Objective: To present the Mobile Toolbox (MTB), comprised of an expandable library of cognitive and other tests, including adapted versions of NIH Toolbox® measures. The MTB provides a complete research platform for app creation, study management, data collection, and data management. We will describe the MTB project and MTB research platform and demonstrate examples of assessments.
Participants and Methods: MTB is the product of an NIH-funded, multi-institutional effort involving Northwestern University, Sage Bionetworks, Penn State, University of California San Francisco, University of California San Diego, Emory University, and Washington University. The MTB assessment library is a dynamic repository built upon Sage Bionetworks mobile health platform. All MTB measures are created or adapted for a mobile interface using iOS and Android smartphones. Guided by the principles of open science, many components are open source to allow researchers and developers to integrate externally developed tests, including supplemental scales (e.g., passively collected contextual factors) assessing variables such as mood and fatigue that might influence cognitive test performance.
Results: The current MTB library includes eight core cognitive tests based on well-established neuropsychological measures: two language tasks (Spelling and Word Meaning), two executive functioning tasks (Arrow Matching and Shape-Color Sorting), an associative memory task (Faces and Names), an episodic memory task (Arranging Pictures), a working memory task (Sequences) and a processing speed task (Numbers and Symbols). Additional cognitive assessments from other popular test libraries including the International Cognitive Ability Resource (ICAR), Cognitive Neuroscience Test Reliability and Clinical Applications for Schizophrenia (CNTRACS) and Test My Brain are currently being implemented, as are non-cognitive measures from the NIH Toolbox Emotion Battery and the Patient-Reported Outcomes Measurement Information System (PROMIS). The MTB library includes measures suitable for use in research studies incorporating point-in-time and burst designs as well as ecological momentary assessment (EMA).
Conclusions: The MTB was created to address many of the scientific, practical, and technical challenges to cognitive assessment by capitalizing on advances in technology measurement and cognitive research. Initial psychometric evaluation of measures has been performed, and additional clinical validation is underway in studies with persons at risk for cognitive impairment or Alzheimer's disease (AD), diagnosed with mild cognitive impairment (MCI) or AD, Parkinson's disease, and HIV-associated Neurocognitive Disorders. Calculation of norms and reliable change indicators is in progress. The MTB is currently available to beta testers with public release planned for Summer, 2023. Clinical researchers will be able to use the MTB system to design smartphone-based test batteries, deploy and manage mobile data collection in their research studies, and aggregate and analyze results in the context of large-scale norming data.

Categories:
Assessment/Psychometrics/Methods (Adult)
Keyword 1: cognitive functioning
Keyword 2: assessment
Keyword 3: technology
Correspondence: Richard Gershon, Northwestern University Feinberg School of Medicine, gershon@northwestern.edu

2 Validity and Reliability of Mobile Toolbox Cognitive Assessments

Cindy J Nowinski[1], Aaron Kaat[1], Jerry Slotkin[2], Erika La Forte[1], Yusuke Shono[1], Miriam Novack[1], Sarah Pila[1], Elizabeth Dworak[1], Stephanie R Young[1], Zahra Hosseinian[1], Hubert Adam[1], Richard Gershon[1]
[1]Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [2]University of Delaware, Newark, Del, USA

Objective: To present validation evidence for the first eight cognitive measures available through Mobile Toolbox (MTB). These measures use a remote self-administered platform to assess language, working memory, episodic memory, executive function, and processing speed.
Participants and Methods: We used two separate samples, recruited as part of a larger study, to validate MTB measures. Sample I, comprised of 92 English-speaking adults ages 18-85, was used to assess internal consistency and construct validity. Participants were first administered "gold standard" cognitive measures (Wechsler Memory Scale-IV Verbal

https://doi.org/10.1017/S1355617723009724 Published online by Cambridge University Press

