SPECIAL COMMUNICATION

Scale Refinement and Initial Evaluation of a Behavioral Health Function Measurement Tool for Work Disability Evaluation

Elizabeth E. Marfeo, PhD, MPH,[a] Pengsheng Ni, MD, MPH,[a] Stephen M. Haley, PT, PhD,[a,†]
Kara Bogusz, BA,[a] Mark Meterko, PhD,[a] Christine M. McDonough, PT, PhD,[a,b]
Leighton Chan, MD, MPH,[c] Elizabeth K. Rasch, PT, PhD,[c] Diane E. Brandt, PT, MS, PhD,[c]
Alan M. Jette, PT, PhD[a]

From the [a]Health and Disability Research Institute, Boston University School of Public Health, Boston, MA; [b]Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine, Lebanon, NH; and [c]National Institutes of Health, Rehabilitation Medicine Department, Mark O. Hatfield Clinical Research Center, Bethesda, MD.

Abstract

Objective: To use item response theory (IRT) data simulations to construct and perform initial psychometric testing of a newly developed instrument, the Social Security Administration Behavioral Health Function (SSA-BH) instrument, that aims to assess behavioral health functioning relevant to the context of work.

Design: Cross-sectional survey followed by IRT calibration data simulations.

Setting: Community.

Participants: Sample of individuals applying for Social Security Administration disability benefits: claimants (n=1015) and a normative comparative sample of U.S. adults (n=1000).

Interventions: None.

Main Outcome Measure: SSA-BH measurement instrument.

Results: IRT analyses supported the unidimensionality of 4 SSA-BH scales: mood and emotions (35 items), self-efficacy (23 items), social interactions (6 items), and behavioral control (15 items). All SSA-BH scales demonstrated strong psychometric properties including reliability, accuracy, and breadth of coverage. High correlations of the simulated 5- or 10-item computer adaptive tests with the full item bank indicated robust ability of the computer adaptive testing approach to comprehensively characterize behavioral health function along 4 distinct dimensions.

Conclusions: Initial testing and evaluation of the SSA-BH instrument demonstrated good accuracy, reliability, and content coverage along all 4 scales. Behavioral function profiles of Social Security Administration claimants were generated and compared with age- and sex-matched norms along 4 scales: mood and emotions, behavioral control, social interactions, and self-efficacy. Using the computer adaptive test-based approach offers the ability to collect standardized, comprehensive functional information about claimants in an efficient way, which may prove useful in the context of the Social Security Administration's work disability programs.

Archives of Physical Medicine and Rehabilitation 2013;94:1679-86

© 2013 by the American Congress of Rehabilitation Medicine

The Social Security Administration's (SSA) work disability insurance programs are the primary U.S. federal programs, providing financial support to over 9.8 million disabled workers and their families.[1] In 2011, mental health impairments represented one of the largest categories of disabling conditions for which individuals receive Social Security Disability Insurance (SSDI) benefits.[1] The latter fact highlights the need for disability

evaluation processes to include systematic, efficient, and accurate assessment of mental and behavioral health functioning relevant to a person's ability to work.

Recent examinations of the SSA's current disability determination procedures have identified significant conceptual and operational challenges to the current SSA work disability determination processes.[2-4] Conceptually, the current definition of disability used by the SSA is limited to a medical perspective and does not encompass key aspects of the interaction between a person's underlying capabilities and the work environment. Under the SSA, disability is assessed by focusing on specific conditions or symptoms rather than characterizing a person's overall functioning. This conceptual gap is especially problematic when applied to mental health conditions because of the episodic and context-specific nature of many mental health-related disabilities.[5,6]

The SSA's current evaluation process for determining disability includes a 5-stage sequential process that collects medical information data from various health care providers in an effort to characterize the extent to which a person's medical impairment may be limiting their capacity to work.[7,8] Logistically, this process is time consuming and lacks standardized methods for collecting information about the claimant's symptoms, impairments, or functioning.[2,9]

