Advances in Mental Health Measurement

Forum on Mental Health and Substance Use Disorders' Public Workshop

Robert D. Gibbons, Ph.D.
Blum-Riese Professor of Biostatistics
The University of Chicago

October 16

Dr. Gibbons is a founder of Adaptive Testing Technologies, which distributes the CAT-MH™ battery of adaptive tests. The terms of this arrangement have been reviewed and approved by the University of Chicago in accordance with its conflict of interest policies.
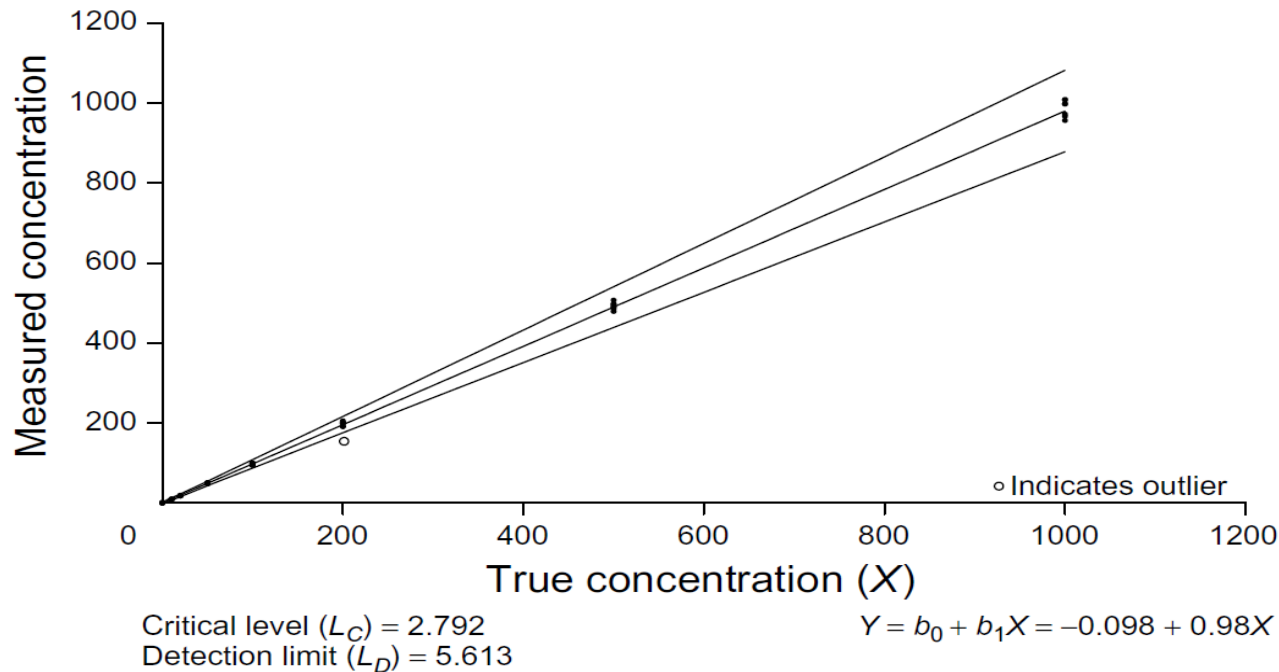
What is Measurement?

Measurement is the process of obtaining the magnitude of a quantity relative to an agreed standard.

