

B.C.

by johnny hart



# Measurement in Psychosocial Research: Do we have a thermometer?

MCUAAAR ANC Scientist Seminar  
November 15, 2023  
Briana Mezuk ([bmezuk@umich.edu](mailto:bmezuk@umich.edu))

# Outline

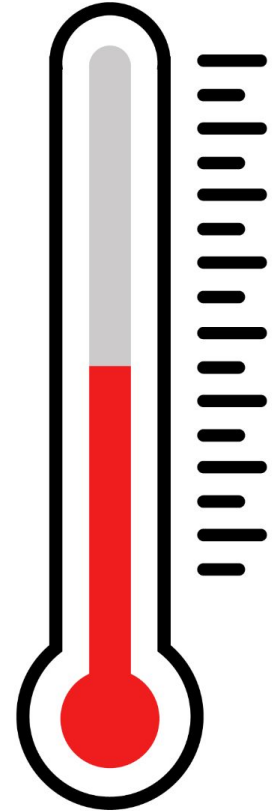
- **Deconstructing your data:** Not all variables have the same sources of error
- **Defining the challenge:** Conceptual and psychometric non-equivalence of psychosocial measures
  - Example: Centers for Epidemiologic Studies - Depression (CESD) Scale
- **Important resources** for finding existing psychosocial measures for adapting measures to, or creating new measures for, specific populations

First rule of survey research:  
If you ask, you will get an answer.

For the purposes of this talk,  
***What is a thermometer?***

A tool that assesses the true value of construct (i.e., core body temperature) reliably, i.e., regardless of age, sex, race, ethnicity, health status, location, etc.

Provides confidence that any differences observed between groups are “true” and not due to measurement bias.



# Typologies of variables and sources of error

- Variables that **must be assessed by self-report** because there is **no other valid source of information**
  - Examples: Emotions, opinions, beliefs, stated preferences
  - Commonly assessed using quasi-continuous scales (e.g., visual analogue scale, Likert scale)
  - **Main challenge:** There may be multiple “validated” scales to choose from
  - **Solution:** Include multiple measures of the same construct to be able to deconstruct the exposure-outcome relationship into “construct” and “measurement”

# Typologies of variables and sources of error

- Variables that **can be assessed by self-report** but for which there is **another source of information that may be more valid (as it relates to your research question)**
  - Examples: History of medical diagnoses, healthcare utilization, job history, education history, dietary history
  - **Main challenge:** Demonstrating that your measurement strategy is a reasonable proxy (i.e., is strongly concordant with) this more valid measure
  - **Solution:** Link your survey data to these external data sources (e.g., medical records), if possible.

# Typologies of variables and sources of error

- Variables that **can only be assessed** through a **biological or clinical test**
  - Examples: White blood cell count, genotype, DNA methylation, blood pressure
  - **Main challenge:** Ensuring reliable implementation of best-practices in sample collection and sample processing (e.g., considerations for things like fasting status, time of day, source of sample - capillary vs. venous blood, etc.)
  - **Solution:** Training staff on the principles of data collection, working with sample processing vendors that have an established track record of analyzing the specific biomarker you are interested in.

# Typologies of variables and sources of error

- Variables that **can only be assessed** through **elicitation or challenge**
  - Examples: Cognition (e.g., memory, processing speed), revealed preferences (through experiments), HPA-axis reactivity, OGTT, Dexamethasone Suppression Test
  - **Main challenge:** Practice effects (e.g., especially for cognition) or one-time deception (e.g., Trier Social Stress Test)
  - **Solution:** Include multiple practice trials (e.g., cognition), use balanced designs (randomize order of measurement) for deception tasks



# Typologies of variables and sources of error

- Variables that are **assigned** (by the investigator) through some sort of **operationalization** in which a **concept** is applied to the original data
  - Examples: “High inflammation,” “depressed,” “lonely,” “disadvantaged,” etc.
  - **Main challenge:** Loss of information through categorization, threshold may not be appropriate
  - **Solution:** Test multiple thresholds for assignment (e.g., “mild”, “moderate”, “severe” depression) to evaluate whether inferences are unduly influenced by the assignment strategy

# Summary

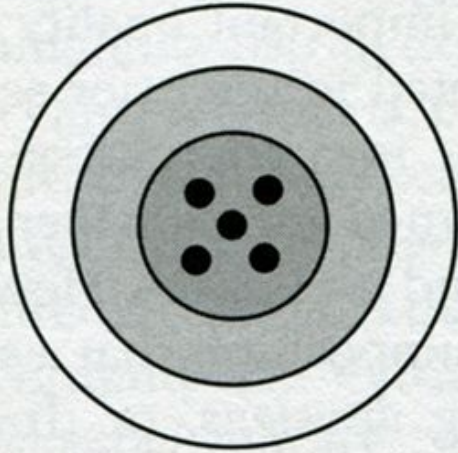
## Typologies of variables and sources of error

1. Variables that **must be assessed by self-report** because there is no other valid source of information
2. Variables that **can be assessed by self-report** but for which there is another source of information that may be more valid
3. Variables that **can only be assessed** through a **biological or clinical test**
4. Variables that **can only be assessed** through **elicitation or challenge**
5. Variables that are **assigned** by applying a concept to original data

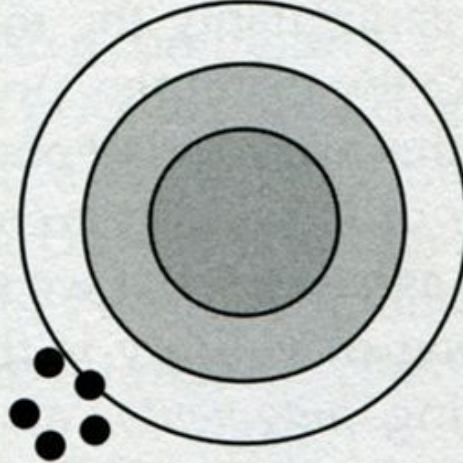
# Things to consider when selecting a measure

- **Purpose:** Why are you assessing this construct in your study anyway?
- **Population being studied:** General population vs. clinic sample, etc.
- **Method of assessment:** mail vs. telephone vs. web vs. in-person
- Copyright/costs
- **Psychometric properties:** Reliability & Validity

