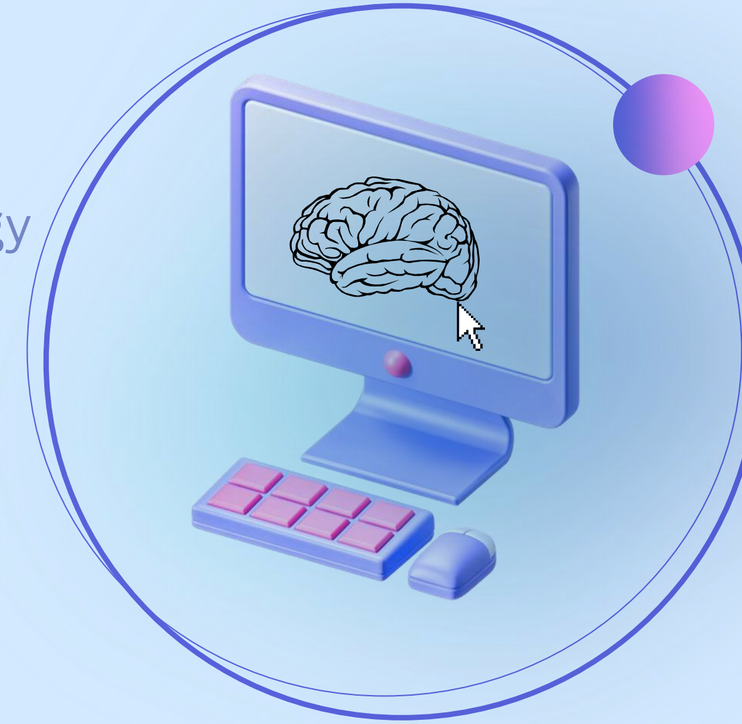
The Making Change Test: Initial validation of a novel digitized performance validity test for tele-neuropsychology

Mira Leese, M.S.

Student Affairs Committee Series

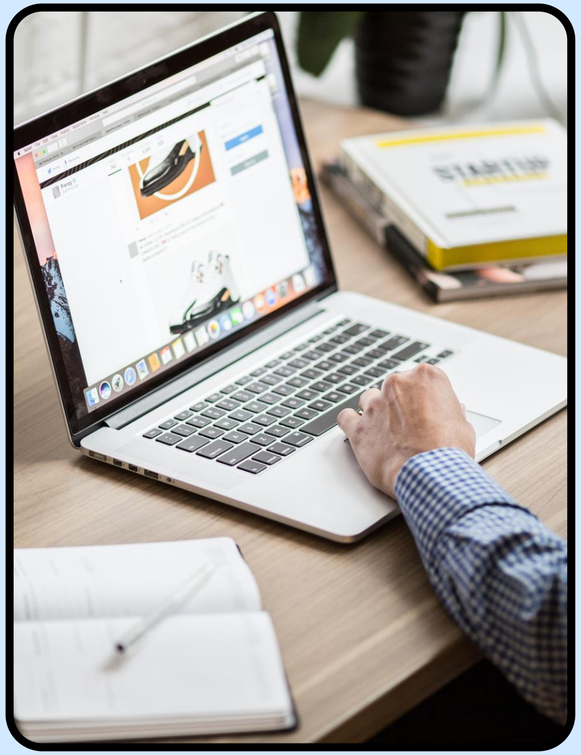
American Academy of Clinical Neuropsychology

