ADAPTIVE TESTING TECHNOLOGIES

UNRIVALED MEASUREMENT

VALIDATED • PRECISE • ACCURATE

The CAT-MH®

A fundamental scientific breakthrough in psychopathology screening and measurement

For additional information, please visit our website at www.adaptivetestingtechnologies.com, send an email to info@adaptivetestingtechnologies.com, or call us at (312) 878-6490 ext. 303

Adaptive Testing Technologies was founded by Professor Robert D. Gibbons, Blum-Riese Professor of Biostatistics at the University of Chicago, and Dr. David Kupfer, former Chair of the Department of Psychiatry at the University of Pittsburgh and Chair of the DSM-5 committee for the American Psychiatric Association, and is led by its C.E.O. Yehuda Cohen. Our technology is based on over a decade of continuous National Institute of Mental Health-funded research and the company provides an advanced and intuitive user interface, technical support, system integration and distribution of these tools. Adaptive Testing Technologies ("ATT") has grown to be the leader in the design, testing, and implementation of large scale computerized adaptive testing (CAT) systems for complex traits, such as mental health assessment. ATT's rollout of the CAT-DI (Computerized Adaptive Test - Depression Inventory) was published in the November 2012 issue of JAMA Psychiatry[1]. It was the first and remains the sole CAT based on multidimensional item response theory (MIRT)[2,3]. In addition to the CAT for depression, the CAT-Mental Health or CAT-MH® now consists of a complete suite of validated measures for diagnostic screening and continuous severity measurement of mental health and substance use disorders and suicidality[4]. We have developed and validated a parallel set of measures for youth[5], the K-CAT®.

Mental Health Measurement has been based primarily on subjective judgment and classical test theory. Typically, impairment level is determined by a total score, requiring that all respondents be administered the same items. This traditional approach weighs equally the responses to items such as "Do you feel sad?" and "Do you feel like everyone would be better off if you were dead?" Clearly, this approach is far from optimal. An alternative to full-scale administration is adaptive testing in which different individuals may receive different symptom items that are targeted to their specific impairment level.

How does the CAT-MH® work?
The CAT-MH® is based on multidimensional item response theory (MIRT). Within CAT, an individuals' initial item responses are used to determine a provisional estimate of her standing on the measured trait to be used for subsequent item selection. Based on MIRT procedures, estimates of items (e.g., difficulty, discrimination) and individuals (e.g., severity of depression) can be obtained to more efficiently identify suitable item subsets for each individual. MIRT weighs more severe items more heavily than less severe items in deriving a test score. MIRT also provides an estimate of uncertainty that can be used to assess the significance of change for an individual. The CAT-MH® adaptively selects a small optimal set of items from a large bank of approximately 1,500 items and the K-CAT® has a separate item bank of 2,120 items. Instead of fixing the items and allowing the precision of measurement to vary, we fix the precision of measurement and allow the items to vary. The result is a dramatic increase in the precision of measurement with a dramatic decrease in the burden of measurement for the patient and the complete elimination of clinician burden.

What CAT-MH® modules are available?
The CAT-MH® is a suite of adult CATs for depression[1], anxiety[6], mania/hypomania[7], suicidality[8], PTSD[9], substance use disorder[10], psychosis[11], and social determinants of health (SDoH). These measures have been optimized for perinatal[12], criminal justice[13], behavioral health [1,4,6–11], primary care[14,15], emergency medicine[16], LGBTQ[17], and Latinx[18] populations. For children and adolescents ages 7 to 17, the K-CAT® has been validated for the measurement of depression, anxiety, mania/hypomania, ADHD, oppositional defiant disorder, conduct disorder, substance use disorder and suicidality, as rated by the child and the parent or caregiver[5]. We have also carefully translated (forward, reverse and adjudicated) our entire adult item bank into Spanish and Chinese and have all our CAT-MH® modules available in all three of these languages[17]. The K-CAT® is available in English and Spanish. Both the CAT-MH® and the K-CAT® have built-in selectable audio to aid in literacy issues and further ease of use. They can be administered on any internet-capable device (smartphone, tablet, notebook, and computers) or questions adaptively administered by telephone.