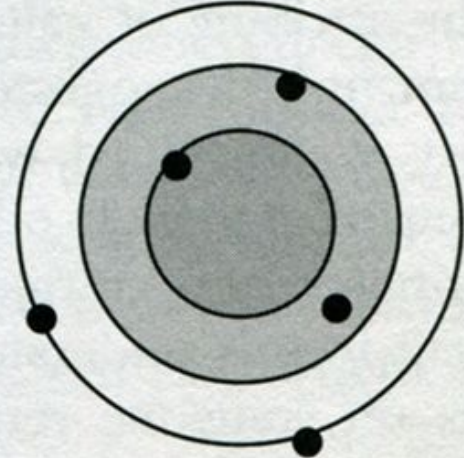
# Relationship between Reliability and Validity



✓ Valid  
✓ Reliable



X Valid  
✓ Reliable



X Valid  
X Reliable

**Sensitivity** and **Specificity** are only meaningful relative to a known status that is a categorical determination that a state/trait is either present or absent based on a “gold standard.”

Such gold standards generally do not exist for most psychosocial constructs.

# Why does measurement matter in minority health research?

Social/behavioral health research largely depends on self-report/survey assessments.

In order to make valid statements of differences (whether between racial/ethnic groups, or within racial/ethnic groups by age, gender, education, over time, etc.) the measures must have minimal bias.

## Two types of bias to consider

**Non-differential** (“random” error that is similar across groups being compared)

**Differential** (more error in one group)

MEDICAL CARE  
Volume 41, Number 11, pp 1207–1220  
©2003 Lippincott Williams & Wilkins, Inc.

## Advancing Health Disparities Research Can We Afford to Ignore Measurement Issues?

ANITA L. STEWART, PhD,\*† AND ANNA M. NÁPOLES-SPRINGER, PhD\*‡

**BACKGROUND.** Research on racial and ethnic health disparities in the United States requires that self-report measures, developed primarily in mainstream samples, are appropriate when applied in diverse groups. To compare groups, mean scores must reflect true scores and have minimal bias, assumptions that have not been tested for many self-report measures used in this research.

**OBJECTIVE.** To identify conceptual and psychometric issues that need to be addressed to assure the quality of self-report measures being used in health disparities research.

**METHODS.** We present 2 broad conceptual frameworks for health disparities research and describe the main research questions and measurement issues for 4 key concepts hypothesized as potential mechanisms of health disparities: socioeconomic status, discrimination, acculturation, and quality of care. This article is based on a small conference convened by 6 Resource Centers for Minority Aging Research (RCMAR) measurement cores. We integrate written materials prepared for the conference

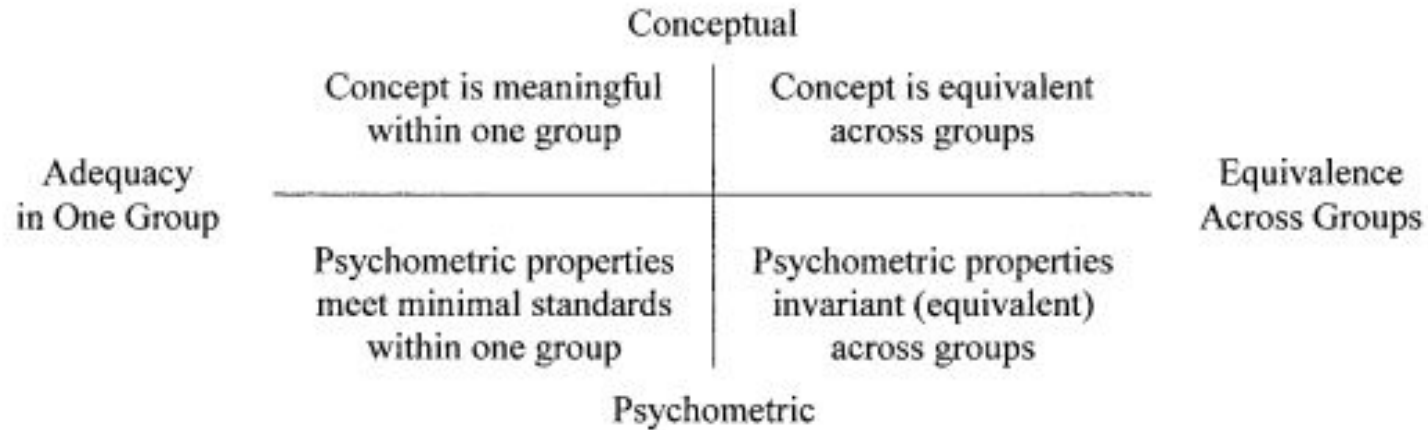
by quantitative and qualitative measurement specialists and cross-cultural researchers, conference discussions, and current literature.

**RESULTS.** Problems in the quality of the conceptualizations and measures were found for all 4 concepts, and little is known about the extent to which measures of these concepts can be interpreted similarly across diverse groups. Many problems also apply to other concepts relevant to health disparities. We propose an agenda for accomplishing this challenging measurement research.

**CONCLUSIONS.** The current national commitment to reduce health disparities may be compromised without more research on measurement quality. Integrated, systematic efforts are needed to move this work forward, including collaborative efforts and special initiatives.

**Key words:** Measurement; health disparities; patient-physician communication; discrimination; minority health; socioeconomic status; acculturation (Med Care 2003;41:1207–1220)