" *Not everything that can be counted counts, and not everything that counts can be counted.*
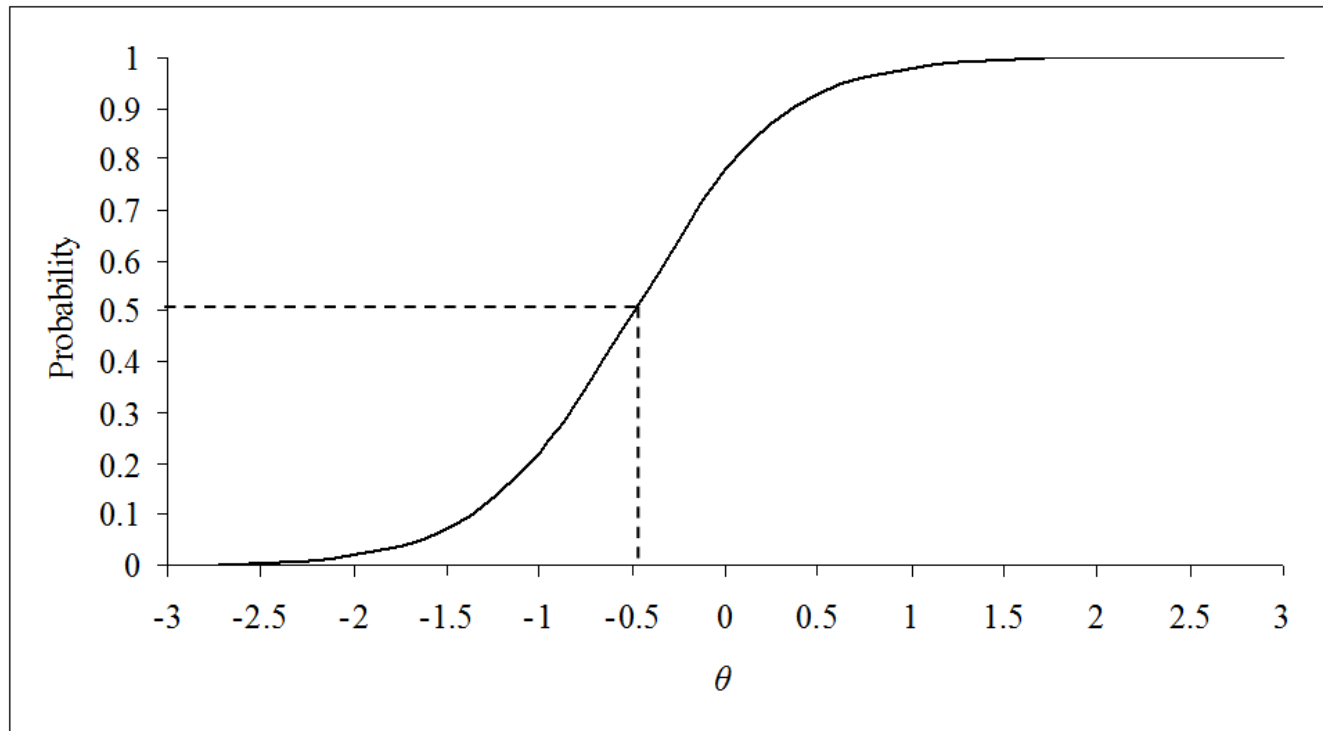
- Albert Einstein

Measurement in the Physical Sciences

Silver in parts per trillion (1 second in 32,000 years)

Critical level (L_C) = 2.792
Detection limit (L_D) = 5.613

$Y = b_0 + b_1 X = -0.098 + 0.98X$

$$\hat{Y}_X \pm s_{Y \cdot X} \left\{ (2F_{2,n-2}^{1-\alpha/2})^{1/2} \left[\frac{1}{n} + \frac{x_i^2}{\sum_{i=1}^{n} x_i^2} \right]^{1/2} + \Phi(P) \left(\frac{n-2}{\alpha/2 \chi_{n-2}^2} \right)^{1/2} \right\}$$

Measurement in the Social Sciences

Classical vs. IRT Measurement

Classical Measurement Model

Classical vs. IRT Measurement

IRT

What is CAT?

Arithmetic   Algebra   Calculus

Imagine a 1000-Item Math Test

Primary

Secondary

The Bifactor Model

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ \alpha_{31} & 0 & \alpha_{33} \\ \alpha_{41} & 0 & \alpha_{43} \end{bmatrix}$$

Preserves multidimensionality

Thresholds

$$P = \int_{-\infty}^{\infty} \left\{ \prod_{v=2}^{d} \int_{-\infty}^{\infty} \left[\prod_{j=1}^{n} \left(\Phi \left[\frac{\gamma_j - \alpha_{j1}\theta_1 - \alpha_{jv}\theta_v}{\sqrt{1 - \alpha_{j1}^2 - \alpha_{jv}^2}} \right] \right)^{u_{jv}} \right] g(\theta_v) d\theta_v \right\} g(\theta_1) d\theta_1 \,,$$

Severity

Discrimination

$$\hat{\theta}_{1i} = E(\theta_{1i} \mid \mathbf{u}_i, \theta_{2i} \ldots \theta_{di}) = \frac{1}{P_i} \int_{\theta_1} \theta_{1i} \left\{ \prod_{v=2}^{d} \int_{\theta_v} L_{iv}(\theta_v^*) g(\theta_v) d\theta_v \right\} g(\theta_1) d\theta_1 \,.$$

$$V(\theta_{1i} \mid \mathbf{u}_i, \theta_{2i} \ldots \theta_{di}) = \frac{1}{P_i} \int_{\theta_1} (\theta_{1i} - \hat{\theta}_{1i})^2 \left\{ \prod_{v=2}^{d} \int_{\theta_v} L_{iv}(\theta_v^*) g(\theta_v) d\theta_v \right\} g(\theta_1) d\theta_1 \,.$$

Uncertainty in severity estimate

Gibbons and Hedeker, 1992, *Psychometrika*
Gibbons et.al., 2007, *Applied Psychological Measurement*

THE UNIVERSITY OF
CHICAGO

What is the CAT-MH?