What are the advantages of the CAT-MH®?

We can dramatically increase precision while eliminating clinician burden and minimizing subject burden. The CAT-MH® provides constant precision of measurement throughout the entire severity continuum for any disorder that we measure. The adaptive nature of the CAT-MH® targets a patient's specific level of severity at that point in time. 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. Because the same questions are not repeatedly administered, the CAT-MH® is ideal for longitudinal assessments essential for measurement-based care[19]. The CAT-MH® is cloud-based and scalable to any size population via a HIPAA secure Amazon Web Services (AWS) platform, meaning patients can be screened, measured and monitored in or out of the clinic. The CAT-MH® and K-CAT® are the only adaptive measurement systems that preserve the multidimensionality of complex mental health constructs and the only suites of mental health measures that have been fully validated against structured clinical diagnostic interviews.

What are some of the applications of the CAT-MH®?

Large-scale screening and monitoring of depression and anxiety in integrated primary care and behavioral health care settings is a natural application of this work, and is already in practice in several major institutions[14/21]. Insurers can now monitor the progress of patients through treatment without the patient needing to be in the clinic for testing. High-frequency monitoring, even hourly in response to fast-acting new psychiatric treatments (e.g. ketamine and deep brain stimulation[19]) are now possible because the same items are not repeatedly administered to the same patient. The CAT-MH® is also uniquely suited to detect falsification of responses that may be made to either give the impression of an illness (e.g. in a jail[13]) or to mask the presence of an illness (e.g. in the military). In collaboration with UCLA's Semel Institute, we have developed a large-scale university student screening and monitoring system (STAND) to screen, triage and provide measurement-based care to college students remotely in real time and to determine which students need more intensive treatment. As a part of the UCLA Depression Grand Challenge, the CAT-MH® will be used to screen 1.8 million people in the Los Angeles area to develop a cohort of 100,000 people with depression who will be followed for a decade. The CAT-MH® perinatal tests have been used to develop an early warning system for the onset of perinatal depression using remote weekly monitoring during pregnancy and post-partum[20]. The CAT-MH® has been used for state-wide assessment of mental health disorders in the State of Tennessee's child welfare system, where over 300 caseworkers have been trained in its use. There are numerous applications of the CAT-MH® in the criminal justice system, including bond court, probation, and the Cook County Jail[13]. Indiana University and the National Opinion Research Center are using the CAT-SUD in a state-wide survey of substance use disorder. Working with the Veterans administration we have developed the most advanced PTSD[9] measurement system and suicide prediction system[8] which will be used in a 100,000 veteran patient cohort. The CAT-MH® is being used a s a 1st stage diagnostic screener in the largest national survey of mental health and substance use disorder survey in history, conducted by SAMHSA (Substance Abuse and Mental Health Services Administration Division of HHS) and implemented by RTI. These are just a handful of the current and potential applications of this technology.

Electronic Health Record Integration

We have fully integrated the CAT-MH® into the Epic EHR at the University of Chicago (and being ported to several other institutions), and developed clinical workflows designed around the CAT-MH® integration[14]. This enables screening and measurement in the clinic using computers, and remote screening and assessment via the MyChart patient portal, with results immediately displayed in the EHR. Using this technology, we have shown that in primary care traditional fixed length screeners such as the PHQ-2/PHQ-9 grossly underestimate the rate of depression at 1%, whereas the CAT-MH® obtains the expected rate of 7%[21].

Validation

Unlike most traditional mental health measures, the CAT-MH® and K-CAT® have been fully validated against lengthy structured clinical interviews (e.g. SCID for DSM-5) and extant clinician-rated measures. In the following, we briefly describe the results of these validations and properties of the different CAT scales for adults and youth.