# Conceptual framework of psychometric “adequacy” and “equivalence” across groups



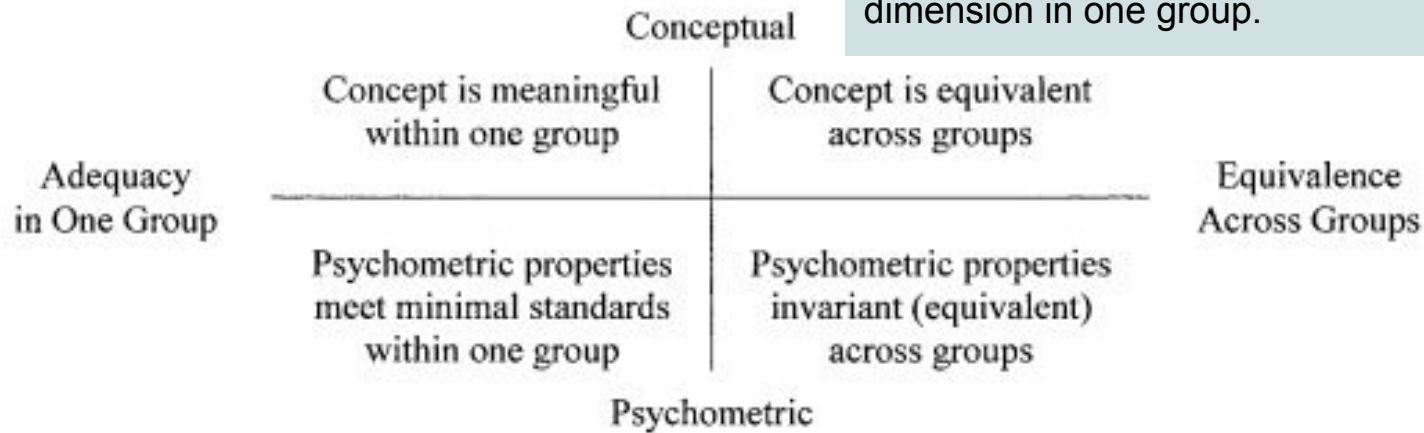
\*Group(s) can refer to any sociodemographic strata being compared

FIG. 1. Conceptual and psychometric adequacy and equivalence within one group and across groups.\*

# Conceptual framework of psychometric “adequacy” and “equivalence” across groups

## ***What causes conceptual non-equivalence?***

Culturally mediated differences in perceptions of the meaning of items and health constructs or because a concept is missing an important dimension in one group.



## ***What causes psychometric non-equivalence?***

Response bias resulting from cultural or group differences in the cognitive processes of answering, using response scales, or differences introduced by inadequate translations and failure to address varying literacy levels.

pared

within one group and across groups.\*



# Why is it that existing measures may be affected by conceptual and/or psychometric non-equivalence?

1. Limited inclusion of concepts relevant to minority populations in the creation of “universal” scales (e.g., stress, quality of life) such as cultural competence and discrimination.
2. Lack of information on the psychometric invariance of measures across diverse groups (e.g., insufficient sample size or limited variability),
3. Traditional survey methods (mail, telephone) fail to reach many minority groups, resulting in small/select samples of these groups;
4. Measures may need to be translated into other languages and written at reading levels appropriate for people with limited English proficiency.

**Goal:** to quantify the agreement between the CIDI and CESD measurement of “depression syndrome” and examine variation by age, sex, and race/ethnicity among older adults

CESD (Radloff 1977) is the most widely-used depression scale in population surveys.

Why? Brief, can be self-administered and has thresholds to indicate “clinically significant” depression.

How was its reliability and validity as a measure of “clinically significant” depression assessed?

## Measurement Article

# Shades of Blue and Gray: A Comparison of the Center for Epidemiologic Studies Depression Scale and the Composite International Diagnostic Interview for Assessment of Depression Syndrome in Later Life

Linh Dang, MPH,<sup>1</sup> Liming Dong, PhD,<sup>1,2</sup> and Briana Mezuk, PhD<sup>1,2,3,\*</sup>

<sup>1</sup>Department of Epidemiology and <sup>2</sup>Center for Social Epidemiology and Population Health, University of Michigan School of Public Health, Ann Arbor. <sup>3</sup>Research Center for Group Dynamics, Institute for Social Research, University of Michigan, Ann Arbor.

\*Address correspondence to: Briana Mezuk, PhD, Center for Social Epidemiology and Population Health, Department of Epidemiology, University of Michigan School of Public Health, Suite 2649B, 1415 Washington Heights, Ann Arbor, MI 48109. E-mail: [bmezuk@umich.edu](mailto:bmezuk@umich.edu)

Received: October 11, 2018; Editorial Decision Date: March 15, 2019

**Decision Editor:** Suzanne Meeks, PhD

## Abstract

**Background and Objectives:** Psychiatric research lacks the equivalent of a thermometer, that is, a tool that accurately measures mental disorder regardless of context. Instead, the psychometric properties of scales that purport to assess psychopathology must be continuously evaluated. To that end, this study evaluated the diagnostic agreement between the eight-item Center for Epidemiologic Studies Depression Scale (CESD-8) and the Composite International Diagnostic Interview—short form (CIDI-SF) in the Health and Retirement Study (HRS).

**Research Design and Methods:** Data come from 17,613 respondents aged >50 from the 2014 wave of the HRS. Kappa coefficients were used to assess the agreement between the 2 instruments on depression classification across a range of thresholds for identifying case status, including variation across subgroups defined by age, race/ethnicity, and gender.

**Results:** The point prevalence of depression syndrome estimated by the CESD was higher than that estimated by the CIDI-SF (CESD: 9.9%–19.5% depending on the cutoff applied to the CESD vs CIDI-SF: 7.7%). Assuming CIDI-SF as the gold standard, the CESD yielded a sensitivity of 56.2%–70.2% and specificity of 84.7%–94.0% across the range of cutoffs. The agreement on depression classification was weak ( $\kappa = 0.32$ – $0.44$ ).

**Discussion and Implications:** Depression cases identified by the CESD have poor agreement with those identified by the CIDI-SF. Conceptually, psychological distress as measured by the CESD is not interchangeable with depression syndrome as measured by the CIDI-SF. Population estimates of depression among older adults based on the CESD should be interpreted with caution.