- The CAT-MH is a suite of 10 computerized adaptive tests (CAT) based on a multidimensional item response theory (MIRT).

- This means that we can, for example, extract the information from 400 depression symptom-items using an average of 10 adaptively administered items, yet maintain a correlation of r>=0.95 with the 400 item test score.

- As such we can dramatically increase precision while eliminating clinician burden and minimize subject burden.

- The K-CAT extends this technology to children.

What are the Advantages of Adaptive Testing?

- Adaptive tests provide constant precision of measurement throughout the entire severity continuum for any disorder that we measure.

- Items are targeted to a patient's specific level of severity at that point in time.

- Adaptive tests ask different questions upon repeat administration eliminating response bias produced by repeatedly asking the same questions over and over again.

- The CAT-MH has been validated against structured clinical diagnostic interviews (e.g. SCID-DSM-5) so it can also provide diagnostic profiles for a large number of disorders and suicide risk.

- The CAT-MH can be used for both screening and measurement.

What are the Advantages of Adaptive Testing?

- Adaptive tests are ideal for longitudinal assessments essential for measurement-based care.

- Adaptive tests not only estimate severity of a disorder (e.g. depression) they also estimate the uncertainty in that severity score and measure all patients to the same level of precision.

- The CAT-MH is cloud-based and scalable to any size population via a HIPAA secure AWS platform. Patients can be screened, measured and monitored in or out of the clinic.

- The CAT-MH can be fully integrated into the electronic health record (e.g. Epic), and we have developed clinical workflows for integrated behavioral health and primary care practices and clinics.
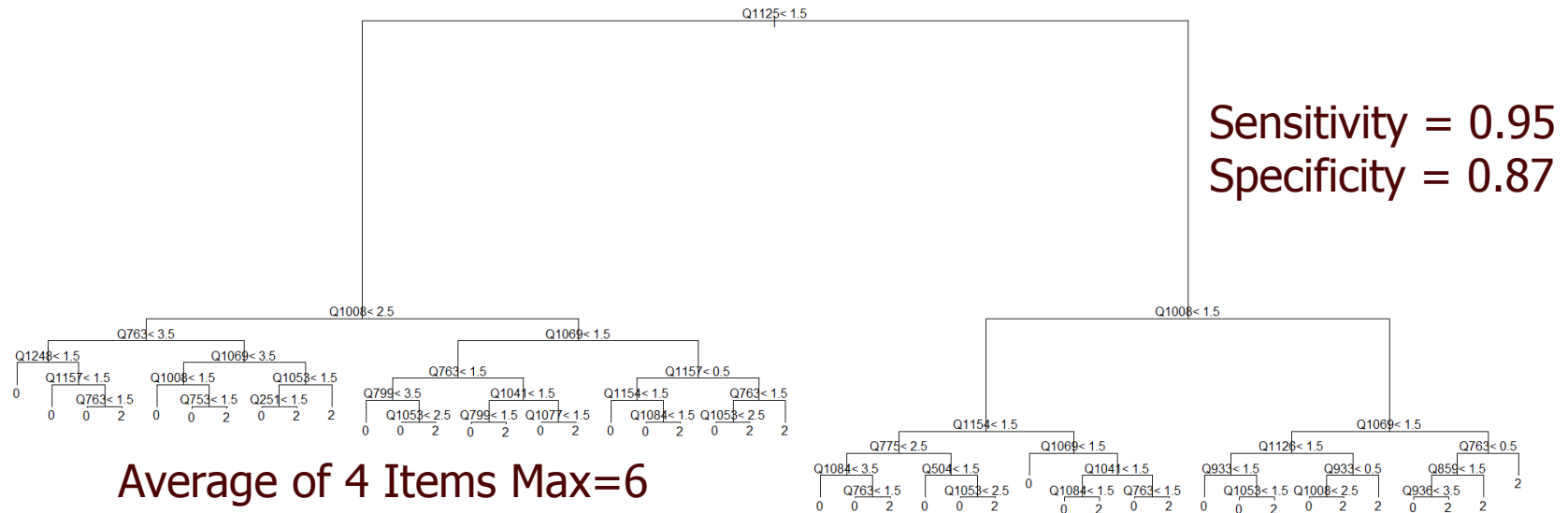
How Do Adaptive Tests Work?

