Emotion assessment using the NIH Toolbox

John M. Salsman, PhD
Zeeshan Butt, PhD
Paul A. Pilkonis, PhD
Jill M. Cyranowski, PhD
Nicholas Zill, PhD
Hugh C. Hendrie, MB, ChB, DSc
Mary Jo Kupst, PhD
Morgen A. R. Kelly, PhD
Rita K. Bode, PhD
Seung W. Choi, PhD
Jin-Shei Lai PhD, OTR/L
James W. Griffith, PhD
Catherine M. Stoney, PhD
Pim Brouwers, PhD
Sarah S. Knox, PhD
David Cella, PhD

Correspondence to Dr. Salsman: j-salsman@northwestern.edu

ABSTRACT

One of the goals of the NIH Toolbox for Assessment of Neurological and Behavioral Function was to identify or develop brief measures of emotion for use in prospective epidemiologic and clinical research. Emotional health has significant links to physical health and exerts a powerful effect on perceptions of life quality. Based on an extensive literature review and expert input, the Emotion team identified 4 central subdomains: Negative Affect, Psychological Well-Being, Stress and Self-Efficacy, and Social Relationships. A subsequent psychometric review identified several existing self-report and proxy measures of these subdomains with measurement characteristics that met the NIH Toolbox criteria. In cases where adequate measures did not exist, robust item banks were developed to assess concepts of interest. A population-weighted sample was recruited by an online survey panel to provide initial item calibration and measure validation data. Participants aged 8 to 85 years completed self-report measures whereas parents/guardians responded for children aged 3 to 12 years. Data were analyzed using a combination of classic test theory and item response theory methods, yielding efficient measures of emotional health concepts. An overview of the development of the NIH Toolbox Emotion battery is presented along with preliminary results. Norming activities led to further refinement of the battery, thus enhancing the robustness of emotional health measurement for researchers using the NIH Toolbox. ***Neurology*® 2013;80 (Suppl 3):S76–S86**

GLOSSARY

IRT = item response theory; **NA** = negative affect; **PROMIS** = Patient Reported Outcomes Measurement Information System; **RFI** = Request for Information.

In everyday terms, the word "emotion" evokes connotations of strong feelings, often negative and distressing when they tax our capacity to maintain our poise. It is unpleasant when we are extended beyond our usual resources by stressful life events including poverty, unemployment, oppression, discrimination, and disease. However, positive emotions can be reflections of well-being in our lives, and positive social relationships can buffer stress and enhance health. Recognizing the full spectrum of emotional life and its impact on health, the mandate for the NIH Toolbox was to develop assessments with a broad focus, beyond just negative emotion, or emotional distress. The original Request for Proposals alluded to several additional aspects of the experience and expression of emotion relevant to health in the general population including the importance of psychological well-being, the role of important aspects of positive functioning, such as adaptability, resilience, and self-efficacy, and the importance of the interpersonal and social context in which emotions arise and may be expressed.

Feedback provided through an NIH Toolbox Expert Request for Information (RFI) was also consistent with this desire for a broad emphasis on emotional health. As part of this RFI, we obtained input from 147 experts (64% of 232 invited experts), including key opinion leaders from the NIH research program staff and NIH-funded investigators with a broad focus in neurologic and behavioral

Supplemental data at www.neurology.org

From the Departments of Medical Social Sciences (J.M.S., Z.B., S.W.C., J.-S.L., J.W.G., D.C.) and Physical Medicine & Rehabilitation (R.K.B.), Northwestern University Feinberg School of Medicine, Chicago; Robert H. Lurie Comprehensive Cancer Center of Northwestern University (J.M.S., J.-S.L., D.C.), Chicago; Comprehensive Transplant Center (Z.B.), Northwestern University, Chicago; Center for Patient-Centered Outcomes (J.M.S., Z.B., J.-S.L., J.W.G., D.C.), Northwestern University, Chicago, IL; Department of Psychiatry (P.A.P., J.M.C., M.A.R.K.), University of Pittsburgh Medical Center, Pittsburgh, PA; Westat, Inc. (N.Z.), Rockville, MD; Center for Aging Research (H.C.H.), Indiana University School of Medicine, Indianapolis, IN; Department of Pediatrics (M.J.K.), Medical College of Wisconsin, Milwaukee; Family Services of Western Pennsylvania, and VA Pittsburgh Healthcare System (M.A.R.K.), Pittsburgh, PA; National Heart, Lung, and Blood Institute (C.M.S.), National Institutes of Health, Bethesda; Division of AIDS Research (P.B.), National Institute of Mental Health, Bethesda; and Department of Community Medicine (S.S.K.), West Virginia University School of Medicine, Morgantown.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

© 2013 American Academy of Neurology

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

research. Fifty-eight percent of those experts who responded indicated sufficient familiarity with emotional health measurement to provide more targeted feedback. Each expert nominated what he or she considered to be the top 4 components of emotional health that would have most relevance in a battery to be used with general or at-risk populations for large-scale longitudinal and epidemiologic studies and clinical trials. The 4 areas most frequently endorsed were negative affect (NA) (98%; n = 83), adaptability (74%; n = 63), positive affect (73%; n = 62), and self/emotional regulation (68%; n = 58).

Important considerations for the work of the Emotion team involved integrating the input of the experts with the goals of the NIH Toolbox and the published literature on emotional health. Our primary objective was to develop tools that reflected health status and function over time, and that would be responsive to change, including the impact of interventions being tested in clinical research with community-dwelling or at-risk individuals. Given the internal nature of emotional experience, self-reports are the most preferred and practical mode of assessment for adults. For younger children, however, measures of emotion often must rely on behavioral observations by knowledgeable informants (e.g., parents, teachers). As a result, we focused on self-report measures for adults and older children and on proxy measures for younger children in the proposed work.

Regarding the particular subdomains assessed within emotional health, there is good evidence that negative and positive emotions are best understood as independent of each other, rather than opposite extremes of a single continuum.[1] This perspective led us to separate development of assessments for negative and positive mood (and of emotional distress vs positive psychological functioning more generally). Emotion has important adaptive and communicative functions in humans; it serves "signal" functions for both self and others. Thus, we also considered emotion in an interpersonal context and assessed its reciprocal relationships with interpersonal functioning and behavior.