**Keywords:** CES-D, Depression classification, Validity, Reliability

## Strategy and sample used to quantify the psychometric properties of the CESD in Radloff (1977)

Recall, the Diagnostic and Statistical Manual of Mental Disorders (DSM) - underwent a major revision in 1980, creating the atheoretical criteria we are familiar with today. *So, where did the CESD even come from?*

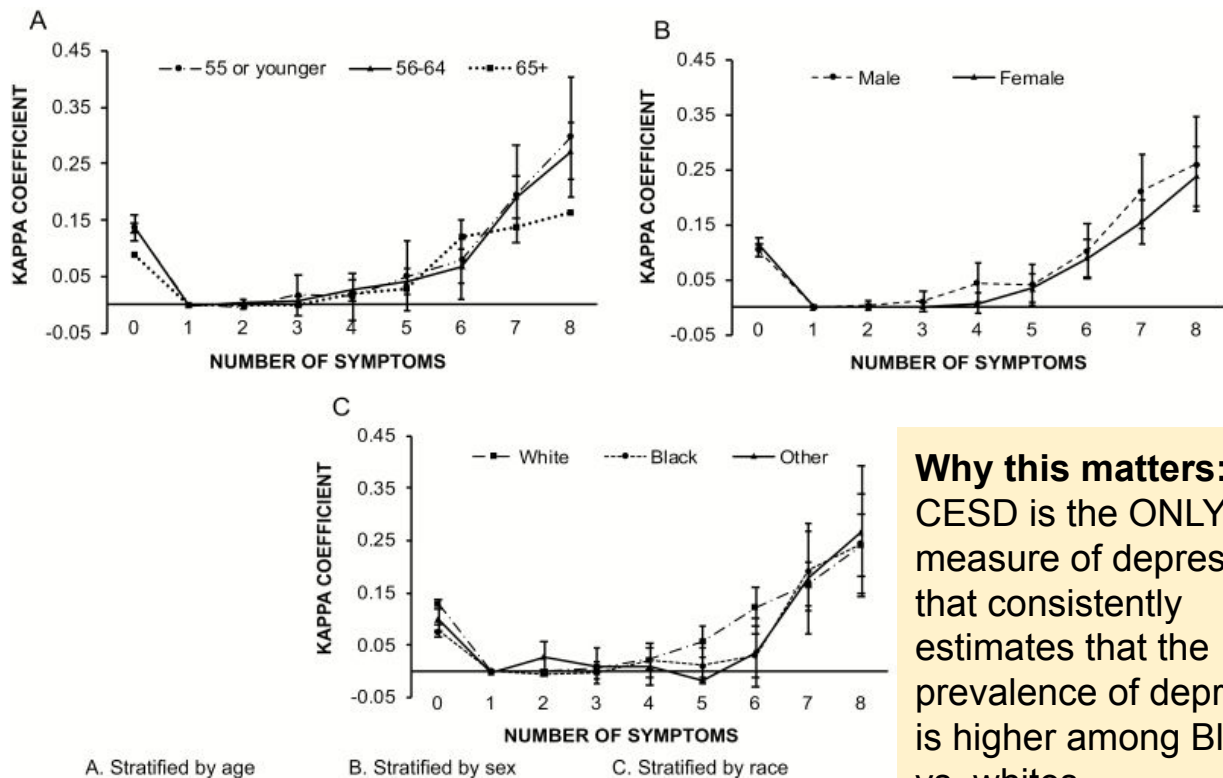
# Strategy and sample used to quantify the psychometric properties of the CESD in Radloff (1977)

## *So, where did the CESD even come from?*

**Study:** Community Mental Health Epidemiology Program (JHU)

- Two sites: Washington County, MD (very rural) and Kansas City, MO
- Total sample size: 3845
  - Number of Black participants: 295, all of whom were from Kansas City, MO
  - **Reliability:** While they don't report the exact numbers, Radloff (1977) says: “Test–retest correlations were moderate (.40 or above) in all but three groups (**Blacks**, age under 25, and “need help” [for mental distress])”
  - **Validity:** CESD was given to “true” cases of depression (n=70 in MD, n=35 in CT) and scores were higher than people given in the general population.
    - Racial/ethnic composition of the “true” cases are never reported, nor is there any text stating whether the validity of the CESD varied by race.

# Chance-corrected (Kappa coefficient) agreement between CIDI and CESD as a function of symptom count



- CESD consistently generates higher point prevalence of depression syndrome: 9.9 - 19.5% depending on threshold used, vs. 7.7% for the CIDI.
- Assuming CIDI as the gold standard, the CESD has a sensitivity of 56.2%–70.2% (**CESD produces many more false-positives**) and specificity of 84.7%–94.0% across the range of cutoffs.
  - Higher symptom counts=higher agreement between the two measures.

**Why this matters:** The CESD is the **ONLY** measure of depression that consistently estimates that the prevalence of depression is higher among Blacks vs. whites.

## **Take-away - Measure twice, cut once**

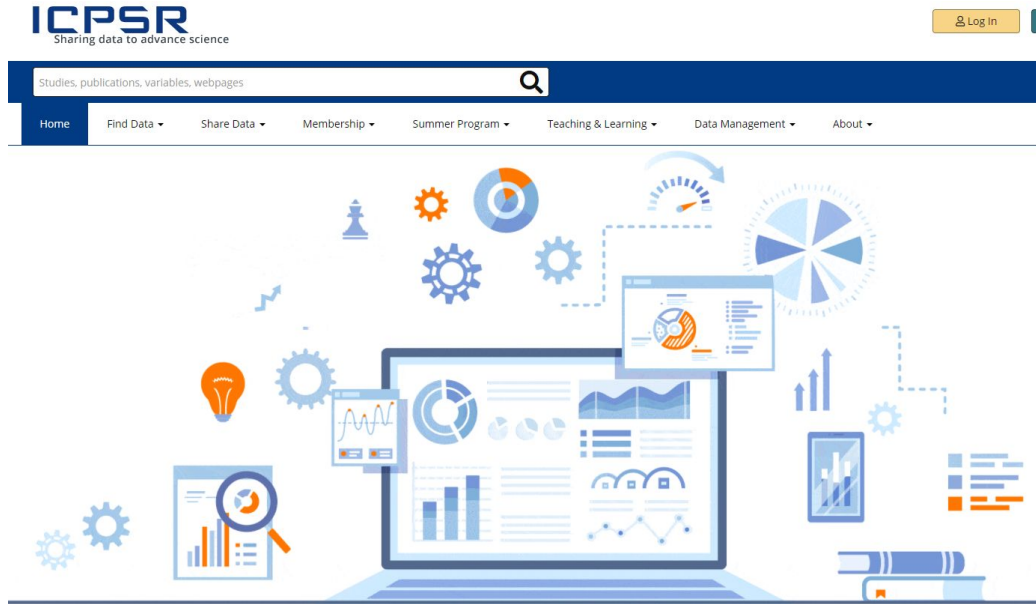
We were only able to do this study because the HRS included multiple measures of the same “concept” - depression.





