1995). Patients were considered credible if all PVT performance fell within normal limits. This resulted in 232 patients in the credible group (Mage = 52.9 years, SDage = 15.2, Medu = 14.0, SDedu = 2.5, 88% male, 71.2% White, 28.3% Black/African American). Patients were considered non-credible if they failed ≥2 PVTs. This resulted in 66 patients in the non-credible group (Mage = 51.6, SDage = 13.79, Medu = 13.1, SDedu = 2.4, 92.4% male, 56.1% White, 43.9% Black/African American). Group assignment was also clinically confirmed. Receiver operating characteristic (ROC) curve analyses were conducted to discriminate between credible and non-credible groups utilizing the established RCFT combination score.
Results: RCFT combination scores distinguished groups, with credible participants scoring higher than non-credible participants ($F[1, 296]=63.76$, $p<.001$, $d=1.11$; M = 56.9, SD = 9.3 vs. M = 46.5, SD = 9.5, respectively). A ROC analysis indicated AUC = .800 (95% CI = .73 to .86). When specificity was set at >90%, a cut-score of ≤46.5 yielded sensitivity at 46.0%. The analogous cut-score from the Lu et al. (2003) study (i.e., ≤47) was associated with a specificity of 88.7 and sensitivity of 46.0% in the current study.
Conclusions: As the Lu et al. (2003) established the combination score of the RCFT with procedures that deviated from the standardized protocol outlined by Meyers and Meyers (1995), clinicians who opted to adhere to Meyers and Meyers' full protocol may have concerns about using the combination score as a PVT. The current study established a similar cut-off score to what Lu et al., (2003) reported (i.e., ≤46.5 vs. ≤47) while following a different administration procedure of the RCFT. Also, the index was moderately sensitive in the current study (i.e., 45.5%) but less so than what Lu et al. reported when using a cut-score that had >90% specificity (i.e., 75.9% sensitivity). This suggests that the index may be robust to deviations in administration procedures. Difference in sensitivity could be related to difference between samples. As the current sample was derived from a clinical, VA setting, current findings extend the generalizability of the index. Future research would benefit exploring if any subgroups would benefit from adjusted cut-scores to reduce the risk of false positive identification.

Categories: Forensic Neuropsychology/Malingering/Noncredible Presentations
Keyword 1: performance validity
Keyword 2: neuropsychological assessment
Keyword 3: psychometrics
Correspondence: Jessica M. Fett, Memphis VA Medical Center, jnich055@gmail.com

85 Use of Embedded Performance Validity Measures Using Verbal Fluency Tests in a Clinical Sample of Military Veterans

Keith P Johnson, Lee Ashendorf, Lauren M Baumann
VA Central Western Massachusetts, Worcester, Massachusetts, USA

Objective: As neuropsychologists aim to collect valid data, maximize the utility of assessments, make effective use of time, and best serve patient populations, measurement of performance validity is considered a critical issue for the field. As effort may vary across an evaluation, including performance validity tests (PVTs) throughout the assessment is important. Incorporating embedded PVTs in addition to free standing PVTs can be particularly useful in this regard. COWAT and animal naming are commonly administered verbal fluency measures. While there have been past investigations into their potential for detecting invalid performance, they are limited, and more research is needed. Perhaps most promising, Sugarman and Axelrod (2015) described a logistic regression derived formula utilizing the combined raw scores of COWAT and animal naming. The current study aimed to investigate the use of embedded PVTs within COWAT and animal naming to provide further support for the use of embedded PVTs in these measures.
Participants and Methods: All subjects were from a mixed clinical sample comprising military veterans from two VA Medical Centers in the northeast U.S., who were referred for neuropsychological evaluation. Subjects deemed credible had zero PVT failures. Subjects were considered non-credible performers if they failed at least two out of a possible eight PVTs administered. Subjects who failed one PVT were excluded from the study (n = 53). The final sample consisted of 116

https://doi.org/10.1017/S1355617723009438 Published online by Cambridge University Press