Guided by a review of the literature (discussed in more detail below), the RFI survey of experts in the area of emotional health, follow-up semi-structured interviews with a subset of these experts, and discussion within the Emotion domain team and among the Emotion expert consultants, we identified 4 subdomains central to the assessment of emotional health: Negative Affect, Psychological Well-Being, Stress and Self-Efficacy, and Social Relationships. Our next task was to identify psychometrically sound and free-for-use instruments. Where these did not exist, we considered development of new measures including item banks wherever possible, which are large pools of questions from which multiple equivalent measurement tools can be derived. A subsequent review of available instruments identified more than 600 existing measures covering relevant concepts. Self-report instruments were targeted for children and adults (ages 8 years and older) and proxy instruments were targeted for children (ages 3–12 years). Instruments that were lengthy or designed for diagnostic purposes were eliminated from consideration. Instruments were reviewed for intellectual property considerations, and confirmation of free-for-use status was sought from authors and/or publishers. Final psychometric review resulted in selection of measures and development of item banks representing the 4 emotional health subdomains to be tested in the calibration and validation stage of the NIH Toolbox.

DATA COLLECTION AND ANALYTIC APPROACH

Our data collection and analytic approach relied heavily on item response theory (IRT), an alternate approach to traditional test construction.[2–5] IRT-based approaches model the probability that a person, at a particular latent trait level, will endorse an item in a particular way. As opposed to the typically used sum of item scores, IRT modeling provides an efficient means to reduce error in the estimate of a particular symptom or domain of emotional health based on an individual's overall pattern of responses to test items. In turn, this information can be used to assess the quality of individual items and to calibrate test scoring. An item bank is composed of carefully calibrated questions that define and quantify a common theme and thus provide an operational definition of a symptom or other aspect of emotion. Because IRT methods make it possible to estimate components of emotional health using any set of items in an item bank, the approach also lends itself to the development of robust short forms and the application of a computerized adaptive test, in which an item from the bank is administered based on the expected yield from that question. Consequently, application of IRT may allow for briefer

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

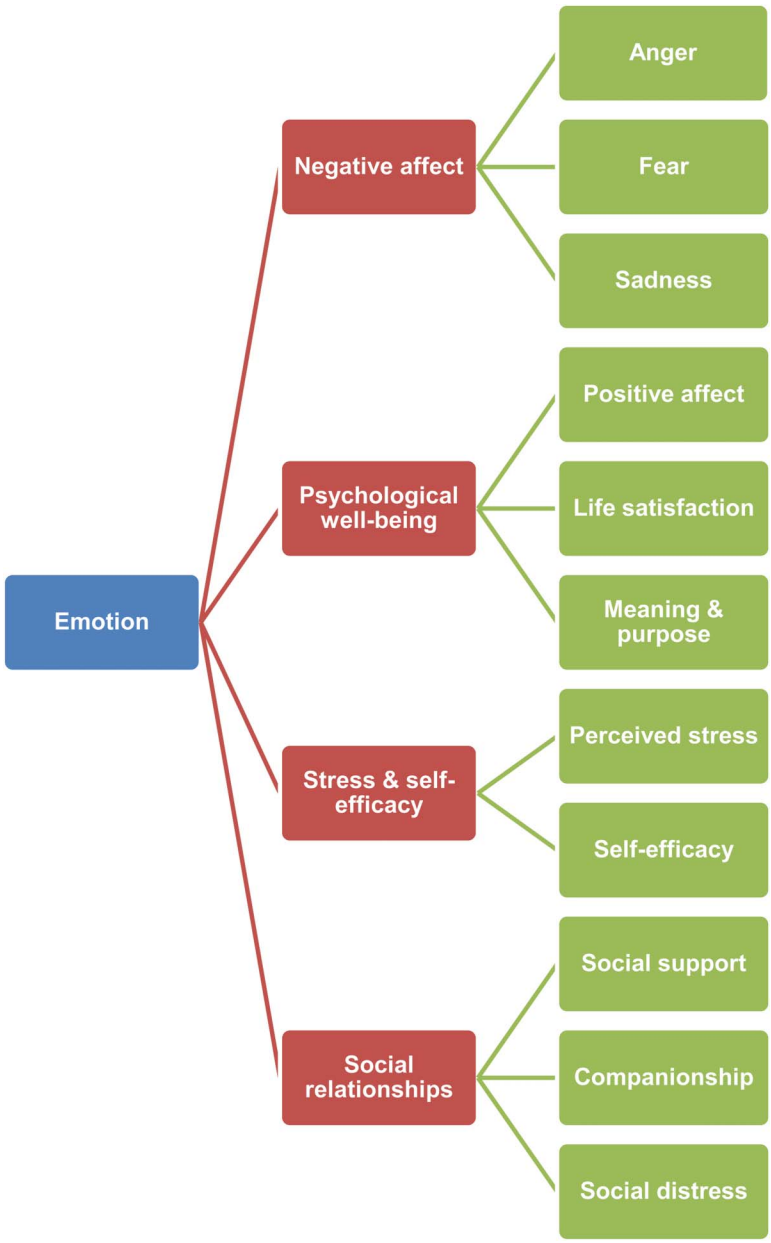
and more efficient assessments and assessment of more symptoms and domains of interest than have been typically feasible, all of which are central goals of the NIH Toolbox.

A detailed data sampling plan was developed for collecting initial item responses to the candidate items and measures from the targeted Emotion subdomains (see figure).

This sampling plan was designed to accommodate a number of purposes: 1) create item calibrations for all of the items in each of the subdomains; 2) when possible, provide convergent validity estimates with legacy questionnaires (e.g., the Center for Epidemiologic Studies–Depression Scale); 3) confirm the factor structure of the subdomains and content areas; and 4) conduct item and bank analyses. However, because of the large total number of items (>600), it was not possible for participants to respond to the entire set of items in each pool. Based on an estimate of 6 questions per minute for adults (5 questions per minute for children), the length of the Emotion questionnaires in the calibration and validation wave of testing was limited to approximately 180 items for adults (150 items for children), which were expected to take approximately 30 minutes to answer. A block administration was implemented for data collection, and 5 community-dwelling Internet panel samples were recruited (see table 1) by an online research panel.

Subsequent data analyses included evaluation of item and scale properties using both traditional (classic test theory) and modern (IRT) psychometric methods. Factor analysis was used to examine the underlying structures of the measured constructs and to evaluate the unidimensionality and local independence assumptions of the IRT model. Differential item functioning was examined to evaluate whether items performed differently across key demographic groups when statistically controlling for the underlying level of the trait assessed by the scale. Finally, items were calibrated using an IRT model to facilitate development of short forms and a computerized adaptive test. The subscales we developed were based on the psychometric analyses and consistent with the NIH Toolbox purposes. Acceptable psychometric properties were defined a priori as a comparative fit index >0.90, a root mean square error of approximation <0.10, and Cronbach α >0.70. In some cases, we recommend modified versions of existing measures (please contact the corresponding author for details and view appendix e-1 on the *Neurology*® Web site at www.neurology.org for measure references).