Advanced methodologies in measurement scale development have emerged that provide an opportunity to measure complex, multifactorial aspects of health and functioning related to physical or mental conditions using a standardized and efficient approach. Specifically, these methodologies use item response theory (IRT) to calibrate an item pool, which is then administered through computer adaptive test (CAT).[10] The IRT methods create an instrument that can characterize a person's functional status along multiple dimensions of function or scales rather than being constrained to a single dimension. Further, IRT modeling techniques provide a method for evaluating a measurement tool at both the item and scale level, and CAT makes it possible to use that information to individualize survey content for each claimant.[11] These standardized, computer-based approaches for collecting health status information could prove to be an efficient and accurate option for the SSA, incorporating a more comprehensive assessment of behavioral health functioning into the disability determination processes.

To address some of the measurement challenges associated with the assessment of behavioral health functioning within the context of SSA disability evaluation, we developed a new IRT-/CAT-based instrument—the Social Security Administration Behavioral Health Function (SSA-BH) instrument. Previous work, reported in complementary articles, describes the initial stages of new measurement scale development.[12,13] Findings from these studies support an underlying unidimensional structure of a set of items designed to assess behavioral health functioning along 4 key domains of behavioral health function relevant to work: mood and emotions, self-efficacy, social interactions, and behavioral control.[12,13]

List of abbreviations:

CAT	computer adaptive test
DIF	differential item functioning
IRT	item response theory
PROMIS	Patient Reported Outcomes Measurement Information System
SSA	Social Security Administration
SSA-BH	Social Security Administration Behavioral Health Function
SSDI	Social Security Disability Insurance

The primary objective of this article is to discuss the use of IRT/CAT methods to refine the SSA-BH measurement scales and conduct initial psychometric testing of the instrument. Specifically, this article will (1) describe the process of applying IRT methods to order the items on a continuum indicative of low to high functional ability in a sample of SSA claimants in a way that is meaningful and useful for the purposes of SSA's disability determination processes, (2) present the development of interval level scales in each domain of behavioral health functioning, and (3) discuss the evaluation and testing of the initial psychometric properties of the SSA-BH instrument. A secondary objective is to discuss a comparison of the SSA-BH score distribution of claimant response profiles versus response profiles of a normative sample of U.S. adults for each of the functional scales.

Methods

Participants

The study included a sample of SSA claimants applying for disability benefits and a second comparative sample of U.S. adults to allow norm-based scoring of the SSA-BH instrument. Eligibility criteria for SSA claimants included: ≥21 years of age, able to speak, read, and understand English, and had to have filed the claim on his or her own behalf because of either a mental or both mental and physical condition. Exclusion criteria specified were related to certain mental conditions or symptoms of paranoia, psychosis, autism, intellectual disability, or Down syndrome. SSA claimants who met eligibility criteria were stratified by both SSA region and urban/rural location then randomly selected for participation in the study. For the comparative normative sample, data were collected on a sample of 1000 U.S. adults. The normative data sampling strategy was developed by the YouGov research organization, which used a proximity sample matching method, drawing respondents from a large opt-in Internet panel.[14] The normative sample was to be representative of a U.S. adult population matched on sex, racial/ethnic background, age, and education, weighted equally. A university institutional review board approved this study, and participants in both samples provided informed consent.

Data collection procedures

The data collected from the SSA-BH instrument used identical methods for both the claimant and normative samples. Trained personnel from Westat research organization collected data from claimants by either phone or Internet. Normative data were collected by YouGov research organization via the Internet. In addition to responding to a set of 165 items used for developing the SSA-BH instrument, participants in both samples provided self-reported demographic data specifying their age, sex, marital status, race/ethnicity, and education. Details of development of the survey item and data collection procedures have been described in a previous article.[13]

Data analytic procedures

Descriptive statistics, including frequency distributions, were calculated for each item response category and for all demographic variables for both the claimant and normative samples. Where applicable, responses to items were reverse coded to