Resources to find existing psychosocial measures or for creating measures of your own

# Existing studies - ICPSR



- Using the same measures that have been included in large, existing studies allows you to compare your (likely small) study's findings to larger/representative samples.
  - Can even [pool your datasets](#)!
- Running cross-tabs of potential measures fielded in ICPSR datasets allows you to see the **observed variation in them within and across groups** to get a sense of whether they are appropriate for your population of interest.



# PhenX: Phenotypes for eXposures

The screenshot shows the PhenX Toolkit website. At the top is a dark blue header with the PhenX Toolkit logo on the left and navigation buttons (Register, Log in, Link Your Study, My Toolkit) on the right. Below the header is a light blue navigation bar with links to Home, Protocols, Search, Resources, News, Help, About, and Contact. A search bar is located below the navigation bar. The main content area has an orange banner with the text "Share your thoughts: Help us improve the PhenX Toolkit!" and a link to the "New COVID-19 Variable Compare Tool". Below this is a section titled "What is the PhenX Toolkit?" with a list of bullet points. To the right of this section is a vertical menu with four items: "Research Domains", "Browse Protocols Tree", and "Research Using PhenX".

PhenX Toolkit

Register Log in Link Your Study My Toolkit

Home Protocols Search Resources News Help About Contact

Search: Search all protocols in the Toolkit using keywords (e.g. diabetes) or PhenX ID (e.g. 011502)

Share your thoughts: Help us improve the PhenX Toolkit!

New COVID-19 Variable Compare Tool, click here.

**What is the PhenX Toolkit?**

- A catalog of recommended measurement protocols
- A toolkit developed by scientific community via a consensus-based process
- A Web-based resource, freely available for use

Research Domains

Browse Protocols Tree

Research Using PhenX

Recommended standard data collection protocols for conducting biomedical research.

Tools selected by expert consensus.

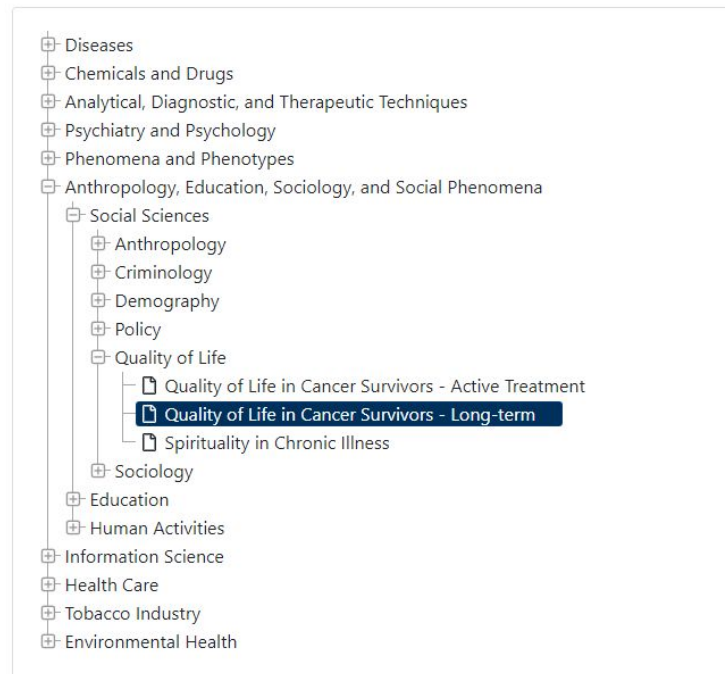
Includes both measures and data collection protocols.

Focus on genetically-informed studies.

# PhenX: Example

## Cross-Domain Concept Tree Browser

The PhenX Cross-Domain Concept Tree Browser allows users to intuitively browse protocols by scientific concept, organized by the National Library of Medicine (NLM)'s [Medical Subject Headings \(MeSH\)](#), the most widely used vocabulary in medical research. Users can identify conceptually related protocols across research domains and access protocols from multiple entry points, e.g. cystic fibrosis (CF) is under both the "Respiratory Tract Diseases" and the "Digestive System Diseases" headings.



### Quality of Life in Cancer Survivors - Long-term

[+ Add to My Toolkit](#)

#### Description

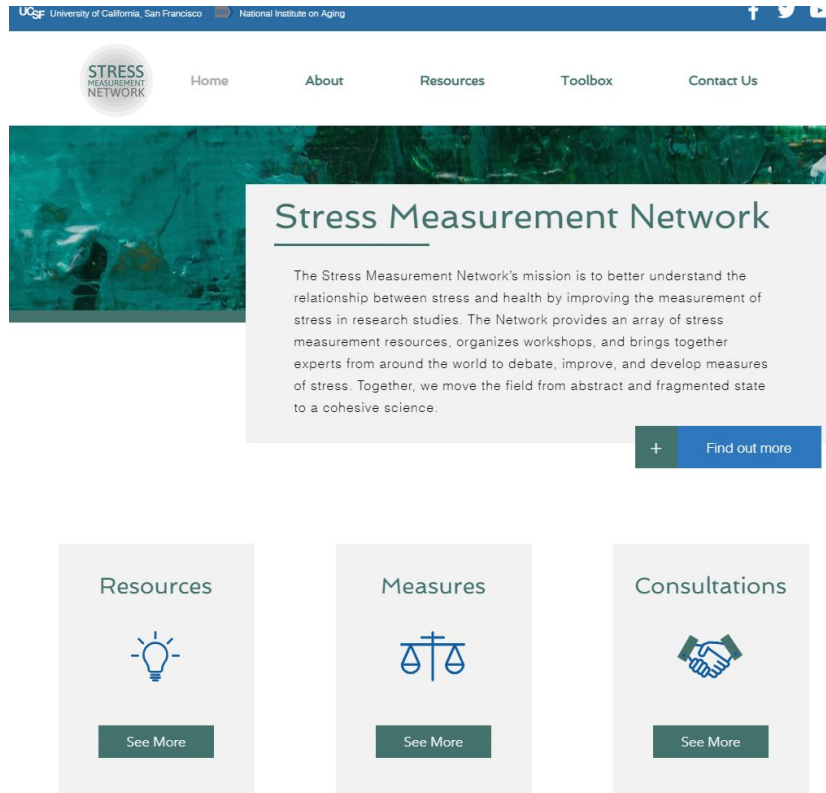
A self-report 41-item Likert style questionnaire representing the four domains of quality of life including physical well-being, psychological well-being, social well-being and spiritual well-being.