Depression: The CAT-Depression Inventory[1] (CAT-DI) reproduces the information in a final bank of 389 items using adaptive administration of 12 items in approximately 2 minutes while maintaining a correlation of 0.95 with the total bank score. In terms of convergent validity, correlations were r=0.81 with the PHQ-9, r=0.75 with the HAM-D, and r=0.84 with the CES-D. In general, the distribution of scores between the diagnostic categories showed greater overlap (i.e., less diagnostic specificity), greater variability, and greater skewness for these other scales relative to the CAT-MH®. In terms of diagnostic validity, the CAD-MDD (our brief adaptive diagnostic screener for MDD) has a sensitivity (Se) of 0.95 and specificity (Sp) of 0.87 for an hour-long SCID DSM-5 diagnosis of MDD22. A thresholded CAT-DI yielded Se of 0.92 and Sp of 0.88 for a SCID DSM-5 diagnosis of MDD[1].

Anxiety: The CAT-ANX[6] reproduces the information in a bank of 437 items using adaptive administration of 12 items in approximately 2 minutes while maintaining a correlation of r=0.94 with the total bank score. Relative to a SCID DSM-5 diagnosis of generalized anxiety disorder (GAD), a thresholded CAT-ANX produced Se of 0.86 and Sp of 0.86.

Suicidality: The CAT-SS[8] reproduces the information in a 111 item bank that provides a crosswalk between symptoms of depression, anxiety and suicidality, using an average of 10 items in less than 2 minutes, while maintaining a correlation of r=0.96 with the total 111 item bank score. The CAT-SS was validated against a 20 minute Columbia scale structured clinical interview and demonstrated Se of 1.0 and Sp of 0.95 for ideation or worse (kappa 0.81); active ideation Se=1.0, Sp=0.95; suicide alert Se=1.0, Sp=0.89; and lifetime attempt Se=0.58 and Sp=0.88 for the CAT-SS high risk versus low-risk groups.

Substance Use Disorder: The CAT-SUD[10] reproduces the information in a bank of 168 items using adaptive administration of 11 items in 2 minutes with a total item bank correlation of r=0.91. The CAT-SUD was validated against the CIDI structured clinical interview diagnosis of SUD and demonstrated an AUC=0.85 across the range of the scale.

Mania/hypomania: The CAT-MANIA[7] reproduces the information in the 89 item bank using adaptive administration of 17 items in 3.4 minutes. This has been reduced to approximately 12 items in 2 minutes through refined CAT termination criteria. The CAT-MANIA was validated against the SCID DSM-5 structured clinical interview for bipolar disorder (BP, BP I, BP II). Across the range of the scale, the probability of a positive diagnosis of bipolar disorder increased 12-fold.

PTSD: The CAT-PTSD measure is an adaptive version of the 20 item PCL-5. The CAT-PTSD can be adaptively administered using an average of 6 items in approximately 1 minute, while maintaining a correlation of r=0.97 with the 20 item fixed-length PCL-5. Since the PCL-5 is a validated instrument and given the almost perfect correlation between the CAT-PTSD and the PCL-5, the CAT-PTSD is also a valid instrument. Working with the Veterans Administration (VA), we have created a new CAT-PTSD based on a 211 item expanded item bank[9]. Adaptive administration reduces the average number of items administered to 10 items, yet maintains a correlation of 0.95 with the 211 item score. The new CAT-PTSD is correlated r=0.88 with the PCL-5 and was also validated against the SCID DSM-5 PTSD diagnosis with AUC=0.85. An adaptive 6-item CAD-PTSD diagnostic screener was also developed[9] which reproduces the DSM-5 PTSD diagnosis with AUC=0.91.

