



THE UNIVERSITY OF
CHICAGO
BIOLOGICAL
SCIENCES

The CAT-MH Project
Development of Computerized Adaptive Tests
(CAT) for Mental Health Measurement

Robert D. Gibbons
University of Chicago

David J. Kupfer
David J. Weiss
Paul Pilkonis

Ellen Frank
R. Darrell Bock

Supported by NIMH Grant R01-MH-66302. The CAT-MH is distributed through Adaptive Testing Technologies (www.adaptive-testingtechnologies.com) for which Drs. Gibbons, Kupfer, Frank, Weiss and Pilkonis have financial interests.

Classical vs. IRT Measurement

Classical Measurement Model

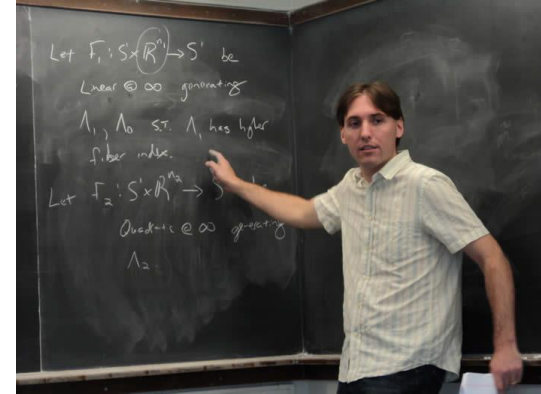


Classical vs. IRT Measurement

IRT



What is CAT?



Arithmetic

Algebra

Calculus

Imagine a 1000 Item Math Test



Bi-Factor IRT Model

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

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

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

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



Paradigm Shift

- Traditional Measurement
 - Fix items allow precision to vary
- IRT-Based CAT
 - Fix precision allow items to vary
- Change precision depending on application
 - Epidemiology – fewer items lower precision ($se=.4$)
 - Primary care screening – medium precision ($se=.3$)
 - RCTs – more items high precision ($se=.2$)
- Suicide Screen – C-SSRS items (2-4)



CAT-DI RESULTS

Results of Simulated CAT (N=308)

Test	SE Term	θ Correlation	Mean No. Items	Min No. Items	Max No. Items
DEP	.30	.95	12	7	22
	.40	.92	6	4	16
ANX	.32	.94	12	6	24
BP	.45	.91	12	6	24

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.

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.

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*, in press.



CAD-MDD – Results

Sensitivity: 0.95

Specificity: 0.87

Average of 4 Items Max=6

Gibbons et.al., J. of Clinical Psychiatry (2013)



Independent Validation Study

- Highly Comorbid Community MH sample (n=150)
 - High sensitivity 0.96 maintained for entire sample
 - Specificity 1.0 for MDD vs Control
 - CAT-DI, CAT-ANX, CAT-BP all predict Dx
 - MDD 28-fold across scale
 - GAD and current BP each 12-fold across scale
- 97% accurately reflected mood
- 86% preferred computer interface (10% preferred pp)
- 97% Comfortable taking CAT-MH
- 98% Answered honestly

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*, in press.



Rates of Detection and Service Utilization

- Emergency Department U of Chicago (n=1000)
 - 26% MDD positive screens (>50% confidence)
 - 22% MDD positive screens (>90% confidence)
 - 7% MDD Positive + moderate or severe CAT-DI
 - 3% suicide screen positive
 - 3-fold increase in ED visits in past year moderate/severe vs. none/mild
 - 4-fold increase in hospitalizations in past year moderate/severe vs. none/mild
 - None of these patients had a psychiatric indication
- Primary Care Spain and US Latino Samples (n=1000)
 - 33% MDD positive screens (>50% confidence)
 - 25% MDD positive screens (>90% confidence)
 - 9% MDD Positive + moderate or severe CAT-DI



Future Directions

- Screening and monitoring in primary care
- Inexpensive phenotyping for GWAS studies
- Psychiatric epidemiology
- Comparative effectiveness and safety
- Differential Item Functioning – Global Health
- Kiddie CAT - Developmental shifts – vertical scaling
- Spend billions on biology but validate using stone age clinical measurements
- Autism, PTSD, RDoC, ...
- Military – Suicide $RR=4$ within 4 years of discharge
- Cloud computing environments