#### Specific Instructions ▼

#### Name from Source ▼

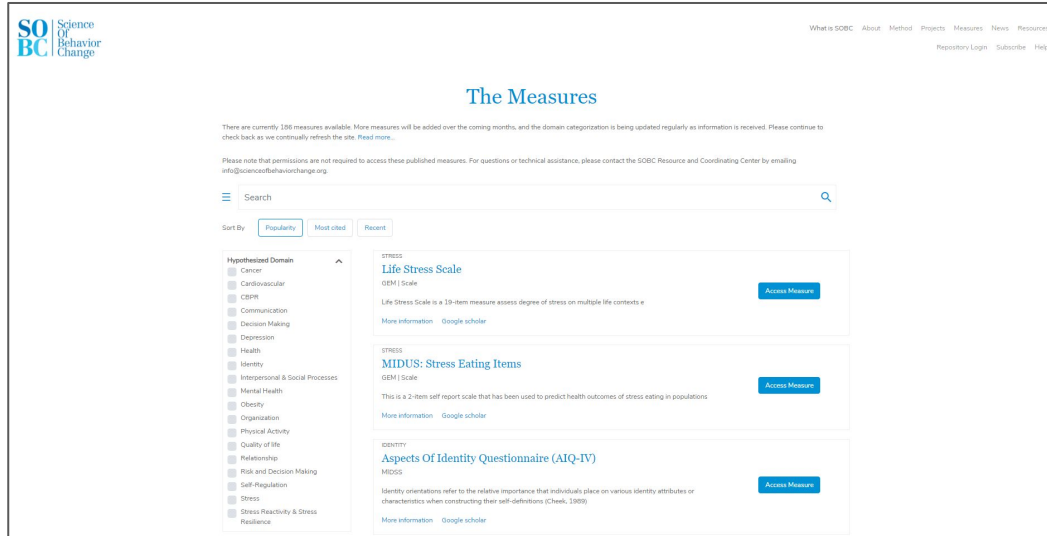
Go to the [protocol page](#).

# Stress Measurement Network



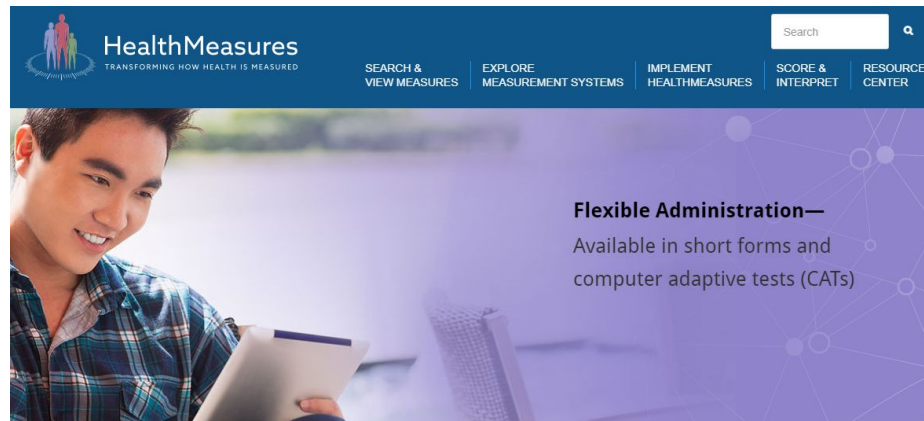
- Detailed explanations of the **rationale** for each stress measure
- **Psychological** measures of “stress” and related constructs.
  - Acute stress, early-life stress, stigma, discrimination, relationships, neighborhood safety, work stress, resilience
- **Physiologic** measures of “stress” exposure and stress response systems
  - Epigenetic clocks/biological aging/telomeres, cortisol, inflammation, skin conductance.
    - Includes some protocols of biological sample collection (ex. [Hair cortisol](#))

# Science of Behavior Change



- NIH-wide initiative to understand the underlying mechanisms of behavior change by promoting basic research on the initiation, personalization, and maintenance of behavior change.
  - Goal is to inform more effective behavioral interventions using core ideas from behavioral economics.
- **Focus on three domains**
  - Self-regulation
  - Stress reactivity/resilience
  - Interpersonal and Social Processes
- Experimental approach: Identify potential mechanisms, Measure those mechanisms, and Develop interventions that influence those mechanisms.
  - Measures are organized as self-report, task, or observational

# PROMIS: Patient Reported Outcome Measurement Information System



- Focus on **clinical care/settings/living with specific health conditions.**
- **Includes development of new measures.**
- Designed to enhance communication between clinicians and patients in diverse research and clinical settings.
- Available in multiple formats and easily integrated into diverse administration platforms.
- Translations available in Spanish and many other languages



#### Administration Guidelines during COVID-19 and Social Distancing

As part of COVID-19 guidelines and social distancing, the HealthMeasures and PROMIS team acknowledge that **recommended administration practices** may not be possible. In cases where modifications are required to continue data collection (e.g., administration via phone interview), this is allowable at this time. We recommend that the required modification be documented and that each assessment that was done using a modified administration is "flagged" for subsequent analysis. When shared devices are used, **disinfection protocols** must be followed.

#### INTRO TO PROMIS

#### OBTAIN & ADMINISTER MEASURES







#### MEASURE DEVELOPMENT & RESEARCH

PROMIS® (Patient-Reported Outcomes Measurement Information System) is a set of person-centered measures that evaluates and monitors physical, mental, and social health in adults and children. It can be used with the general population and with individuals living with chronic conditions.

# Resource Centers for Minority Aging Research Resources:

## How to conduct focus groups & cognitive interviews for developing new measures

Keep up with the latest on RCMARCH4ME?