Figure Emotion Domain framework



Emotion Domain and subdomains.

SUBDOMAIN-SPECIFIC DISCUSSION AND PRELIMINARY RESULTS

Negative affect. Negative affect is a phrase used to describe unpleasant feelings or emotions, which exist on a continuum ranging from common feelings to more extreme indicators of these same affects (appendix e-2). Negative emotions have been associated with both the incidence and progression of some diseases, such as cardiovascular disease and certain autoimmune diseases. The mechanisms underlying these associations are not well understood, but they can be organized into 2 general families: direct effects via physiologic pathways and indirect effects via the impact of NA on health-related behaviors. Accurate measurement of NA may lead to improvements in our understanding of the impact of emotions on general health of the individual as well as mechanisms of certain disease processes.

Our focus in the NA subdomain is on 3 principal negative emotions: sadness, fear, and anger. Given the overlap among the cognitive, affective, and somatic

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Table 1 Demographic characteristics for the NIH Toolbox calibration and validation samples

	Pediatric proxy 3-7 (n = 1,000)	Pediatric proxy 8-12 (n = 1,086)	Pediatric 8-12 (n = 1,553)	Pediatric 13-17 (n = 1,525)	Adults 18+ (n = 2,551)[a]
Mean age, y	5.05	10.10	10.09	14.99	45.56
Sex, %					
Male	51.0	49.2	50.5	50.3	43.3
Ethnicity, %					
Not Hispanic/Latino	87.8	88.8	88.5	92.5	90.0
Race					
American Indian/Alaska native	1.8	2.0	1.8	1.8	2.7
Asian	3.4	1.9	2.4	2.0	3.6
Black/African American	12.0	12.9	11.7	10.0	10.3
Native Hawaiian/other Pacific Islander	1.3	0.6	0.5	0.4	0.6
White	80.4	80.7	82.0	84.7	81.6
Other	6.8	5.2	5.3	3.5	3.9
Proxy relationship, %					
Mother/female guardian	75.0	71.6			
Father/male guardian	17.4	22.5			
Grandmother	4.8	4.3			

[a] Note: 41% of participants in the adult sample had a high school education or less.

symptoms of sadness, fear, and anger, a number of conceptual models have been proposed to account for the shared vs unique variance captured in measures of NA. In the Watson and Clark model,[6] a second-order, nonspecific factor reflecting high levels of negative affect (or "general distress") was common to these 3 emotions. Angry mood tends to have smaller loadings on the general factor than fear and sadness, but it still is a strong marker of the dimension. In addition, the Watson and Clark model included first-order factors that are specific to, and help to differentiate, the 3 affects. Symptoms specific to sadness are those that reflect low levels of positive affect. Symptoms that best differentiate fear are those that reflect autonomic arousal and perceptions of threat. The affective expression of anger includes symptoms that reflect irritation, frustration, and resentment.[7,8] NA can be understood as a set of related but distinguishable concepts.

Sadness, fear, and anger are experienced throughout the lifespan, and, thus, they are amenable to assessment by self-report. We have selected age-appropriate measures for respondents aged 8 years and older. Given the lifespan emphasis of the NIH Toolbox, special assessment issues arise with young children. Measures of NA in children often must rely on behavioral observation reports by parents and teachers, and we propose to include such proxy measures for younger children. Factor analytic studies of such ratings are generally consistent with the Watson and Clark hierarchical model of NA described above. At the same time, the analyses also capture separable subdimensions, specifically, an "internalizing" factor with item loadings reflecting sadness and fear and an "externalizing" factor reflecting anger as well as stubbornness, sullenness, impulsivity, and defiance.

Our initial psychometric (calibration) data have been used to identify brief scales that can assess these negative emotions efficiently. In some cases, our explorations of the dimensionality of the item banks and legacy scales used in the initial psychometric testing led us to articulate specific facets of these negative emotions. In adults, for example, somatic anxiety was distinguished from generalized anxiety, which has primarily cognitive and affective components. The larger goal, however, has remained consistent throughout, that is, to identify brief measures of sadness, fear, and anger that are age-appropriate and can be used to measure these constructs throughout the lifespan. Table 2 provides summary data on the scales recommended for inclusion in the NIH Toolbox.

Psychological Well-Being. The Psychological Well-Being subdomain includes positive affect (happiness, serenity, and cognitive engagement), life satisfaction, and meaning and purpose in life. Psychological well-being has been conceptualized as having hedonic and eudaimonic components.[9] Hedonic aspects of well-being are often more experiential in nature and emphasize pleasure (e.g., positive affect), whereas eudaimonic aspects of well-being are more evaluative in nature and emphasize human flourishing (e.g., meaning, life satisfaction). Positive affect has been characterized as "feelings that reflect a level of pleasurable engagement with the environment such as happiness, joy, excitement, enthusiasm, and contentment"

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Table 2 Confirmatory factor analysis by Negative Affect instrument—single-factor model

CFA model	No. of items	CFI	RMSEA	Cronbach α	Convergent validity
Adult 18+ self-report					
NIH Toolbox Anger—Physical Aggression	5	0.981	0.097	0.83	
NIH Toolbox Anger—Hostility	5	0.975	0.160	0.85	
NIH Toolbox Anger—Affect SF	8	0.988	0.085	0.94	0.61[a]
NIH Toolbox Fear—Affect SF	7	0.995	0.086	0.95	0.86[b]
NIH Toolbox Fear—Somatic Arousal	6	0.958	0.124	0.85	
NIH Toolbox Sadness SF	8	0.985	0.171	0.97	0.88[c]
8–17 Self-report					
NIH Toolbox Pediatric Anger SF	6	0.993	0.082	0.92	0.73[d]
NIH Toolbox Pediatric Fear SF	8	0.979	0.102	0.93	0.68[e]
NIH Toolbox Pediatric Sadness SF	8	0.974	0.180	0.95	0.75[f]
8–12 Proxy					
NIH Toolbox Anger	10	0.975	0.101	0.92	0.54[g]
NIH Toolbox Fear	10	0.975	0.063	0.88	0.57[g]
NIH Toolbox Sadness	11	0.986	0.051	0.89	0.67[g]
3–7 Proxy					
NIH Toolbox Anger	9	0.983	0.053	0.85	0.64[h]
NIH Toolbox Fear—Overanxious	6	0.969	0.090	0.79	0.60[i]
NIH Toolbox Fear—Separation Anxiety	7	0.982	0.057	0.83	0.71[j]
NIH Toolbox Sadness	7	0.992	0.031	0.77	0.38[k]