June 12th, 2025



OUTLINE

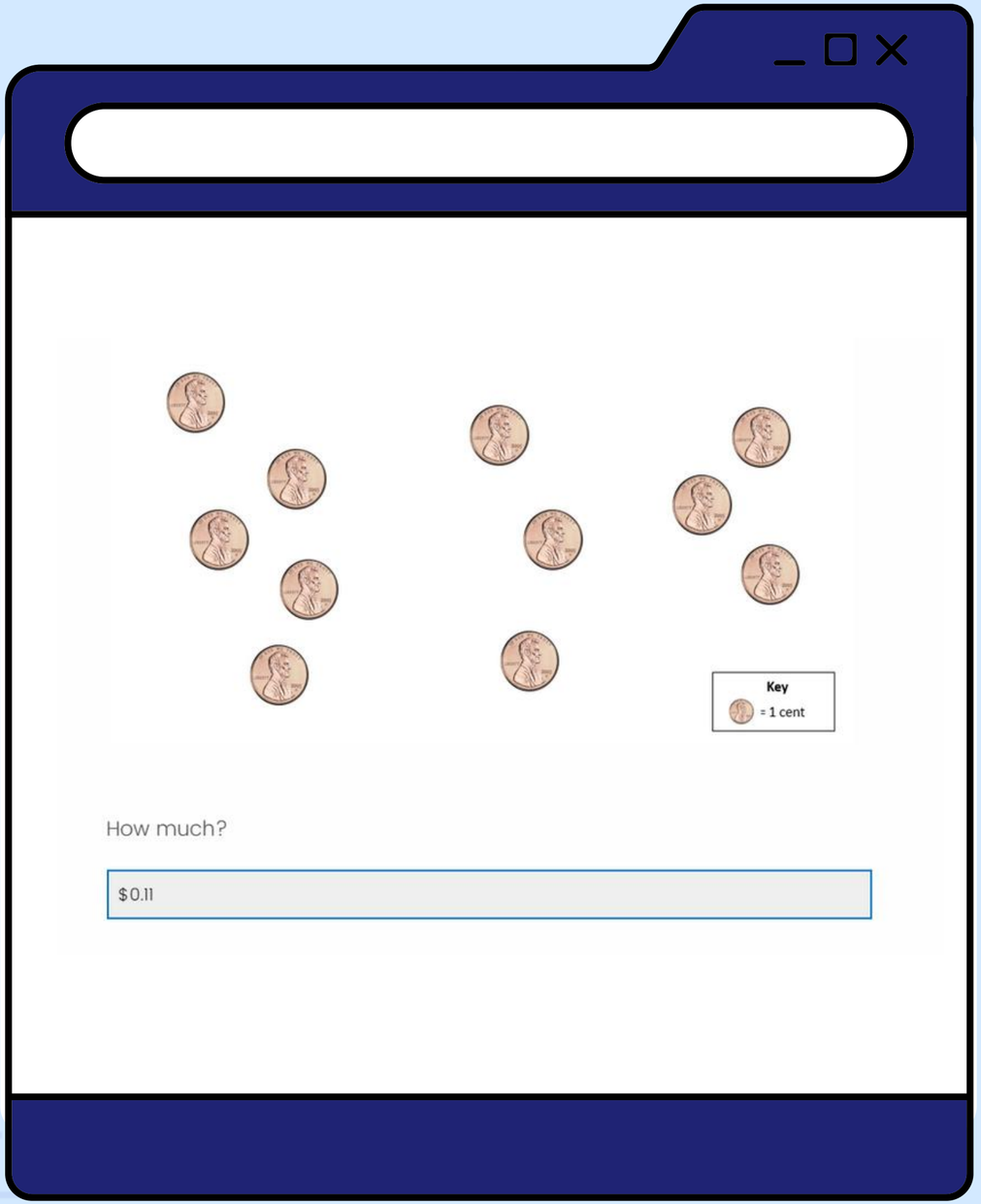
- Background
  - The Making Change Test (MCT)
  - Present Study
- Methods
  - Participants & Procedures
  - Criterion PVTs
  - Group Classification
- Results/Conclusions
- Limitations & Future Directions



