–15. [\[PMID: 962494\]](#)
63. Sheehan DV. *The anxiety disease*. New York (NY): Scribner; 1983.
64. Moos RH, Fenn CB, Billings AG. Life stressors and social resources: An integrated assessment approach. *Soc Sci Med*. 1988;27(9):999–1002. [\[PMID: 3227394\]](#)
[http://dx.doi.org/10.1016/0277-9536\(88\)90291-2](http://dx.doi.org/10.1016/0277-9536(88)90291-2)
65. Marx BP, Schnurr P, Rodriguez P, Holowka DH, Lunney C, Weathers F, Sloan DM, Kean TM. Development of a functional impairment scale for active duty service members and veterans. In: Lester KM, chair. *Beyond PTSD symptom reduction: Social and health-related benefits of trauma focused treatment*. Meeting for the International Society for Traumatic Stress Studies; 2009; Atlanta, GA.
66. Frisch MB, Cornell J, Villanueva M, Retzlaff PJ. Clinical validation of the Quality of Life Inventory: A measure of life satisfaction for use in treatment planning and outcome assessment. *Psychol Assess*. 1992;4(1):92–101. <http://dx.doi.org/10.1037/1040-3590.4.1.92>
67. Meehan S. Improving healthcare for women veterans. *J Gen Intern Med*. 2006;21:S1–S2. <http://dx.doi.org/10.1111/j.1525-1497.2006.00382.x>
68. Haskell SG, Gordon KS, Mattocks K, Duggal M, Erdos J, Justice A, Brandt CA. Gender differences in rates of depression, PTSD, pain, obesity, and military sexual trauma among Connecticut war veterans in Iraq and Afghanistan. *J Womens Health*. 2010;19(2):267–71. [\[PMID: 20109115\]](#)
<http://dx.doi.org/10.1089/jwh.2008.1262>
69. Garb HN. Computer-administered interviews and rating scales. *Psychol Assess*. 2007;19(1):4–13. [\[PMID: 17371119\]](#)
<http://dx.doi.org/10.1037/1040-3590.19.1.4>
70. Goodwin MS, Velicer WF, Intille SS. Telemetric monitoring in the behavior sciences. *Behav Res Methods*. 2008;40(1):328–41. [\[PMID: 18411557\]](#)
<http://dx.doi.org/10.3758/BRM.40.1.328>
71. Donker T, Van Straten A, Marks I, Cuijpers P. Brief self-rated screening for depression on the internet. *J Affect Disord*. 2010;122(3):253–59. [\[PMID: 19679358\]](#)
<http://dx.doi.org/10.1016/j.jad.2009.07.013>
72. Donker T, Van Straten A, Marks I, Cuijpers P. A brief Web-based screening questionnaire for common mental disorders: Development and validation. *J Med Internet Res*. 2009;11(3):e19. [\[PMID: 19632977\]](#)
<http://dx.doi.org/10.2196/jmir.1134>
73. Lin CC, Bai YM, Liu CY, Hsiao MC, Chen JY, Tsai SJ, Ouyang WC, Wu CH, Li YC. Web-based tools can be used reliably to detect patients with major depressive disorder and subsyndromal depressive symptoms. *BMC Psychiatry*. 2007;7:12. [\[PMID: 17425774\]](#)
<http://dx.doi.org/10.1186/1471-244X-7-12>
74. Muñoz RF, McQuaid JR, González GM, Dimas J, Rosales VA. Depression screening in a women's clinic: Using automated Spanish- and English-language voice recognition. *J Consult Clin Psychol*. 1999;67(4):502–10. [\[PMID: 10450620\]](#)
<http://dx.doi.org/10.1037/0022-006X.67.4.502>
75. Muñoz RF, Lenert LL, Delucchi K, Stoddard J, Perez JE, Penilla C, Pérez-Stable EJ. Toward evidence-based Internet

- interventions: A Spanish/English Web site for international smoking cessation trials. *Nicotine Tob Res.* 2006;8(1):77–87. [PMID: 16497602]
<http://dx.doi.org/10.1080/14622200500431940>
76. Marks IM, Cavanaugh K, Gega L. *Hands-on help: Computer-aided psychotherapy.* New York (NY): Psychology Press; 2007.
77. Nasreddine ZS, Phillips N, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* 2005;53(4):695–99. [PMID: 15817019]
<http://dx.doi.org/10.1111/j.1532-5415.2005.53221.x>
78. Rosen MI. Compensation examinations for PTSD—An opportunity for treatment? *J Rehabil Res Dev.* 2010;47(5):

xv–xxii. [PMID: 20803385]
<http://dx.doi.org/10.1682/JRRD.2010.04.0075>

Submitted for publication August 5, 2010. Accepted in revised form February 23, 2011.

This article and any supplementary material should be cited as follows:

McQuaid JR, Marx BP, Rosen MI, Bufka LF, Tenhula W, Cook H, Keane TM. Mental health assessment in rehabilitation research. *J Rehabil Res Dev.* 2012;49(1):121–38.
<http://dx.doi.org/10.1682/JRRD.2010.08.0143>