Abbreviations: CFA = confirmatory factor analysis; CFI = comparative fit index; RMSEA = root mean square error of approximation; SF = short form.
Validation scales:
[a] Buss-Perry Aggression Questionnaire, combined subscales for anger and verbal aggression.
[b] Generalized Anxiety Disorder-7.
[c] Center for Epidemiologic Studies Depression Scale.
[d] Anger Expression Scale for Children.
[e] Screen for Anxiety Related Disorders.
[f] Short Mood and Feelings Questionnaire.
[g] Positive and Negative Affect Schedule for Children-negative subscale.
[h] Children's Behavior Questionnaire-anger/frustration subscales.
[i] Preschool Anxiety Scale-general anxiety subscale.
[j] Preschool Anxiety Scale-separation anxiety subscale.
[k] Children's Behavior Questionnaire-sadness subscale.

and describe both activated, excited and nonactivated, peaceful components of positive affect.[10] Diener et al.[11] note that well-being may include correlated but separable components of high positive affect, life satisfaction, and low NA. A closely related construct to life satisfaction is life purpose or meaning, often considered a key aspect of "living well."[12] Together, these concepts are considered critical components of psychological well-being throughout the lifespan.

The relationship between psychological well-being and health is well documented, with a majority of this research focusing on the role of positive affect.[13–20] However, the mechanisms by which psychological well-being affects health remain unclear. Pressman and Cohen[14] have suggested that positive affect has both direct and stress-buffering effects on health outcomes. In the direct effect model, physiologic states associated with positive affect (e.g., sympathetic/parasympathetic and brain activation patterns) are directly associated with health outcomes, and in the stress-buffering model, positive affect moderates the relationships between stress and poor health outcomes. Fredrickson[21] suggests that positive affect is also beneficial because it serves a "broaden-and-build" function in individuals, enhancing one's repertoire of social cognitive and problem-solving capabilities, and thereby strengthening personal and interpersonal resources.

Parent proxy reports were used for the youngest children (aged 3–12 years) and self-report for children aged 8 to 17 years and adults aged 18 to 85 years. Our initial psychometric (calibration) data were used to shorten scales to more efficiently assess psychological well-being across the lifespan. The expression of positive emotions demonstrated a developmental sequence with greater

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

variability in range and intensity in adulthood. As a consequence, less-activated components of positive affect (serenity and cognitive engagement) were more readily differentiated and measured in adults relative to children whereas the more activated expressions of positive affect (i.e., happiness) were easily detectable throughout the lifespan.

Similarly, although aspects of global life satisfaction were easily assessed throughout the lifespan, the life purpose/meaning component was expressed differently at different ages. For children, life purpose and meaning was expressed through satisfaction with domain-specific activities (e.g., I am satisfied with the friends I have, I am satisfied with my skills and talents), whereas for adults it took on a more abstract and philosophical connotation closely linked to life satisfaction and associated with human flourishing and "living well" (e.g., My life has a clear sense of purpose, I value my activities a lot). We anticipate that meaning and purpose expressed by children through domain-specific satisfaction will be associated with the more abstract perspectives of adulthood, thus enabling a lifespan assessment of meaning and purpose. Table 3 provides summary data on the scales recommended for inclusion in the NIH Toolbox.

Stress and Self-Efficacy. Instruments in this subdomain measure perception of the nature of events, cognitions, and experiences and one's ability to respond to these challenges. The specific concepts assessed include perceived stress and self-efficacy. In the NIH Toolbox assessment, we are concerned with measuring how individuals perceive and report these variables in order to explore relationships with health status and function.

Table 3 Confirmatory factor analysis by Psychological Well-Being instrument—single-factor model (unless otherwise noted)

CFA model	No. of items	CFI	RMSEA	Cronbach α	Convergent validity
Adult 18+ self-report					
NIH Toolbox Positive Affect	21	0.932	0.106	0.96	0.92[a]
Satisfaction with Life Scale	5	0.943*	0.100*	0.91	
Students' Life Satisfaction Scale	7			0.91	
NIH Toolbox Meaning & Purpose SF	8			0.89	0.64[b]
13–17 Self-report					
NIH Toolbox Pediatric Positive Affect	16	0.926	0.112	0.97	0.96[c]
Students' Life Satisfaction Scale	7	0.980*	0.084*	0.91	
Domain-Specific Life Satisfaction	10			0.89	0.50[d]
8–12 Self-report					
NIH Toolbox Pediatric Positive Affect	15	0.913*	0.133*	0.95	0.97[c]
Students' Life Satisfaction Scale	7	0.944*	0.104*	0.91	
Domain-Specific Life Satisfaction	9			0.88	0.52[d]
8–12 Proxy					
NIH Toolbox Positive Affect	12	0.951*	0.090*	0.93	0.95[c]
Students' Life Satisfaction Scale	7	0.955*	0.082*	0.86	
Domain-Specific Life Satisfaction	10			0.88	0.28[d]
3–7 Proxy					
NIH Toolbox Positive Affect	9	0.973*	0.073*	0.92	0.95[c]
Students' Life Satisfaction Scale	7	0.981*	0.058*	0.86	
Domain-Specific Life Satisfaction	8			0.87	0.28[d]

Abbreviations: CFA = confirmatory factor analysis; CFI = comparative fit index; RMSEA = root mean square error of approximation; SF = short form.
Validation scales:
[a] Positive and Negative Affect Schedule—positive subscale.
[b] Satisfaction with Life Scale.
[c] Positive and Negative Affect Schedule for Children—positive subscale.
[d] Students' Life Satisfaction Scale.
*Results from a bifactor analysis in which these scales together load on a common factor but have separate factor loadings of their own. CFA statistics reflect the overall fit of this model. In cases where items were sampled from larger item banks, CFI and RMSEA statistics are reported on the larger item bank. The CFA statistics reported for life satisfaction items represent fit statistics for models that include life satisfaction and meaning & purpose item content (Ages 18+) or general and domain-specific life satisfaction item content (Ages 3–17).

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Psychological definitions of stress focus on perceptions about the nature of events and their relationship to the values and perceived coping resources of an individual.[22] We chose perceived stress rather than an objective list of stressors because we are more interested in people's experience and perception of situations. In general, psychological stress is said to occur when an individual perceives that environmental or internal demands that are personally meaningful exceed his or her perceived adaptive capacity.[23] This definition was chosen because it is the perception of distress that triggers the cascade of neuroendocrine changes that can negatively affect health outcomes. This definition of stress integrates the nature of the threat and the perceived coping capacities of the individual.

