The CAT-MH Project:
Development of Comprehensive CAT-based Instruments for Measuring Mental Health

Robert D. Gibbons
David J. Kupfer          Ellen Frank
David J. Weiss           R. Darrell Bock
Paul Pilkonis

Supported by NIMH Grants R01-MH-66302 and R01-MH-100155-01. The CAT-ANX will be made available for routine administration (by the end of 2014) 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 Test Theory    Item Response Theory
What is CAT?

Imagine a 1000 Item Math Test
Bi-Factor IRT Model

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

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

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

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

Gibbons and Hedeker, 1992, Psychometrika; Gibbons et al., 2007, Applied Psychological Measurement
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=0.4)
  – Primary care screening – medium precision (se=0.3)
  – RCTs – more items high precision (se=0.2)

• Create large item banks rather than short tests
Results

1008 item bank (Depression, Anxiety, Mania)

Depression – 12 items – r=0.95 with 389 item score
Anxiety – 12 items – r=0.94 with 431 item score
Mania – 12 items – r=0.91 with 88 item score

MDD Screener – 4 items – sensitivity 95% spec 87%
2-4 item C-SSRS Suicide Screener
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
Future Directions

- Screening and monitoring in primary care
- Inexpensive phenotyping for GWAS studies
- Psychiatric epidemiology
- 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