- We administer a question with medium severity.

- We estimate severity based on the response to the question (symptom).

- We select the next most informative question out of the remaining symptoms-item questions.

- We stop when we reach the desired precision of measurement (e.g. 5 points on a 100 point scale).

Diagnosis and Measurement are Fundamentally Different Things
CAD-MDD – Decision Tree

Sensitivity = 0.95
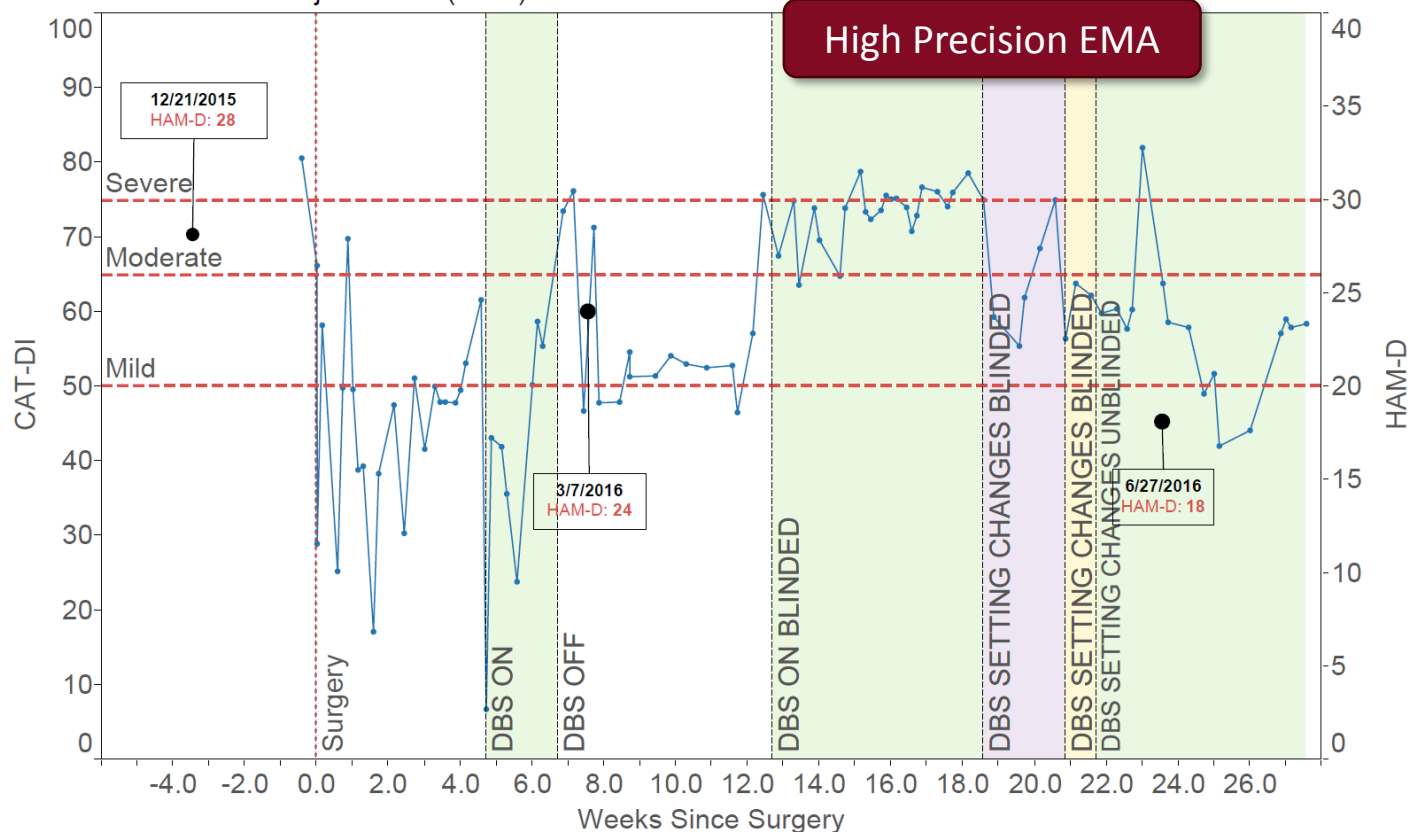Specificity = 0.87

Average of 4 Items Max=6

Rates of Detection and Service Utilization

- Emergency Department U of Chicago (n=1000)
  - 22% MDD positive screens (>90% confidence)
  - 7% MDD Positive + moderate or severe depression
  - 3% suicide screen positive
  - **3-fold increase in ED visits**
  - **4-fold increase in hospitalizations**
  - None of these patients had a psychiatric indication