TELE-NEUROPSYCHOLOGY & PERFORMANCE VALIDITY TESTS

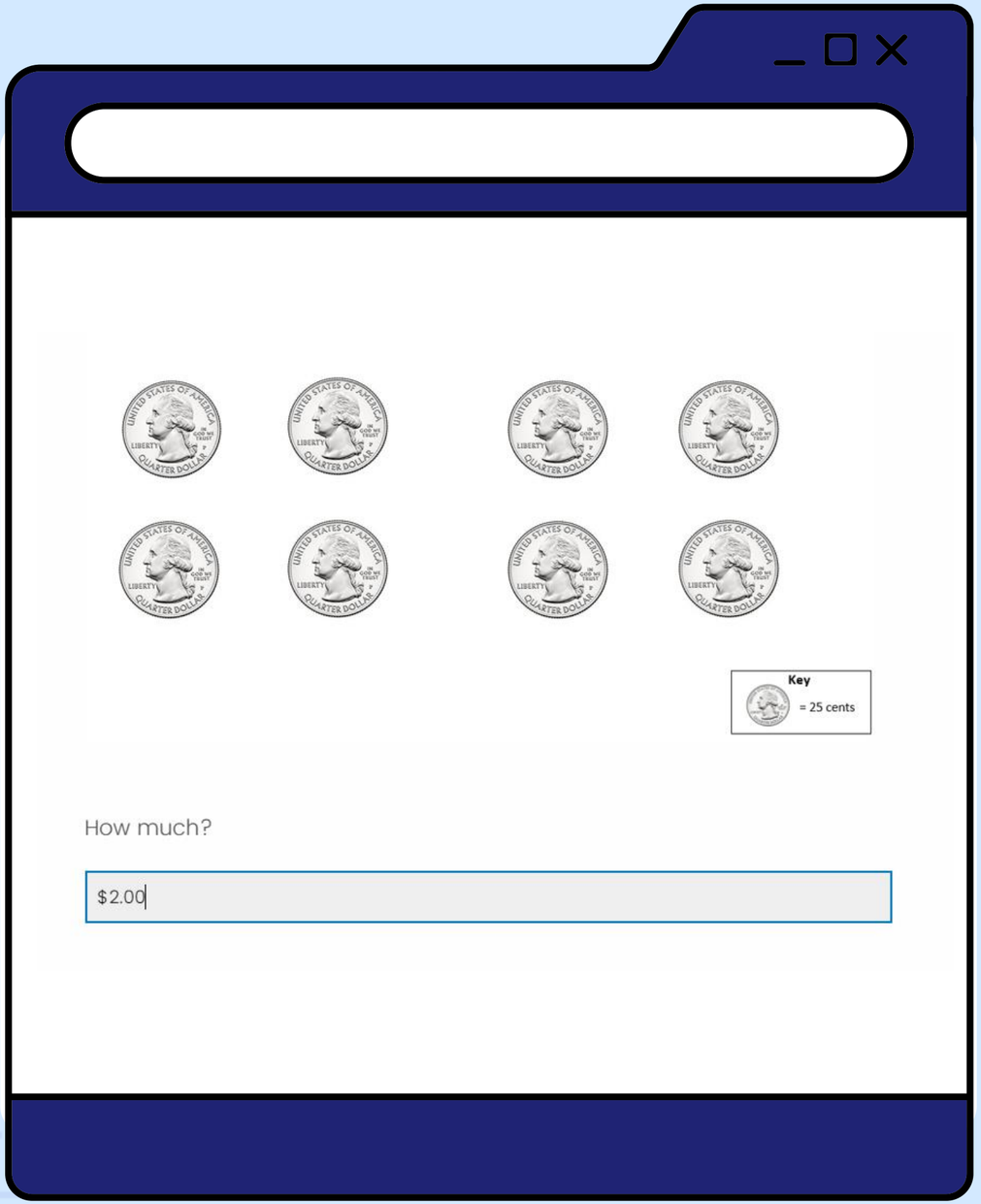
- Defining Tele-Neuropsychology (Tele-NP)
- Advantages of Tele-NP
- Future directions: PVTs
- Development and validation of the Making Change Test (MCT)

The Making Change Test (MCT)



Leese et al. (2024)

The Making Change Test (MCT)



Leese et al. (2024)

The Making Change Test (MCT)

Leese et al. (2024)

PRESENT STUDY

AIMS & HYPOTHESES

- **Aim 1:** Develop cut scores for two MCT validity indicators.