Self-efficacy is a person's belief in his/her capacity to manage his/her life and have control over meaningful events.[24] Self-efficacy varies across situations and skills needed to accomplish necessary tasks. Perceived self-efficacy can serve as a mediator between stress and health outcomes.

There is much literature in behavioral health research on the relationships among perceived stress, self-efficacy, and health. Mechanisms of these associations with health have typically supported mediation effects. For example, models of stress and immunity suggest that dysregulated hormone secretion or maladaptive health practices may operate as mediational pathways and result in suppressed immune functioning.[25,26] Self-efficacy has been shown to mediate the link between therapeutic interventions and health behaviors, including pain experience and management, control of eating and weight, success of recovery from myocardial infarction, and adherence to preventive health programs.[27]

The essential cognitive and emotional abilities to perceive situations, to reflect on one's efficacy, and to manage situations and emotional reactions are dependent on developmental processes. We have selected age-appropriate, self-report items and measures for children aged 8 years and older and supplemented these self-reports with proxy reports from parents or other caregivers for children aged 8 to 12 years. Thus, although assessment of these constructs may differ with respect to item content, our goal was to measure the same constructs equivalently across the lifespan. We were able to accomplish this in most cases. Given the limited development of cognitive and emotional abilities for children aged 3 to 7 years and the inherent difficulty for parents or other caregivers to speculate on their children's perceptions of stress and self-efficacy, we do not assess these concepts for this age group. Table 4 shows the summary data on the scales recommended for inclusion in the NIH Toolbox for Stress and Self-Efficacy.

Table 4 Confirmatory factor analysis by Stress and Self-Efficacy instrument—single-factor model (unless otherwise noted)

CFA model	No. of items	CFI	RMSEA	Cronbach α
Adult 18+ self report				
Perceived Stress Scale-14	14	0.976[a]	0.091[a]	0.91
General Self-Efficacy Scale	10	0.993[b]	0.073[b]	0.93
13–17 Self-report				
Perceived Stress Scale-14	14	0.988[a]	0.057[a]	0.89
General Self-Efficacy Scale	10	0.994[b]	0.060[b]	0.90
How I Feel	10	0.948	0.182	0.93
8–12 Self-report				
Perceived Stress Scale-14	14	0.975[a]	0.072[a]	0.87
General Self-Efficacy Scale	10	0.991[b]	0.076[b]	0.90
How I Feel	10	0.939	0.176	0.92
8–12 Proxy				
Perceived Stress Scale-14	14	0.986[a]	0.056[a]	0.87
General Self-Efficacy Scale	10	0.995[b]	0.060[b]	0.92
How I Feel	10	0.948	0.203	0.94

Abbreviations: CFA = confirmatory factor analysis; CFI = comparative fit index; RMSEA = root mean square error of approximation.
[a] Results from a bifactor analysis with group factors for positively and negatively scored items.
[b] Results from a 1-factor model with 1 residual dependency between 2 items: "I can always manage to solve difficult problems if I try hard enough," and "I can solve most problems if I invest the necessary effort."

Social Relationships. There are several dimensions of social relationships, including their structure, extent, and quality.[28,29] For the purpose of the NIH Toolbox, we have focused on assessing concepts related to the perceived availability and quality of one's social relationships, including perceptions of social support, loneliness, and social distress. Self-report scales have been developed when developmentally appropriate (i.e., for respondents aged 8 years and older). In contrast, parental proxy scales have been developed to measure analogous aspects of these social relationship constructs in younger age cohorts (i.e., for children aged 3–12 years), and to assess early behavioral indicators of positive social development (i.e., empathic behaviors).

Social relationships can have a significant impact on both emotional and physical health outcomes across the lifespan.[30–33] Whereas positive and supportive social relationships may improve health and well-being, perceptions of social isolation, loneliness, or relationship distress can be associated with negative health outcomes. Mechanisms by which social relationships affect health may include both the direct and indirect impact of social relationships on emotional and physiologic responses (e.g., oxytocin release), the ability to cope with changing life stressors, or the patterning of sleep and health behaviors.[28,29]

Social support refers to the extent to which an individual views social networks as available to provide aid in times of need, and may serve to protect

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

individuals from the negative emotional and physical effects of life stress.[29,33] In our adult self-report forms, perceived social support is assessed with 2 related but distinguishable scales, emotional support, or the perceived availability of someone to provide empathy or advice in times of need, and instrumental support, or the perceived availability of someone to provide aid with basic daily activities when needed, such as shopping, cleaning, or preparing meals. Given the particular importance of perceived parental support for pediatric samples, self-report assessment of perceived social support in samples for ages 8 to 17 years includes brief scales assessing perceived closeness within parental relationships, as well as more general perceptions of emotional support experienced across relationships (paralleling the adult emotional support scale). Similarly, assessment of social support by parental proxy in children aged 3 to 12 years focuses on the assessment of parent-child relationships.

Companionship refers to the perceived availability and closeness of friends or companions with whom to interact, vs feelings of loneliness or social isolation. Research suggests that self-reported loneliness predicts a variety of negative health outcomes, including depression, sleep disturbance, cardiovascular risk indicators, and increased mortality.[30–32] Our initial psychometric testing identified 2 related but distinguishable scales related to the companionship construct: loneliness and friendship. Analogous constructs are also included in parental proxy forms for assessment of children aged 3 to 12 years, in which parents rate the target child's engagement in social withdrawal from peers and positive peer interactions.

Social distress refers to the extent to which individuals perceive their daily social relationships or interactions to be negative. Negative aspects of social relationships, such as perceptions of hostility, criticism, or rejection experienced in one's daily social interactions, may have powerful, deleterious effects on emotional and physical health across the lifespan. For the purpose of the NIH Toolbox, social distress items were designed to capture the perceived frequency of hostile, critical, insensitive, or rejecting social exchanges. In our adult (age 18 years and older) self-report forms, this construct is assessed with 2 related but distinguishable scales: perceived hostility and perceived rejection. Similar self-report scales have been developed for use in pediatric (ages 8–17 years) samples. Analogous assessment of social distress experienced by 3- to 12-year-old children is included in parental proxy reports of the target child's experience of rejection by peers and rejection by siblings. Finally, given its importance both as an indicator of a child's current emotional health and positive social development, and as a predictor of positive and supportive social relationships in adolescence and adulthood, parental proxy assessment includes the parent's evaluation of the target child's empathic behaviors. Table 5 provides a summary of the psychometric data for the scales recommended for the Social Relationships subdomain of Emotion.