Psychosis: The CAT-Psychosis measures are adaptive clinician-administered and patient self-report psychosis measures[11]. The measures extract the information from the 144 item bank in an average of 12 items, maintaining a correlation of r=0.92 with the total bank score in 5 minutes for the clinician guided interview and 1 minute and 20 seconds for the patient self-report. The measures were validated against the SCID DSM-5 psychosis diagnosis with AUC=0.97 for the clinician ratings and AUC=0.85 for the patient self-report. Test-retest reliability was r=0.86 for the clinician and r=0.82 for the patient self-report. For the clinician-administered version, inter-rater reliability was ICC=0.73.

The K-CAT® is a parallel tool that extends the adult technology contained in the CAT-MH® to children and adolescents ages 7-17, based on parent/caregiver and youth self-reports[5]. It measures depression, anxiety, mania/hypomania, ADHD, oppositional defiant disorder, conduct disorder, and suicidality. The total item bank contains 2120 items distributed across parent and child ratings. It can be adaptively administered in 5 minutes for the parent/caregiver and 7 minutes for the child. It has been validated against the KSADS structured clinical diagnostic interview. In terms of validity, it tracks KSADS diagnoses for each of the above disorders with AUCs ranging from 0.83 for GAD, to 0.92 for MDD, and suicidal ideation or worse with an AUC=0.996. Strong correlations with extant clinical measures also demonstrated convergent validity (r>0.60). Test-retest reliability averaged r=0.80.

For more information, reference the K-CAT® Overview document.

Academic Press

1. 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. *JAMA Psychiatry*, *69(11)*, 1104-1112, 2012.
2. Gibbons R.D., & Hedeker D.R. Full-information item bi-factor analysis. *Psychometrika*, *57*, 423-436, 1992.
3. 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.
4. Gibbons R.D., Computerized adaptive diagnosis and testing of mental health disorders. *Annual Review of Clinical Psychology*, *12*, 83-104, 2016.
5. 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, 59(11),* 1264-1273, 2020.
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. 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.
8. Gibbons R.D., Kupfer D., Frank E., Moore T., Boudreaux E.D. Development of a computerized adaptive suicide scale. *Journal of Clinical Psychiatry, 78,* 1376-1382, 2017.
9. Brenner L., Betthauser L., Germain A., Chattopadhyay I., Frank E., Kupfer D.J., Gibbons R.D. Development and Validation of a Computerized Adaptive Assessment Tools for the Screening and Measurement of Post-Traumatic Stress Disorder. JAMA Network Open, *4(7)*, e2115707, 2021.
10. Gibbons RD, Alegria M, Markle S, Fuentes L, Zhang L, Carmona R, Collazos F, Wang Y, Baca-Garcia E. Development of a computerized adaptive substance use disorder scale – The CAT-SUD. *Addiction, 115(7),* 1382-1394, 2020.
11. Guinart D., de Filippis R., Rosson S., Patil B., Prizgint L., Talasazan N., Meltzer H., Kane J.M., Gibbons R.D. Development and validation of a computerized adaptive assessment tool for severity of psychotic symptoms and diagnostic prediction. *Schizophrenia Bulletin, 47(3)*, 644-652, 2021.
12. 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.
13. 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 Services, 70(11),* 1040-1043, 2019.
14. 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.
15. 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*, 1053-1059, 2019.
16. 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, 26(8)*, 878-888, 2019.
17. Mustanski B., Gibbons R.D. et.al. Predicting the transition from suicidal ideation to suicide attempt among sexual and gender minority youths. *Psychiatric Services, 72(11),* 1261-1267, 2021.
18. 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 Assessments, 30(10)*, 1267-1276, 2019.
19. 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.
20. 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, 23(2),* 181-188, 2020.
21. Laiteerapong N., Franco M, Knitter A., Deehan W., Keahey A. Moses J.S., Shah S., Gibbons R.D., Yohanna D., Beckman N., Vinci L., Staab E.M. Implementation of a Computerized Adaptive Test for Depression Screening in Primary Care. *Society of General Internal Medicine Abstract*, 2020.
22. 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.