individuals with credible performance (Mean Age = 35.5, SD = 8.8; Mean Edu = 13.6, SD = 2; Mean Est. IQ = 106, SD = 7.9) and 94 individuals with psychometrically determined non-credible performance (Mean Age = 38.5, SD = 9.4; Mean Edu = 113, SD = 2.1; Mean Est. IQ = 101, SD = 8.7). Performance of COWAT and animals in detecting non-credible performances was evaluated through calculation of classification accuracy statistics and use of the logistic regression formulas reported in Sugarman and Axelrod (2015).
Results: For COWAT, the optimal cutoff was a raw score of ≤27 (specificity = 89%; sensitivity = 31%), and a T-score of ≤35 (specificity = 92%; sensitivity = 31%). For animal naming, optimal cutoffs were ≤16 for raw score (specificity = 92%, sensitivity = 38%) and ≤37 for T-score (specificity = 91%; sensitivity = 33%). The logistic regression formula based on raw scores for both COWAT and animal naming was inadequately sensitive at the recommended cutoff in this sample, but a coefficient of ≥ .28 was revealed to be optimal (91% specificity; 42% sensitivity). When the formula for T-scores was used, a coefficient of ≥ .38 was optimal (91% specificity; 28% sensitivity).
Conclusions: Results of the current research suggest that PVTs embedded within the commonly administered COWAT and animal naming verbal fluency tests can effectively detect low effort, in concordance with generally accepted standards. A logistic regression formula using raw scores in particular appears to be most effective, consistent with findings reported by Sugarman and Axelrod (2015).

Categories: Forensic Neuropsychology/Malingering/Noncredible Presentations
Keyword 1: performance validity
Keyword 2: effort testing
Keyword 3: fluency
Correspondence: Keith P. Johnson Ph.D., VA Central Western Massachusetts, keith.johnson10@va.gov

86 The Examination Between Credible and Non-Credible Groups on Embedded PVT Tests

Krissy E. Smith[1,2], Tara L. Victor[1], Matthew J. Wright[2], Kyle B. Boone[2], Daniel W. Lopez-Hernandez[2]
[1]California State University Dominguez Hills, Carson, California, USA. [2]The Lundquist Institute, Torrance, California, USA

Objective: Performance validity tests (PVTs) are included in neuropsychological testing to ensure examinees are performing to the best of their abilities. There are two types of PVTs: embedded and free standing. Embedded PVTs are tests that are derived from standard neuropsychological tests of various cognitive domains. Freestanding PVTs are tests that are designed with the intention of being a PVT. Research studies show that undergraduate samples do not always performed to the best of their abilities. The purpose of this study was to cross-validate previous research on the topic of performance validity in a college sample. It was predicted that the non-credible group would demonstrate higher failure rates on embedded PVTs compared to the credible group.
Participants and Methods: The sample consisted of 198 neurologically and psychologically healthy undergraduate students with a mean age of 19.69 (SD = 2.11). Participants were broken into two groups: non-credible (i.e., participants that failed two or more PVTs) and credible (i.e., participants that did not failed two or more PVTs). The Rey-Osterrith copy test, Comalli Stroop part A (CSA), B (CSB), and C (CSC), Trail Making Test part A and B, Symbol Digit Modalities Test written (SDMT-W) and oral (SDMT-O) parts, Controlled Oral Word Association Test (COWAT) letter fluency, and Finger Tapping Test were used to evaluate failure rates in our sample. PVT cutoff scores were use from previously validated in the literature. Chi-square analysis was used to evaluate failure rates between the groups.
Results: Chi-square analysis revealed significant failure rate differences between groups on several PVTs. Results revealed that 15% of the non-credible group failed the CSA compared to 1% of the credible group, X2=14.77, p=.000. Meanwhile, 26% of the non-credible group failed the CSB compared to 2% of the credible group, X2=24.72, p=.000. Furthermore, results showed that 11% of the non-credible group failed the CSC compared to 1% of the credible group, X2=13.05, p=.000.Next, 48% of the non-credible group failed the Trail Making Test part A compared to

https://doi.org/10.1017/S1355617723009438 Published online by Cambridge University Press