LIMITATIONS The convergent validity coefficients, reported in the tables, are initial attempts to examine relationships with other existing measures that assess similar concepts. Some of these relationships are modest (coefficient range = 0.10–0.30). In some cases, the conceptual link between the NIH Toolbox measure and the comparison ("validating") measure is weak. In other cases, the psychometric properties (e.g., internal consistency reliability) of the identified comparison measures were not strong, which will attenuate correlation coefficients. In addition, some of these data are from different informants, which also tends to decrease the magnitude of association. These mitigating considerations can account for some of the relatively lower correlations between the NIH Toolbox measures and the selected comparison measures. Another limitation is the use of informants in general for some concepts in children. Although this was deemed necessary in order to obtain reasonable measures among young children, there is potential for the informant ratings to be inaccurate reflections of the child's emotional health. This can be especially problematic when the informant may be a cause of NA or stress in the child. A more detailed and comprehensive treatment of the psychometric issues will be addressed in forthcoming publications on specific measures developed within the Emotion Domain for Negative Affect, Psychological Well-Being, Stress and Self-Efficacy, and Social Relationships.

CONCLUSIONS The NIH Toolbox is a broad and extensive measurement project that provides tools to assess emotional, cognitive, sensory, and motor *health*, as opposed to symptoms, impairment, or disability. That said, many of the concepts measured by the NIH Toolbox bear a close and at times even direct relationship to other large NIH initiatives, such as the Patient Reported Outcomes Measurement Information System (PROMIS; www.nihpromis.org), Quality of Life Outcomes in Neurological Disorders (Neuro-QOL; www.neuroqol.org), and the National Institute of Neurological Disorders and Stroke Common Data Elements project (http://www.commondataelements.ninds.nih.gov). For example, in the emotional health area, NIH Toolbox NA assessments for sadness, fear, and anger draw from the same pool of questions as the PROMIS item banks do for depression, anxiety, and anger, respectively. We see this overlap as advantageous, allowing researchers across a wide variety of ongoing and future studies to share a common metric for these important concepts.

In summary, the NIH Toolbox mandate was to develop brief, psychometrically sound measures for

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Table 5 Confirmatory factor analysis by Social Relationships instrument—single-factor model

CFA model	No. of items	CFI	RMSEA	Cronbach α	Convergent validity
Adult 18+ self-report					
NIH Toolbox Emotional Support	8	0.994	0.112	0.97	0.78[a]
NIH Toolbox Instrumental Support	8	0.966	0.166	0.95	0.69[a]
NIH Toolbox Loneliness	5	0.992	0.267	0.94	0.83[b]
NIH Toolbox Friendship	8	0.985	0.141	0.95	−0.80[b]
NIH Toolbox Perceived Hostility	8	0.979	0.101	0.94	0.64[c]
NIH Toolbox Perceived Rejection	8	0.975	0.142	0.93	0.60[c]
13–17 Self-report					
NIH Toolbox Pediatric Emotional Support	7	0.981	0.092	0.91	0.66[d]
NIH Toolbox Pediatric Loneliness	7	0.987	0.087	0.92	0.70[e]
NIH Toolbox Pediatric Friendship	5	0.992	0.116	0.86	−0.55[e]
NIH Toolbox Pediatric Perceived Hostility	5	0.993	0.073	0.88	0.28[f]
NIH Toolbox Pediatric Perceived Rejection	7	0.999	0.020	0.90	0.18[f]
8–12 Self-report					
NIH Toolbox Pediatric Emotional Support	7	0.950	0.111	0.91	0.55[d]
NIH Toolbox Pediatric Loneliness	7	0.954	0.168	0.93	0.68[e]
NIH Toolbox Pediatric Friendship	5	0.995	0.080	0.87	−0.59[e]
NIH Toolbox Pediatric Perceived Hostility	5	0.979	0.119	0.87	0.38[f]
NIH Toolbox Pediatric Perceived Rejection	7	0.995	0.059	0.90	0.33[f]
8–12 Proxy					
NIH Toolbox Positive Parental Relationships	4	0.980	0.114	0.81	0.21[g]
NIH Toolbox Negative Parental Relationships	4	>0.999	<0.001	0.76	−0.07[g]
NIH Toolbox Social Withdrawal	4	>0.999	<0.001	0.85	−0.15[g]
NIH Toolbox Positive Peer Interactions	4	0.971	0.198	0.79	0.26[g]
NIH Toolbox Peer Rejection	9	0.951	0.182	0.94	−0.18[g]
NIH Toolbox Sibling Rejection	9	0.961	0.208	0.94	0.02[g]
NIH Toolbox Empathic Behaviors	10	0.963	0.085	0.90	0.24[g]
3–7 Proxy					
NIH Toolbox Positive Parental Relationships	4	0.984	0.073	0.71	0.13[g]
NIH Toolbox Negative Parental Relationships	4	0.987	0.120	0.77	−0.15[g]
NIH Toolbox Social Withdrawal	4	0.999	0.036	0.80	−0.03[g]
NIH Toolbox Positive Peer Interactions	4	0.958	0.195	0.73	0.25[g]
NIH Toolbox Peer Rejection	9	0.944	0.160	0.92	0.07[g]
NIH Toolbox Sibling Rejection	9	0.956	0.156	0.93	0.08[g]
NIH Toolbox Empathic Behaviors	10	0.970	0.086	0.90	0.31[g]

Abbreviations: CFA = confirmatory factor analysis; CFI = comparative fit index; RMSEA = root mean square error of approximation.
Validation scales:
[a] Interpersonal Support Evaluation List.
[b] UCLA Loneliness Scale.
[c] Negative Interactions Scale.
[d] Monitoring the Future-support scale.
[e] Monitoring the Future-loneliness scale.
[f] NCS Sibling Scale.
[g] Social Network Index-number of high-contact roles.

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

use in large-scale, longitudinal, epidemiologic, or clinical trial studies. The Emotion Domain has identified 4 subdomains of particular relevance to health outcomes: Negative Affect, Psychological Well-Being, Stress and Self-Efficacy, and Social Relationships. Preliminary calibration and validation data have guided item and measure selection across the lifespan for each emotional health concept. Subsequent norming activites, to be reported in the future, enabled cross-domain comparisons, further refined the assessments, and enhanced the quality of the NIH Toolbox to ensure efficient measurement of key emotional health factors linked to health outcomes.