UCLA      

HOME ABOUT APPLY EVENTS PEOPLE PUBLICATIONS RESEARCH CORES LINKS


### QUALITATIVE METHODS

---

Qualitative research methods resources

- [Focus Group Protocol](#)
- [Focus Group PowerPoint Presentation](#)
- [Focus Group Write-up Example](#)
- [Cognitive Interview Protocol](#)
- [Cognitive Interview PowerPoint Presentation](#)
- [Cognitive Interview Guide](#)

Cognitive Interview demonstration. Below is a 15 minute video demonstrating a cognitive interview.



The video player shows a demonstration of a cognitive interview. It features two women in a professional setting, one of whom is speaking. The video player interface includes a play button, a progress bar, and a share icon.



# An example of creating a new scale of financial exploitation risk among older adults

JOURNAL OF ELDER ABUSE & NEGLECT  
2017, VOL. 28, NO. 3, 134–151  
<http://dx.doi.org/10.1080/08946566.2016.1168333>



Innovation in Aging  
2017, Vol. 40, No. 1, 9  
doi:10.1080/geron/40.003  
Advance Access publication May 28, 2017



JOURNAL OF ELDER ABUSE & NEGLECT  
2017, VOL. 29, NO. 4, 213–228  
<https://doi.org/10.1080/08946566.2017.1338170>



## The Lichtenberg Financial Decision Screening Scale (LFDSS): A new tool for assessing financial decision making and preventing financial exploitation

Peter A. Lichtenberg, PhD, ABPP<sup>a,b</sup>, Lisa Ficker, PhD<sup>c</sup>, Analise Rahman-Filipiak, MA<sup>a,b</sup>, Ron Tatro, BA<sup>c</sup>, Cynthia Farrell, MSW<sup>d</sup>, James J. Speir, MSW<sup>a</sup>, Sanford J. Mall, JD<sup>f</sup>, Patrick Simasko, JD<sup>g</sup>, Howard H. Collens, JD<sup>h</sup>, and John Daniel Jackman Jr., MD<sup>i</sup>

<sup>a</sup>Institute of Gerontology, Wayne State University, Detroit, Michigan, USA; <sup>b</sup>Department of Psychology, Wayne State University, Detroit, Michigan, USA; <sup>c</sup>Center for Elder Rights Advocacy, Elder Law of Michigan, Lansing, Michigan, USA; <sup>d</sup>Aging and Adult Services, Adult Protective Services, State of Michigan Department of Health and Human Services, Lansing, Michigan, USA; <sup>e</sup>Speir Financial Services, Southfield, Michigan, USA; <sup>f</sup>Mall, Malisow and Cooney, PC, Farmington Hills, Michigan, USA; <sup>g</sup>Simasko and Simasko Law Firm, Mount Clemens, Michigan, USA; <sup>h</sup>Galloway and Collens, PLLC, Huntington Woods, Michigan, USA; <sup>i</sup>Private Practice Cardiologist (Ret.), Tyler, Texas, USA

### ABSTRACT

One of the challenges in preventing the financial exploitation of older adults is that neither criminal justice nor noncriminal justice professionals are equipped to detect capacity deficits. Because decision-making capacity is a cornerstone assessment in cases of financial exploitation, effective instruments for measuring this capacity are essential. We introduce a new screening scale for financial decision making that can be administered to older adults. To explore the scale's implementation and assess construct validity, we conducted a pilot study of 29 older adults seen by APS (Adult Protective Services) workers and 79 seen by other professionals. Case examples are included.

### KEYWORDS

Financial decision making; financial exploitation; financial judgment; protective services

### Original Research Article

## Reliability and Validity of the Lichtenberg Financial Decision Screening Scale

Peter A. Lichtenberg, PhD, ABPP<sup>a,\*</sup>, Jeanne A. Teresi, EdD, PhD<sup>a,c</sup>, Katja Ocepek-Welikson, MPhil<sup>b</sup>, Joseph P. Eimicke, MS<sup>d,4</sup>

<sup>a</sup>Institute of Gerontology, Wayne State University, Detroit, Michigan. <sup>b</sup>Columbia University Stroud Center at New York State Psychiatric Institute. <sup>c</sup>Research Division, Hebrew Home at Riverdale, River Spring Health, New York. <sup>d</sup>Department of Geriatrics and Palliative Medicine, Weill Cornell Medical Center, New York.

\*Address correspondence to Peter A. Lichtenberg, PhD, ABPP, Institute of Gerontology, Wayne State University, 87 E. Ferry Street, Detroit, MI 48202. E-mail: [p.lichtenberg@wayne.edu](mailto:p.lichtenberg@wayne.edu)

Received: January 11, 2017; Editorial Decision Date: March 20, 2017

Decision Editor: Laura P. Sands, PhD

### Abstract

The scarcity of empirically validated assessment instruments continues to impede the work of professionals in a number of fields, including medicine, finance, and estate planning; adult protective services; and criminal justice—and, more importantly, it impedes their ability to effectively assist and, in some case, protect their clients. Other professionals (e.g., legal, financial, medical, mental health services) are in a position to prevent financial exploitation and would benefit from access to new instruments. The Lichtenberg Financial Decision Screening Scale (LFDSS) was introduced in 2016, along with evidence for its convergent validity (Lichtenberg, P. A., Fickern, L., Rahman-Filipiak, A., Tatro, R., Farrell, C., Speir, J. J., ... Jackman, J. D. (2016b). The Lichtenberg Financial Decision Screening Scale: A new tool for assessing financial decision making and preventing financial exploitation (2016). *Journal of Elder Abuse and Neglect*, 28, 134–151. doi:10.1080/08946566.2016.1168333). Using a sample of 213 participants, this study investigated the internal consistency of the LFDSS and its criterion validity based on ratings by professionals using the scale. Results demonstrate that the LFDSS has excellent internal consistency and clinical utility properties. This paper provides support for use of the LFDSS as a reliable and valid instrument. The LFDSS and instructions for its use are included in the article, along with information about online tools and support.

**Translational Significance:** The screening scale presented in this article can be used in a variety of settings to help assess decisional capacity and prevent financial exploitation (e.g., Adult Protective Services, Medical offices, legal services, financial advisors). Few empirically validated, efficient scales are available for these professionals to assess the older adult's decision making capacity around specific financial transactions. This scale is offered to help fill that void.