- Spain and US Latino Samples (n=1000)
  - 25% MDD positive screens (>90% confidence)
  - 9% MDD Positive + moderate or severe depression

What if you could assess treatment response on a daily basis?

Daily Home Monitoring of a Deep Brain Stimulation Patient for 6 months

What Can the CAT-MH and K-CAT Measure?

- Adult (English and Spanish)

  - **Depression***

  - **Anxiety***

  - **Mania/Hypomania***

  - **Suicidality***

  - **PTSD***

  - **Substance Abuse***

  - **Psychosis***

  - **Functional Impairment**

  - **Quality of Life**

  - **Functional status and well being (Cancer Patients and Survivors)**

  - Borderline Personality Disorder

- Perinatal (English and Spanish)

  - **Depression***

  - **Anxiety***

  - **Mania/Hypomania***

- Child and adolescent (child and parent ratings) Ages 7-17

  - **Depression***

  - **Anxiety***

  - **Mania/Hypomania***

  - **ADHD***

  - **Conduct disorder***

  - **Oppositional defiant disorder***

  - **Suicidality***

  - Autism

* Completed, validated and in use
Completed, being validated and/or integrated into the CAT-MH

Example Applications

- **UCLA Grand Challenge**
  - Screen all undergraduates at UCLA and triage to iCBT
  - Screen 1.8 million to develop a Registry of 100,000 patients

- **Indiana University Grand Challenge**
  - Survey the state of Indiana for substance abuse disorder, screening, monitoring and treatment for SUD

- **University of Chicago**
  - Emergency Medicine – Depression and suicide risk
  - Integrated Primary and Behavioral Health Care

- **Rush University Medical Center**
  - Orthopedic Surgery – Does depression lead to poor outcomes?

- **UK - National Health Service**
  - Pilot Study for National roll-out

- **NorthShore University Health Systems**
  - Perinatal depression screening and follow-up

- **Cook County Corrections Department**
  - Screen all inmates in Bond Court, the Cook County Jail and Probation

- **Veteran's Administration/Department of Defense**
  - Develop new PTSD scale and further validate suicidality scale

- **State of Tennessee**
  - Foster Care, Juvenile Justice, Detention Centers – 300 case workers

- **SAMHSA/RTI**
  - Large-Scale National Survey: Mental and Substance Use Disorder Prevalence Study (MDPS) – US Adult population

- **India/SCARF**
  - Large-scale screening and measurement-based care

We spend billions on biological measurements, yet we validate them using stone-age clinical measurements.

" *Statistics ... is the most important science in the whole world, for upon it depends the practical application of every other (science) and of every art.*