AUTHOR CONTRIBUTIONS

Drs. Cella, Pilkonis, Choi, Butt, Salsman, and Zill: study concept and design. Drs. Butt, Salsman, Lai, Choi, Bode, and Griffith: acquisition of data. Drs. Salsman, Butt, Pilkonis, Cyranowski, Zill, Hendrie, Kupst, Kelly, Bode, Choi, Lai, Griffith, Stoney, Browers, and Cella: analysis and interpretation. Drs. Salsman, Butt, Pilkonis, Cyranowski, Zill, Hendrie, Kupst, Kelly, Bode, Choi, Lai, Griffith, Stoney, Browers, Knox, and Cella: critical revision of the manuscript for important intellectual content. Drs. Cella and Pilkonis: study supervision.

ACKNOWLEDGMENT

The authors express their appreciation to Tara L. Moore, MA, MPH, who coordinated the NIH Toolbox work at the University of Pittsburgh Medical Center. The authors thank the subdomain consultants and Gitanjali Taneja, PhD, of the NIH project team, for providing critical and constructive expertise during the development of the NIH Toolbox Emotion measurement battery.

STUDY FUNDING

This study is funded in whole or in part with Federal funds from the Blueprint for Neuroscience Research and the Office of Behavioral and Social Sciences Research, NIH, under contract no. HHS-N-260-2006-00007-C.

DISCLOSURE

J. Salsman served as an independent contractor for the RTOG (Study no. 0841) and has received research support from the NIH (Contract no. HHSN265200423601C), GlaxoSmithKline, the American Cancer Society–Illinois Division (PSB-08-15), and the NIH (U01 AR052177-05). He currently receives research support from the National Cancer Institute (5K07CA158008-01A1). Z. Butt served as a consultant for Johnson & Johnson and the American Society of Transplant Surgeons. He received support from Boehringer Ingelheim, Pfizer, the Patient-Centered Outcomes Research Institute (PCORI), and NIH grant KL2RR0254740. He currently receives research support from Daiichi Sankyo, Inc., the Frankel Foundation, and NIH grants HD067440, DK091786, DK062467-10, and DK062467-08S1. P. Pilkonis served as a consultant for Lundbeck Pharmaceuticals. He currently receives support from NIH grants AR052155, AT006453, MH056888, MH066302, and MH090333. J. Cyranowski has received research funding from the Pittsburgh Foundation and from NIH grants HL076852, AR52155, and MH37896; she currently receives research support from NIH grants MH085874, MH083647, AG012546, AR052186, and MH086637. N. Zill received research support from the National Council for Adoption, the Brookings Institution, and the Marriage and Religion Research Institute. He served as a reviewer for an NIH SBIR Review Panel. He holds a TIAA-CREF Retirement Annuity Contract that invests in US Treasury Bonds and an International Stock Index Fund. He received consulting income and income from selling stock and exercising stock options from Westat, an employee-owned S Corporation. He holds an IRA and brokerage account with Vanguard that includes holdings in the Vanguard Health Care, Precious Metals and Mining, High-Yield Corporate Bond, High-Yield Tax Exempt Bond, and Long-Term Corporate Bond mutual funds, as well as the Vanguard Consumer Discretionary, Consumer Staples, FTSE International Small Cap, and Corporate Long-Term Bond Exchange Traded Funds. He has holdings in a number of closed-ended mutual funds, including the Aberdeen Asia Pacific Income Fund, Templeton Global High Income Fund, Templeton Dragon Fund, Alliance-Bernstein Global High Income Fund, Alliance-Bernstein Corporate Income Fund, Blackrock Income Opportunity Trust, Putnam Master Intermediate Income Trust, Nuveen Floating Rate Income Opportunity Fund, India Fund, Morgan Stanley India Investment Fund, Latin American Discovery Fund, Aberdeen Latin American Equity Fund, Singapore Fund, and Turkish Investment Fund. He has stock holdings in AT&T, Boardwalk Pipeline Partners, Banco de Columbia, Consolidated Edison, Exxon, Frontier Communications, GlaxoSmithKline, Honeywell, Huntington Ingalls Industries, IBM, Intel, 3M, MeadWestvaco, Newell Rubbermaid, Northrop Grumman, Occidental Petroleum, SCANA, Sherwin-Williams, Siemens, Southern, Verizon, Johnson Controls, American Superconductor, Ocean Power Technology, and Maxwell Technology. Dr. Zill's wife, Karen, prepared discussion guides for the *Independent Lens* program on PBS. She holds a TIAA Retirement Annuity Contract that invests in US Treasury Bonds. H. Hendrie currently receives research funding from NIH/NIA grant R01AG009956, R24MH080827, 5R01AG026096-05, UF20303/U01AG022376, R01AG031222, R01AG019181, and R01AG029884. M. Kupst reports no disclosures. M. Kelly has been employed by the University of Pittsburgh School of Medicine and Family Services of Western Pennsylvania. Dr. Kelly is the co–principal investigator on a study sponsored by Ortho McNeil Janssen for which Family Services of Western Pennsylvania receives remuneration. Dr. Kelly does not receive direct remuneration for her role in that study. R. Bode reports no disclosures. S. Choi has received research support from Boehringer-Ingelheim, Novartis, and the NIH. J.-S. Lai has received research support from the NIH, Agency for Healthcare Research and Quality, and Pfizer, Inc. J. Griffith has received financial support from NorthShore University HealthSystem, the Cleveland Clinic Foundation/Teva Neurosciences, Inc., Ironwood Pharmaceuticals, Inc., and Forest Laboratories, Inc., the NIH, the Department of Defense (DOD)–United States Army, and the FWO, Belgium. In addition to NIH Toolbox funding, he receives funding from the NIH for other research (grant U01 DK082342). He has also been a paid consultant to Dr. Kathryn Grant of DePaul University, and maintains a clinical psychology practice for which he bills for his services. C. Stoney reports no disclosures. P. Brouwers and S. Knox report no disclosures. D. Cella serves on the editorial board of the *Journal of Supportive Oncology*, has received travel support from Pfizer, honoraria from Virginia Commonwealth University and Moffit Cancer Center for speaking engagements, and research support from Boehringer-Ingelheim, Novartis, and the NIH. He also receives royalties from *Up to Date*. Go to Neurology.org for full disclosures.

Received June 6, 2012. Accepted in final form October 3, 2012.