**Keywords:** Financial decision making, Financial capacity, Financial exploitation

## Item response theory analysis of the Lichtenberg Financial Decision Screening Scale

Jeanne A. Teresi, EdD, PhD<sup>a,b,c</sup>, Katja Ocepek-Welikson, MPhil<sup>a</sup>, and Peter A. Lichtenberg, PhD, ABPP<sup>a</sup>

<sup>a</sup>Institute of Gerontology and Merrill Palmer Skillman Institute, Wayne State University, Detroit, Michigan, USA; <sup>b</sup>Research Division, Hebrew Home at Riverdale, River Spring Health, Riverdale, New York, USA; <sup>c</sup>Division of Geriatrics and Palliative Medicine, Weill Cornell Medical College, Cornell University, New York, New York, USA

### ABSTRACT

The focus of these analyses was to examine the psychometric properties of the Lichtenberg Financial Decision Screening Scale (LFDSS). The purpose of the screen was to evaluate the decisional abilities and vulnerability to exploitation of older adults. Adults aged 60 and over were interviewed by social, legal, financial, or health services professionals who underwent in-person training on the administration and scoring of the scale. Professionals provided a rating of the decision-making abilities of the older adult. The analytic sample included 213 individuals with an average age of 76.9 ( $SD = 10.1$ ). The majority (57%) were female. Data were analyzed using item response theory (IRT) methodology. The results supported the unidimensionality of the item set. Several IRT models were tested. Ten ordinal and binary items evidenced a slightly higher reliability estimate (0.85) than other versions and better coverage in terms of the range of reliable measurement across the continuum of financial incapacity.

### KEYWORDS

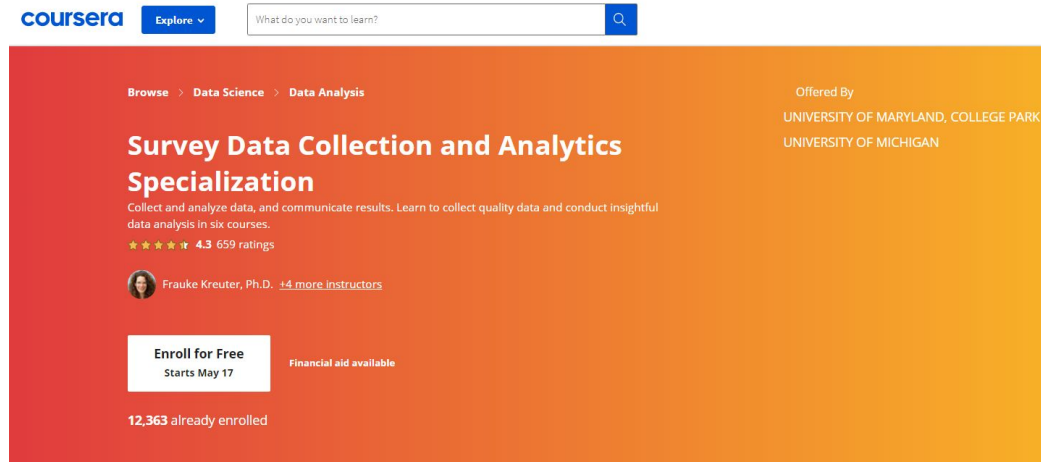
Competency screening; financial abuse; financial decision screening; information; item response theory; reliability

# Additional resources for conducting survey research

Lumen Learning [Course](#)

Open textbook: [Practical strategies for psychological measurement](#)

Coursera Survey Research [Courses](#)



The screenshot shows the Coursera website interface. At the top, there is a navigation bar with the Coursera logo, an 'Explore' button, and a search bar containing the text 'What do you want to learn?'. Below the navigation bar, the main content area has a red-to-orange gradient background. On the left, there is a breadcrumb trail: 'Browse > Data Science > Data Analysis'. The main heading is 'Survey Data Collection and Analytics Specialization'. Below the heading, there is a description: 'Collect and analyze data, and communicate results. Learn to collect quality data and conduct insightful data analysis in six courses.' followed by a rating of 4.3 stars from 659 ratings. Below the rating, there is a profile picture of Frauke Kreuter, Ph.D., and a link to '+4 more instructors'. At the bottom left, there is a white button that says 'Enroll for Free' and 'Starts May 17'. To the right of the button, it says 'Financial aid available'. At the bottom, it says '12,363 already enrolled'. On the right side of the main content area, there is a section titled 'Offered By' with the text 'UNIVERSITY OF MARYLAND, COLLEGE PARK' and 'UNIVERSITY OF MICHIGAN'.

coursera Explore

What do you want to learn?

Browse > Data Science > Data Analysis

**Survey Data Collection and Analytics Specialization**

Collect and analyze data, and communicate results. Learn to collect quality data and conduct insightful data analysis in six courses.

★★★★★ 4.3 659 ratings

Frauke Kreuter, Ph.D. +4 more instructors

Enroll for Free  
Starts May 17

Financial aid available

12,363 already enrolled

Offered By  
UNIVERSITY OF MARYLAND, COLLEGE PARK  
UNIVERSITY OF MICHIGAN



# Summary

- Social/behavioral health research lacks the equivalent of a *thermometer*, that is, a tool that accurately measures psychosocial factors regardless of context - whether that context is gender, race/ethnicity, age, SES, etc.
- Instead, the **psychometric properties of scales** that purport to assess psychosocial factors must be **continuously evaluated**, within and across populations, across language, across place and across time.
- Doing so is a multi-step, multi-modal process.

# Post-script:

## Even thermometers have measurement issues

- **The Metabolic Equivalent of a Task (MET) Rate - the amount of energy you “burn”/heat you generate doing a specific activity.**
  - 1 MET=the amount of oxygen consumed while sitting at rest and is equal to 3.5 ml O<sub>2</sub> per kg body weight x min.
  - **Reference group for the MET? Men who are ~150lbs and 40 year old.**
- **Why does this matter? Because it has shape our environments.**
  - Engineers have normed “Ideal” office temperature ranges predominantly on the preferences of men, who typically prefer colder spaces than women because of their higher MET rate and because their clothing tends to be more insulating vs. women’s clothing.

ANNALS OF TECHNOLOGY

## IS YOUR THERMOSTAT SEXIST?

By Anthony Lydgate

August 3, 2015



Since the nineteen-sixties, the science of thermal comfort has tended to neglect women. Illustration by Tomi Uemura