- Florence Nightingale

Scientific Literature

1. Gibbons R.D., & Hedeker D.R. Full-information item bi-factor analysis. *Psychometrika*, *57*, 423-436, 1992.
2. Gibbons R.D., Bock R.D., Hedeker D., Weiss D., Segawa E., Bhaumik D.K., Kupfer D., Frank E., Grochocinski V., Stover A. Full-Information Item Bi-Factor Analysis of Graded Response Data. *Applied Psychological Measurement*, *31*, 4-19, 2007.
3. Gibbons R.D., Weiss D.J., Kupfer D.J., Frank E., Fagiolini A., Grochocinski V.J., Bhaumik D.K., Stover A. Bock R.D., Immekus J.C. Using computerized adaptive testing to reduce the burden of mental health assessment. *Psychiatric Services, 59*, 361-368, 2008.
4. Gibbons R.D., Weiss D.J., Pilkonis P.A., Frank E., Moore T., Kim J.B., Kupfer D.K. The CAT-DI: A computerized adaptive test for depression. *Archives of General Psychiatry*, 69, 1104-1112, 2012.
5. Gibbons R.D., Hooker G., Finkelman M.D., Weiss D.J., Pilkonis P.A., Frank E., Moore T., Kupfer D.J. The CAD-MDD: A computerized adaptive diagnostic screening tool for depression. *Journal of Clinical Psychiatry*, 74, 669-674, 2013.
6. Gibbons R.D., Weiss D.J., Pilkonis, P.A., Frank E., Moore T., Kim J.B., Kupfer D.J. Development of the CAT-ANX: A computerized adaptive test for anxiety. *American Journal of Psychiatry*, 171, 187-194, 2014.
7. Gibbons R.D., Bi-factor Analysis. Encyclopedia of Quality of Life Research, Springer, Netherlands, 386-394, 2014.
8. Achtyes E.D., Halstead S., Smart L., Moore T., Frank E., Kupfer D., Gibbons R.D. Validation of computerized adaptive testing in an outpatient non-academic setting. *Psychiatric Services,* 66, 1091-1096, 2015.
9. Beiser D., Vu, M., Gibbons, R.D. Test-retest reliability of a computerized adaptive depression test. *Psychiatric Services*, 67, 1039-1041, 2016.
10. Gibbons R.D., Computerized adaptive diagnosis and testing of mental health disorders. *Annual Review of Clinical Psychology*, 12, 83-104, 2016.
11. Kim J.J., Silver R.K., Elue R., Adams M.G., La Porte L.M., Cai L., Kim J.B., Gibbons R.D. The experience of depression, anxiety and mania among perinatal women. *Archives of Women's Mental Health*, 19, 94-100, 2017.
12. Gibbons R.D., Kupfer D., Frank E. Moore T., Boudreaux E. Development of a computerized adaptive suicide scale., *Journal of Clinical Psychiatry*, 78, 1376-1382, 2017.
13. Gibbons R.D., Beiser D., Boudreaux E., Kupfer DJ. Einstein, measurement and prediction. *Journal of Affective Disorders*, 2017, published on-line ahead of print.
14. Sani S., Busnello J., Kochanski R., Cohen Y., Gibbons R.D. High frequency measurement of depressive severity in a patient treated for severe treatment resistant depression with deep brain stimulation. *Translational Psychiatry*, 7, e1207, 2017.
15. Aschebrook-Kilfoy, B., Ferguson, B.A., Angelos, P., Kaplan, E.L., Grogan, R.H., Gibbons, R.D. Development of the ThyCAT: A clinically useful computerized adaptive test to assess quality of life in thyroid cancer survivors. *Surgery*, 163, 137-142, 2017.
16. Gibbons. R.D., Alegria, M., Cai, L., Herrera, L., Markle, S.L., Collazo, F. Garcia, E.B. Successful validation of the CAT-MH scales in a sample of Latin American migrants in the U.S. and Spain. *Psychological Assessme*nts, 2019, published on-line ahead of print.
17. Graham A.K., Minc A., Staab E., Beiser D.G., Gibbons R.D., Laiteerapong N. Validation of a computerized adaptive test for mental health in primary care. *Annals of Family Medicine*, 17, 23-20, 2019.
18. Beiser, D.J. Ward, C.E., Vu, M. Laiteerapong, N., Gibbons, R.D. Depression in Emergency Department Patients and Association with Healthcare Utilization. *Academic Emergency Medicine*, 2019, published on-line ahead of print.
19. Gibbons R.D., Smith J.D., Brown C.H., Sajdak M., Tapia N., Kulik A., Epperson M.W., Csernansky J. Improving the evaluation of adult mental health disorders in the criminal justice system using computerized adaptive testing. *Psychiatric Ser*vices. Published on-line ahead of print.
20. Gibbons R.D., deGruy F.V. Without wasting a word: Extreme improvements in efficiency and accuracy using computerized adaptive testing for mental health disorders (CAT-MH). *Current Psychiatry Reports*. 21, 1053-9, 2019.
21. LaPorte L.M., Lim J.J., Adams M.G., Zagorsky B.M., Gibbons R.D., Silver R.K. Feasibility of perinatal mood screening and text messaging on patients' personal smartphones. *Archives of Women's Health*. 2019, Published on-line ahead of print.
22. Gibbons R.D., Kupfer D., Frank E., Lahey B.B., George-Milford B., Biernesser C., Porta G., Moore T., Kim J.B., Brent D. Computerized adaptive tests for rapid and accurate assessment of psychopathology dimensions in youth. *Journal of the American Academy of Child and Adolescent Psychiatry*, in press.

THE UNIVERSITY OF
CHICAGO

More details on the science and access to the CAT-MH available at

www.adaptivetestingtechnologies.com