REFERENCES

1. Watson D, Tellegen A. Toward a consensual structure of mood. Psychol Bull 1985;98:219–235.
2. Anastasi A, Urbina S. Psychological Testing. Upper Saddle River, NJ: Prentice Hall; 1997.
3. Lord FM. Applications of Item Response Theory to Practical Testing Problems. Hillsdale, NJ: Lawrence Erlbaum; 1980.
4. Richardson M. The relation between the difficulty and the differential validity of a test. Psychometrika 1936;1:33–49.
5. Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use. New York: Oxford University Press; 1995.
6. Watson D, Clark LA. Affects separable and inseparable: on the hierarchical arrangement of the negative affects. J Pers Soc Psychol 1992;62:489–505.
7. Stith SM, Hamby SL. The anger management scale: development and preliminary psychometric properties. Violence Vict 2002;17:383.
8. Suchday S, Carter MM, Ewart CK, Larkin KT, Desiderato O. Anger cognitions and cardiovascular recovery following provocation. J Behav Med 2004;27:319–341.
9. Samman E; University of Oxford Poverty and Human Development Initiative. Psychological and Subjective Wellbeing: A Proposal for Internationally Comparable

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Indicators. Oxford: University of Oxford Poverty and Human Development Initiative; 2007.
10. Cohen S, Pressman SD. Positive affect and health. Curr Dir Psychol Sci 2006;15:122–125.
11. Diener E, Suh EM, Lucas RE, Smith HL. Subjective well-being: three decades of progress. Psychol Bull 1999;125: 276–302.
12. Ryan RM, Huta V, Deci E. Living well: a self-determination theory perspective on eudaimonia. J Happiness Stud 2008;9: 139–170.
13. Danner DD, Snowdon DA, Friesen WV. Positive emotions in early life and longevity: findings from the nun study. J Pers Soc Psychol 2001;80:804–813.
14. Pressman SD, Cohen S. Does positive affect influence health? Psychol Bull 2005;131:925–971.
15. Fredrickson BL, Levenson RW. Positive emotions speed recovery from the cardiovascular sequelae of negative emotions. Cogn Emot 1998;12:191–220.
16. Richman LS, Kubzansky L, Maselko J, Kawachi I, Choo P, Bauer M. Positive emotion and health: going beyond the negative. Health Psychol 2005;24:422–429.
17. Diener E. Subjective well-being: the science of happiness and a proposal for a national index. Am Psychol 2000;55:34–43.
18. Koizumi M, Ito H, Kaneko Y, Motohashi Y. Effect of having a sense of purpose in life on the risk of death from cardiovascular diseases. J Epidemiol 2008;18:191–196.
19. Verduin PJ, de Bock GH, Vliet Vlieland TP, Peeters AJ, Verhoef J, Otten W. Purpose in life in patients with rheumatoid arthritis. Clin Rheumatol 2008;27:899–908.
20. Mascaro N, Rosen DH. Existential meaning's role in the enhancement of hope and prevention of depressive symptoms. J Pers 2005;73:985–1013.
21. Fredrickson BL. The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions. Am Psychol 2001;56:218–226.
22. Lazarus RS, Folkman S. Stress, Appraisal, and Coping. New York: Springer Publishing Co.; 1984.
23. Cohen S, Kessler RC, Gordon LU. Strategies for measuring stress in studies of psychiatric and physical disorder. In: Cohen S, Kessler RC, Gordon LU, editors. Measuring Stress: A Guide for Health and Social Scientists. New York: Oxford University Press; 1997:3–26.
24. Bandura A. Self-efficacy: The Exercise of Control. New York: Freeman; 1997.
25. Cohen S, Miller GE, Rabin BS. Psychological stress and antibody response to immunization: a critical review of the human literature. Psychosom Med 2001;63:7–18.
26. Segerstrom SC, Miller GE. Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. Psychol Bull 2004;130:601–630.
27. O'Leary A. Self-efficacy and health. Behav Res Ther 1985; 23:437–451.
28. Cohen S, Gordon LU, Gottlieb BH, Fetzer I. Social Support Measurement and Intervention: A Guide for Health and Social Scientists. New York: Oxford University Press; 2000.
29. Cohen S. Social relationships and health. Am Psychol 2004;59:676–684.
30. Cacioppo JT, Hawkley LC, Crawford LE, et al. Loneliness and health: potential mechanisms. Psychosom Med 2002; 64:407–417.
31. Cacioppo JT, Hughes ME, Waite LJ, Hawkley LC, Thisted RA. Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. Psychol Aging 2006;21:140–151.
32. Seeman TE. Health promoting effects of friends and family on health outcomes in older adults. Am J Health Promot 2000; 14:362–370.
33. Uchino BN. Social Support and Physical Health Outcomes: Understanding the Health Consequences of Our Relationships. New Haven: Yale University Press; 2004.

© 2013 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.

Neurology®

Emotion assessment using the NIH Toolbox

John M. Salsman, Zeeshan Butt, Paul A. Pilkonis, et al.

Neurology 2013;80;S76-S86
DOI 10.1212/WNL.0b013e3182872e11

This information is current as of March 11, 2013

Neurology ® is the official journal of the American Academy of Neurology. Published continuously since 1951, it is now a weekly with 48 issues per year. Copyright © 2013 American Academy of Neurology. All rights reserved. Print ISSN: 0028-3878. Online ISSN: 1526-632X.

AMERICAN ACADEMY OF NEUROLOGY®

Updated Information & Services	including high resolution figures, can be found at: http://www.neurology.org/content/80/11_Supplement_3/S76.full.html
Supplementary Material	Supplementary material can be found at: http://www.neurology.org/content/suppl/2013/03/09/80.11_Supplement_3.S76.DC1
References	This article cites 24 articles, 2 of which you can access for free at: http://www.neurology.org/content/80/11_Supplement_3/S76.full.html#ref-list-1
Citations	This article has been cited by 2 HighWire-hosted articles: http://www.neurology.org/content/80/11_Supplement_3/S76.full.html#otherarticles
Subspecialty Collections	This article, along with others on similar topics, appears in the following collection(s): **All Neuropsychology/Behavior** http://www.neurology.org//cgi/collection/all_neuropsychology_behavior **Neuropsychological assessment** http://www.neurology.org//cgi/collection/neuropsychological_assessment **Outcome research** http://www.neurology.org//cgi/collection/outcome_research
Permissions & Licensing	Information about reproducing this article in parts (figures,tables) or in its entirety can be found online at: http://www.neurology.org/misc/about.xhtml#permissions
Reprints	Information about ordering reprints can be found online: http://www.neurology.org/misc/addir.xhtml#reprintsus

Neurology ® is the official journal of the American Academy of Neurology. Published continuously since 1951, it is now a weekly with 48 issues per year. Copyright © 2013 American Academy of Neurology. All rights reserved. Print ISSN: 0028-3878. Online ISSN: 1526-632X.

AMERICAN ACADEMY OF NEUROLOGY®