GPS Learning and Memory was related to Working Memory, Mental Flexibility, and Inhibition. GPS Initiation was not related to neuropsychological measures.

Conclusions: Current findings build upon prior work establishing validity of GPS functional assessment measure (Novakovic-Agopian et al., 2012). Seven of 8 GPS subdomains were related to at least one aspect of executive functioning assessed with neuropsychological measures, with the majority related to mental flexibility. Taken together, findings suggest that the GPS converges with traditional measures, offering a method to capture multiple aspects of executive functioning applied together. Further, it may also be useful tool capturing aspects of executive functioning in complex, ecologically-valid settings often not captured with traditional neuropsychological assessment.

Correspondence: Jillian M. Tessier, PhD; San Francisco VA Medical Center, San Francisco, VA; jillian.tessier@va.gov

---

27 Green's Word Memory Test (WMT) Immediate Recall as a Screening Tool for Performance Invalidity

Jonathan D Sober, Nicholas J Pastorek, J. Parks Fillauer, Brian I Miller, Cheyanne C Barba
Michael E. DeBakey VAMC, Houston, TX, USA

Objective: Assessment of performance validity during neuropsychology evaluation is essential to reliably interpret cognitive test scores. Studies (Webber et al., 2018; Wisdom, et al., 2012) have validated the use of abbreviated measures, such as Trial 1 (T1) of the Test of Memory Malingering (TOMM), to detect invalid performance. Only one study (Bauer et al., 2007) known to these authors has examined the utility of Green's Word Memory Test (WMT) immediate recall (IR) as a screening tool for invalid performance. This study explores WMT IR as an independent indicator of performance validity in a mild TBI (mTBI) veteran population.

Participants and Methods: Participants included 211 (M_{age} = 32.1, SD = 7.4; M_{edu} = 13.1, SD = 1.64; 94.8% male; 67.8% White) OEF/OIF/OND veterans with a history of mTBI who participated in a comprehensive neuropsychological evaluation at one of five participating VA Medical Centers. Performance validity was assessed using validated cut scores from the following measures: WMT IR and delayed recall (DR); TOMM T1; WAIS-IV reliable digit span; CVLT-II forced choice raw score; Wisconsin Card Sorting Test failure to maintain set; and the Rey Memory for Fifteen Items test, combo score. Sensitivity and specificity were calculated for each IR score compared with failure on DR. In addition, sensitivity and specificity were calculated for each WMT IR score compared to failure of at least one additional performance validity measure (excluding DR), two or more measures, and three or more measures, respectively.

Results: Results indicated that 46.8% participants failed to meet cut offs for adequate performance validity based on the standard WMT IR cut score (i.e., 82.5%; M = 81.8%, SD = 17.7%); however, 50.2% participants failed to meet criteria based on the standard WMT DR cut score (M = 79.8% SD = 18.6%). A cut score of 82.5% or below on WMT IR correctly identified 82.4% (i.e., sensitivity) of subjects who performed below cut score on DR, with a specificity of 94.2%. Examination of IR cutoffs compared to failure of one or more other PVTs revealed that the standardized cut score of 82.5% or below had a sensitivity of 78.2% and a specificity of 72.4%; whereas, a cut score of 65% or below had a sensitivity of 41% and a specificity of 91.3%. Similarly, examination of IR cutoffs compared to failure of two or more additional PVTs revealed that the cut score of 60% or below had a sensitivity of 45.7% and specificity of 93.1% ; whereas, a cut score of 57.5% or below had a sensitivity of 57.9% and specificity of 90.9% when using failure of three or more PVTs as the criterion.

Conclusions: Results indicated that a cut score of 82.5% or below on WMT IR may be sufficient to detect invalid performance when considering WMT DR as criterion. Furthermore, WMT IR alone, with adjustments to cut scores, appears to be a reasonable way to assess symptom validity compared to other PVTs. Sensitivity and specificity of WMT IR scores may have been adversely impacted by lower sensitivity of other PVTs to independently identify invalid performance.

conducted followed by a 6-factor solution. Promax rotation was used for both EFA models.
Results: The 6-factor EFA solution is reported here. Results indicated the following 6 factors: working memory (Digit Span forward, backward, and sequencing, PASAT trials 1 and 2, NIH-Toolbox List Sorting, LNS), speed/executive function (Stroop color naming, word reading, and color-word interference, NIH-Toolbox Flanker, Dimensional Change, and Pattern Comparison, Trail Making Test A&B), verbal fluency (COWA letters F-A-S), crystallized intelligence (WTAR, NIH-Toolbox Oral Recognition and Picture Vocabulary), visual memory (BVMT immediate and delayed), and verbal memory (HVLT immediate and delayed. COWA animals and NIH-Toolbox Picture Sequencing did not adequately load onto any EFA factor and were excluded from the subsequent CFA.
Conclusions: Findings indicate that in a sample of healthy older adults, these collected measures and those obtained through the NIH-Toolbox battery represent 6 domains of cognitive function. Results suggest that in this sample, picture sequencing and COWA animals did not load adequately onto the factors created from the rest of the measures collected. These findings should assist in interpreting future research using combined NIH-TCB and neuropsychological batteries to assess cognition in healthy older adults.

---

28 Factor Structure of Conventional Neuropsychological Tests and NIH-Toolbox in Healthy Older Adults

Kailey Langer[1], Cheshire Hardcastle[1], Hanna Hausman[1], Jessica Kraft[1], Alejandro Albizu[1], Nicole Evangelista[1], Emanuel Boutzoukas[1], Andrew O'Shea[1], Emily Van Etten[2], Samantha Smith[2], Hyun Song[2], Pradyumna Bharadwaj[2], Georg Hishaw[2], Samuel Wu[1], Steven DeKosky[1], Gene Alexander[2], Eric Porges[1], Michael Marsiske[1], Ronald Cohen[1], Adam Woods[1]
[1]University of Florida, Gainesville, FL, USA.
[2]University of Arizone, Tuscon, AZ, USA

Objective: The National Institutes of Health-Toolbox cognition battery (NIH-TCB) is widely used in cognitive aging studies and includes measures in cognitive domains evaluated for dimensional structure and psychometric properties in prior research. The present study addresses a current literature gap by demonstrating how NIH-TCB integrates into a battery of traditional clinical neuropsychological measures. The dimensional structure of NIH-TCB measures along with conventional neuropsychological tests is assessed in healthy older adults.
Participants and Methods: Baseline cognitive data were obtained from 327 older adults. The following measures were collected: NIH-Toolbox cognitive battery, Controlled Oral Word Association (COWA) letter and animals tests, Wechsler Test of Adult Reading (WTAR), Stroop Color-Word Interference Test, Paced Auditory Serial Addition Test (PASAT), Brief Visuospatial Memory Test (BVMT), Letter-Number Sequencing (LNS), Hopkins Verbal Learning Test (HVLT), Trail Making Test A&B, Digit Span. Hmisc, psych, and GPARotation packages for R were used to conduct exploratory factor analyses (EFA). A 5-factor solution was

---

29 Examining the Relationship between Symptom and Performance Validity Measures Across Referral Subtypes

Kaley Boress[1], Owen Gaasedelen[2], Jeong Hye Kim[1], Doug Whiteside[3], Michael Basso[4]
[1]University of Iowa Hospitals and Clinics, Iowa City, Iowa, USA. [2]Private Practice, Albuquerque, New Mexico, USA. [3]University of Minnesota, Minneapolis, Minnesota, USA. [4]Mayo Clinic, Rochester, Minnesota, USA