Paired Associates I and II; Wechsler Adult Intelligence Scale-IV Symbol Search, Digit Span, Coding, and Letter-Number Sequencing; Delis-Kaplan Executive Function System Color-Word Interference Test, Peabody Picture Vocabulary Test, Wechsler Individual Achievement Test-4 Spelling, and the Wisconsin Card Sorting Test), after which they completed MTB (pre-loaded on a study-provided smartphone) on their own. Internal consistency was evaluated using measure-appropriate indices (split-half reliability, Cronbach's alpha or IRT-based indices). Pearson correlation coefficients between MTB tests and measures of similar constructs were used to evaluate concurrent validity. For two tests with timing-dependent scores, Arrow Matching and Shape-Color Sorting, separate analyses were performed for iOS and Android devices. Sample II, with 1,120 English-speaking participants ages 18-90, was used to evaluate age-related change. Participants completed MTB measures remotely on their own smartphones, in a preset order, within a 14-day period. Spearman correlation coefficients, corrected for education, were calculated to evaluate relationships between age and test scores.
Results: Sample I participants were 67% female, 52% white, 99% non-Hispanic; average age=48 (SD= 17). Education was: < high school (1%); high school (55%); some college (21%); college (15%); graduate degree (8%). Internal consistency estimates ranged from 0.81 to 0.99. Pearson correlations between MTB and external measures ranged from 0.41 to 0.86 (all $p < .01$). Of the timed tests, only Shape-Color sorting showed significant score differences between Android and iOS devices.
Sample II was 57% female, 13% Hispanic, 72% white, mean age = 45 (SD = 21). Education distribution was: < high school (2%); high school (34%); some college (34%), college (20%); graduate degree (11%). Measures of executive function (r = -0.50; r=-0.57) and processing speed (r= -0.61) showed the expected negative correlation with age (all $p <0.001$). Negative correlations, although weaker, were also seen on measures of working memory (r=-0.2) and episodic memory (r=-0.2, r=-0.37; p.<.001). Vocabulary performance improved with age (r=0.4; p<.001), while spelling scores remained stable (r=0.09).
Conclusions: Initial studies support the validity and reliability of the first eight MTB cognitive measures in two diverse samples. MTB tests showed satisfactory construct validity, as demonstrated by the associations between MTB and well-established tests. Furthermore, most MTB measures correlated with age in the expected directions. Executive function, processing speed and memory typically decrease with age and this decrease was reflected in MTB test performance. In contrast, spelling and vocabulary, typically preserved as we age, did not decrease in our sample. Our results support the use of MTB in cognitive aging research.

Categories:
Assessment/Psychometrics/Methods (Adult)
Keyword 1: cognitive functioning
Keyword 2: assessment
Keyword 3: psychometrics
Correspondence: Cindy J. Nowinski, Northwestern University Feinberg School of Medicine

3 Mobile Toolbox: Enrollment of a Large Normative Sample Using the UCSF Brain Health Registry

Rachel L Nosheny[1,2], Monica R Camacho[2], Bernard Landavazo[2], Aaron J Kaat[3], Zahra Hosseinian[3], Richard C Gershon[3], Michael W Weiner[1,2]
[1]University of California San Francisco, San Francisco, CA, USA. [2]Veteran's Administration Advanced Research Center, San Francisco, CA, USA. [3]Northwestern University, Chicago, IL, USA

Objective: A critical need in the neuropsychology field is development and validation of efficient, scalable assessments of cognition. The Mobile Toolbox (MTB), a novel suite of mobile device-compatible, app-based cognitive assessments, was developed to address this need. The goals of this study were (1) To collect longitudinal normative data for the MTB assessments in a large, ethnoculturally and educationally diverse cohort; (2) To assess the feasibility and usability of remote assessment using MTB.
Participants and Methods: Participants were recruited from the UCSF Brain Health Registry (BHR), an online cohort (N>100,000) that collects longitudinal cognitive, functional,

https://doi.org/10.1017/S1355617723009724 Published online by Cambridge University Press