- **Aim 2:** Examine construct validity of two MCT validity indicators.

    - **Hypothesis**: The MCT was predicted to be more strongly associated with non-memory PVTs (e.g., Dot Counting Test) compared to memory-based PVTs (e.g., Test of Memory Malingering)

METHODS

PARTICIPANTS & PROCEDURES

Inclusion Criteria	Exclusion Criteria
18+	Hx of compromised developmental trajectory (i.e., estimated pre-morbid IQ <80 based on the WRAT-4 reading subtest)
English language proficiency	Diagnosis of a major neurocognitive disorder
Referred for a neuropsychological evaluation and administered the MIL/MCT and 4 criterion measures	Known external incentive for poor performance (e.g., not involved in litigation, not undergoing a disability evaluation)

SAMPLE CHARACTERISTICS

Table 1. Demographic characteristics of valid and invalid performance groups.

Measures and characteristics	Valid group (n = 115)	Invalid group (n = 21)	F	Effect size (η^2)
Age	M = 52.65 (SD = 15.88)	M = 50.76 (SD = 11.55)	0.42	.00
Years of education	M = 14.34 (SD = 2.14)	M = 13.30 (SD = 1.81)	4.23	.03*
Female Sex	88 (77%)	14 (67%)	0.91	.01
Racial identity				
White	105 (91%)	17 (81%)	1.28	.02
Black	4 (3%)	2 (9.5%)	0.08	.00
Hispanic	2 (2%)	0 (0%)	0.37	.00
Asian	2 (2%)	0 (0%)	0.37	.00
Multiracial	2 (2%)	2 (9.5%)	2.90	.08
Estimated premorbid IQ	M = 99.68 (SD = 9.44)	M = 94.62 (SD = 11.52)	4.75	.04*
MCT AR-Score	M = 17.03 (SD = 4.63)	M = 25.83 (SD = 13.12)	9.24	.19**
MCT Abbreviated Index	M = −0.34 (SD = 0.90)	M = 1.84 (SD = 3.19)	9.90	.23**
Overall mean cognitive performance				
High average	17 (15%)	0 (0%)	–	–
Average	76 (66%)	2 (9%)	–	–
Low average	19 (16%)	13 (62%)	–	–
Below average	3 (3%)	5 (24%)	–	–
Exceptionally low	0 (0%)	1 (5%)	–	–

Note: N = 136. Overall Mean Cognitive Performance was derived by averaging test scores across 16 performance-based measures.
M: mean; SD: standard deviation; MCT: Making Change Test
*p < .05. **p < .01.

Table 2. Clinical diagnoses for mixed neuropsychiatric sample.

Diagnosis	Valid group (n = 115) (%)	Invalid group (n = 21) (%)
None	10 (9)	0 (0)
Mood disorder (depression, anxiety, and PTSD)	82 (71)	19 (90)
Neurodevelopmental disorder (ADHD, learning disorder)	27 (23	1 (5)
Long COVID-19	7 (6)	1 (5)
Prolonged post-concussive syndrome	4 (3)	3 (14)
Mild traumatic brain injury	6 (5)	2 (10)
Moderate-severe traumatic brain injury	1 (1)	1 (5)
Substance use disorder	2 (2)	0 (0)
Encephalopathy	2 (2)	0 (0)
Epilepsy/brain tumor	6 (5)	0 (0)
Cerebrovascular disease	3 (3)	3 (14)
Chronic pain and fatigue syndrome	15 (13)	5 (24)
Other/unspecified	9 (8)	3 (14)

Note. N=136.
PTSD: post-traumatic stress disorder; ADHD: attention-deficit/hyperactivity disorder

Leese et al. (2024)

CRITERION PVTS

Table 3. Criterion for determining group status.

