

*For the Assessment of
Neurological and Behavioral Function*



TOOLBOX



Welcome to Toolbox

Richard Gershon, PhD, Principal Investigator, ENHRI

Welcome to the first edition of the NIH Toolbox newsletter. On a quarterly basis we will keep you up to date on our development efforts. This initial newsletter is being sent to the over three hundred scientists, consultants and other professionals who have direct contract responsibilities, or who have been identified by their NIH project officer as having specific expertise which may prove helpful in meeting the goals of the Toolbox. In this month's newsletter you can read about the "Why This, Why Now?" of the Toolbox contract from NIH Toolbox project officer Molly Wagster. We also outline the results of our initial Request for Information and overview the final criteria for sub-domain and instrument selection. Feel free to contact me with your comments, questions or insights. Also, if you know of someone who would benefit by receiving this newsletter please send me their contact information (gershon@northwestern.edu).

Why This, Why Now?

Molly Wagster, PhD, NIA



At present, there are many studies that collect information on aspects of neural function (cognition, sensation, movement, emotion); unfortunately, there is little uniformity among the measures used to capture these constructs. Investigators have expressed the need for brief assessment tools that could be used as a form of 'common currency' across diverse study designs and populations. When individual studies employ unique assessment batteries, comparisons between studies and combining data from multiple studies can be problematic, if not impossible.

In 2005, a report (Hendrie et al., *Alzheimer's & Dementia*, vol. 2, 12-32, 2006) was delivered to the staff of the trans-NIH Cognitive and Emotional Health Project (CEHP) that suggested, among other ideas, that "standard questionnaires to measure cognitive and emotional health" should be developed, for many of the reasons stated above. Around the same time the report was being generated, fifteen Institutes and Centers at NIH that support neuroscience research formed a coalition called the Blueprint for Neuroscience Research. The stated aim of the NIH Blueprint is to develop new tools, resources, and training opportunities to accelerate the pace of discovery in neuroscience research. Armed with the CEHP report, NIH staff successfully petitioned the NIH Blueprint to consider the development of an assessment tool for neurological and behavioral function for funding through this initiative. As a result, in September 2006 a contract was awarded to Dr. Richard Gershon and the ENHRI to develop an innovative approach to neurological and behavioral health measurement that would be responsive to the needs of researchers in a variety of settings, with a particular emphasis on measuring outcomes in large cohort such as epidemiological, large longitudinal, and prevention or intervention trials. (cont'd page 4)

Criteria Meeting

The recent Criteria meeting held in January brought together 27 experts in the field of neuroscience from across the country for the first time on this project. The purpose of the meeting was to set criteria for selection of sub-domains and adoption of tools to use over the next four years. Now each of the four domain teams is busy gathering literature and reviewing our request for information from outside experts, to help inform the decision. Experts are also being asked to establish priorities from the 8-9 sub-domains in each of the four broad areas. (see Table 1).

During the study the following criteria were agreed upon as a result of a modified Delphi consensus gathering technique:

Validity	43.7%
Precision/Accuracy across a broad range of function	19.3%
Innovative methods - both assessment and evaluative/psychometric	15.7%
Suitable for Use	14.7%
Usability	6.6%

Validity

This was one of the most important elements of the discussion in that it covers a wide array of criteria. Some of the criteria that the panel decided to group with validity include:

- predict gold standard criteria for current and later function:
- strong conceptual/theoretical foundation
- cross-walk to other scales
- provide results that predict real-world functioning

Precision/Accuracy across a broad range of function

This category addresses the importance of the Toolbox measures to be sensitive across a full range of ability or function and across the normal population:

- help define the normal range (average) and cross-over
- cover full range of construct

Innovative Methods

Discussions were based around the importance of the Toolbox measures to incorporate innovative methods, such as the use of Item Response Theory (IRT), Computer Adaptive Testing (CAT), and Item Banking:

- brief & practical
- adaptability
- incorporate item banking (i.e. be adaptable over time)
- computer administration/scoring
- utilize IRT & CAT

Suitability for Use

- with variety of racial/ethnic groups and education levels / usable with broad range of population (ethnic, regional, literacy)
- norms for different age groups
- usable across literacy levels

Usability

- provide results that are easy to understand
- ease of administration and scoring
- select instruments with minimal/no cost intellectual property hurdles

Table 1. Candidate Subdomains for Potential Toolbox Measurements

Emotional Subdomains

1. Negative Affect
2. Positive Affect
3. Externalizing Problems
4. Coping /Resilience
5. Sense of Control / Self-Efficacy
6. Emotion Regulation
7. Social Integration
8. Attachment/Intimate Relationships

Motor Subdomains

1. Locomotion
2. Strength
3. Coordination
4. Balance
5. Endurance
6. Dexterity
7. Upper Extremity Function
8. Flexibility

Sensory Subdomains

1. Vision
2. Hearing
3. Vestibular Balance
4. Touch
5. Olfaction
6. Taste
7. Pain
8. Proprioception

Cognitive Subdomains

1. Executive Function
2. Processing Speed
3. Learning & Memory
4. Language
5. Attention & Working Memory
6. Visuospatial Functioning

Request for Information (RFI) Summary

A key initial task was to identify criteria for the acceptance of behavioral measures for widespread use in clinical and epidemiological research. The data obtained will inform subsequent stages of the project to ensure that the final set of instruments will meet with widespread acceptance by likely users.