Criterion performance validity tests	Invalid cut score	Sensitivity	Specificity	Failure cut score reference(s)	Sample failure rate (%)
TOMM Trial 1 (raw score)	≤41	.65	.95	Martin et al. (2020)	24/136 (18)
Dot Counting Test *E*-score	≥13.8	.70	.90	McCaul et al. (2018)	30/136 (22)
Reliable digit span (raw score)	≤6	.30–.35	.96–.97	Schroeder et al. (2012)	8/136 (6)
HVLT-R recognition discrimination (raw score)	≤5	.53	.93	Bailey et al. (2018)	3/72 (4)
RBANS effort index	≥3	.48	.93	Shura et al. (2018)	8/63 (13)

Note: $N = 136$. Participants were administered either the RBANS Effort Index or HVLT-R recognition discrimination.
TOMM: Test of Memory Malingering; HVLT-R: Hopkins Verbal Learning Test-Revised; RBANS: Repeatable Battery for the Assessment of Neuropsychological Status

Leese et al. (2024)

RESULTS

Results: Construct Validity

Table 4. Zero order correlations among performance validity tests.

	MCT AR-Score	MCT abbreviated index	TOMM Trial 1	Dot Counting Test *E*-score	Reliable digit span	RBANS effort index	HVLT-R recognition discrimination
MCT AR-score	–	.78***	.18*	.55***	.31***	.21**	.24**
MCT Abbreviated index		–	.16*	.45***	.37***	.20**	.22**

Note: $N = 136$. Participants were administered either the RBANS Effort Index or HVLT-R Recognition Discrimination.
*p < .05; **p < .01; ***p <.001.
MCT: Making Change Test; TOMM: Test of Memory Malingering; HVLT-R: Hopkins Verbal Learning Test-Revised; RBANS: Repeatable Battery for the Assessment of Neuropsychological Status

Leese et al. (2024)

Results: Cut Scores and Classification Accuracy

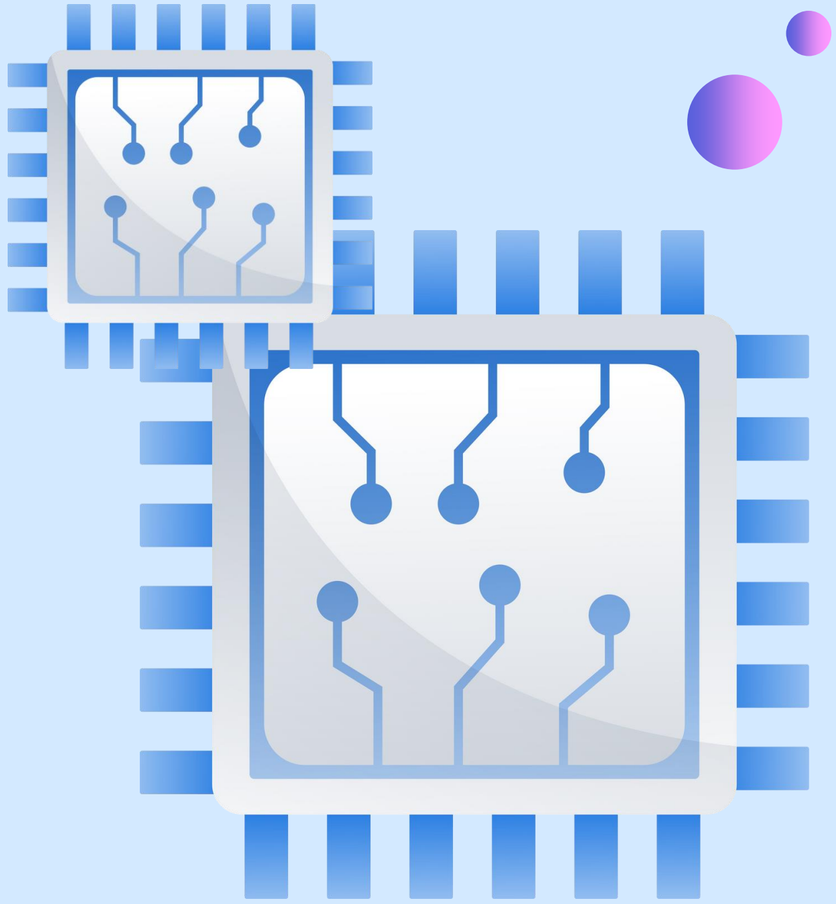
Table 5. Classification accuracy and optimal cut scores for the Making Change Test.