Toolbox Characteristics

Overall, the majority of respondents thought the maximum amount of administration time an individual domain module (e.g., all cognition instruments in the Toolbox) should be between 20-30 minutes, while the maximum amount of time an assessment battery (including all four domains of health and function: Cognitive, Sensory, Motor and Emotional) should be no longer than two hours. Using a 4-point Likert scale, participants were asked to rate how important certain characteristics were for the Toolbox measures to possess, from “not important at all” to “very important.” Of the characteristics rated the majority of the time as “very important”, the most highly rated responses were the ability of the Toolbox measures to: a) provide consistent results even when administered by different people (95%, n=132), b) be responsive to real change (93%, n=128), c) measure what is supposed to be measured (89%, n=124), and d) be stable over time unless there is a true change in what is being measured (83%, n=116). Please see Table 2 for the percentages of criteria sorted by “very important”.

TABLE 2: Potential Toolbox Criteria

Potential Criteria	Very Important	Somewhat Important	A little Important	Not at all Important	N
provide consistent results even when administered by different people	95% (132)	5% (7)	0% (0)	0% (0)	139
be responsive to real change (sensitivity)	93% (128)	7% (10)	0% (0)	0% (0)	138
measure what it is supposed to measure (general vitality)	89% (124)	9% (13)	1% (2)	0% (0)	139
be stable over time unless there is a true change in what is being measured (test-retest reliability)	83% (116)	14% (20)	1% (2)	1% (0)	139
be suitable for use with a variety of racial and ethnic groups	69% (96)	27% (37)	4% (6)	0% (0)	139
provide results that predict real-world functioning	64% (89)	31% (43)	5% (7)	0% (0)	139
have established norms for different age groups	58% (81)	27% (38)	14% (19)	1% (1)	139
provide scores/results that are easy to understand	52% (72)	36% (50)	9% (13)	2% (3)	138
predict a gold standard criterion at the time of assessment (concurrent validity)	50% (69)	42% (59)	8% (11)	0% (0)	139
provide consistent results across items within a single test (internal consistency)	49% (67)	38% (53)	12% (16)	1% (2)	138
have a Spanish translation	45% (62)	42% (58)	13% (18)	1% (1)	139
be easy to administer	45% (61)	39% (54)	15% (20)	1% (2)	137
correlate strongly with outside measures of the same construct (convergent validity)	42% (59)	49% (68)	9% (12)	0% (0)	139
provide profile of scores for different areas	40% (55)	41% (57)	17% (23)	3% (4)	139
predict a gold standard criterion at some later point in time (predictive validity)	39% (54)	48% (66)	12% (17)	0% (0)	137
be suitable for all age groups (or have versions that can be given to all age groups)	37% (52)	39% (54)	17% (24)	6% (9)	139
make sense to the respondent (face validity)	34% (47)	46% (63)	17% (24)	3% (4)	138



Steering Committee Members

Richard Gershon, PhD (PI)
ENH Research Institute

David Cella, PhD
ENH Research Institute

Susan Coldwell, PhD
University of Washington

Sureyya Dikmen, PhD
University of Washington

Winnie Dunn, PhD
University of Kansas

Richard Havlik, MD, MPH
Westat, Inc.

Hugh Hendrie, MB, ChB, DSc
Indiana Univ./Purdue Univ.

Nina Kraus, PhD
Northwestern University

Dan Mungas, PhD
University of California-LA

Cindy Nowinski, MD, PhD
ENH Research Institute

Paul Pilkonis, PhD
University of Pittsburgh

David Reuben, MD
University of California-LA

William Rymer, MD, PhD
Rehabilitation Institute of Chicago

Terry Svaglic, MAcc
ENH Research Institute

David Tulskey, PhD
Kessler Corporation (KMRREC)

Molly Wagster, PhD
National Institute on Aging

Robert Wallace, MD
University of Iowa

Margaret Wallhagen, PhD
University of California-SF

Sandra Weintraub, PhD
Northwestern University



Newsletter Editor

Maggie Heydt

Director of
Communications

ENH Research Institute

847.570.7304
mheydt@enh.org

Why This, Why Now?

(cont'd from page 1)

The overall goal of the contract is to develop unified/integrated methods and measures of cognitive, emotional, motor and sensory health for use in the types of studies just mentioned. Such measures are rarely included in studies of this type, due in no small part to the lack of instruments that are well-validated and have a short administration time; further, where such domains are sampled, the heterogeneity of measures hinders the ability to compare data readily across studies. With an available toolbox of measures, yields from these large and very expensive studies can be maximized by allowing a much larger number of important research questions regarding neurological and behavioral health to be studied – both with respect to the primary study aims, and to those arising from secondary data analyses – with a minimal increment in subject burden and cost. By ensuring that the assessment methods are capable of comparison to existing and completed studies and can incorporate modifications/improvements in the future, a truly “economic” and valuable resource for the entire neuroscience community will result.

Advances in psychometric/educational testing research methodology, including computerized adaptive testing and real-world setting (virtual reality), may lead to more efficient, flexible, and responsive assessment of cognitive, emotional, sensory and motor function in a variety of settings. Thus, a secondary goal of the contract is to integrate these methodological advances with statistical techniques (e.g., item response theory) so that the resultant measurement tool has the capability of being adaptive over time in response to advances in science and/or advances in technology.



NIH Toolbox
1001 University Place
Evanston, IL 60201
www.nihtoolbox.org