					10% base rate		20% base rate		30% base rate	
Performance validity test indicator	Area under the curve (95% CI)	Cut score	Sensitivity	Specificity	PPV	NPV	PPV	NPV	PPV	NPV
MCT	.77***	≥23	.43	.89	.30	.93	.49	.86	.63	.78
AR-Score	(.68, .88)	≥24	.38	.90	.30		.49		.62	.77
		≥25	.29	.92	.29	.93	.48	.85	.61	.75
		≥26	.29	.94	.35	.92	.55	.84	.67	.76
						.92		.84		
MCT	.77***	≥.87	.51	.89	.34	.94	.54	.88	.67	.81
Abbreviated index	(.66, .89)	≥.90	.50	.90	.36	.94	.56	.88	.68	.81
		≥.95	.48	.91	.37	.94	.57	.88	.70	.80
		≥.98	.43	.92	.37	.94	.57	.87	.70	.79

Note: $N = 136$. AR-Score values are expressed as raw scores. Abbreviated Index values are expressed as z-scores. CI: confidence intervals; PPV: positive predictive value; NPV: negative predictive value; MCT: making change test
***$p < .001$.

Leese et al. (2024)

LIMITATIONS & FUTURE DIRECTIONS



REFERENCES

1.Bilder, R. M., Postal, K. S., Barisa, M., Aase, D. M., Cullum, C. M., Gillaspy, S. R., Harder, L., Kanter, G., Lanca, M., Lechuga, D. M., Morgan, J. M., Most, R., Puente, A. E., Salinas, C. M., & Woodhouse, J. (2020). InterOrganizational practice committee recommendations/guidance for teleneuropsychology (TeleNP) in response to the COVID-19 pandemic. The Clinical Neuropsychologist, 34(7–8), 1314–1334. https://doi.org/10.1080/13854046.2020.1767214

2.Finley, J. C. A. (2024). Performance validity testing: the need for digital technology and where to go from here. Frontiers in Psychology, 15, 1452462. https://doi.org/10.3389/fpsyg.2024.1452462

3.Kanser, R. J., Rapport, L. J., Bashem, J. R., & Hanks, R. A. (2019). Detecting malingering in traumatic brain injury: Combining response time with performance validity test accuracy. The Clinical Neuropsychologist, 33(1), 90–107. https://doi.org/10.1080/13854046.2018.1440006

4.**Leese, M. I., Finley, J. C. A., Roseberry, J. E., & Hill, S. K. (2024). The Making Change Test: Initial validation of a novel digitized performance validity test for tele-neuropsychology. The Clinical Neuropsychologist, 1-14. https://doi.org/10.1080/13854046.2024.2352898**

5.Martin, P. K., & Schroeder, R. W. (2020). Base rates of invalid test performance across clinical nonforensic contexts and settings. Archives of Clinical Neuropsychology, 35(6), 717–725. https://doi.org/10.1093/arclin/acaa017

6.McCaul, C., Boone, K. B., Ermshar, A., Cottingham, M., Victor, T. L., Ziegler, E., Zeller, M. A., & Wright, M. (2018). Cross-validation of the Dot Counting Test in a large sample of credible and non-credible patients referred for neuropsychological testing. The Clinical Neuropsychologist, 32(6), 1054–1067. https://doi.org/10.1080/13854046.2018.1425481

7.Parsons, T. D., McMahan, T., & Kane, R. (2018). Practice parameters facilitating adoption of advanced technologies for enhancing neuropsychological assessment paradigms. The Clinical Neuropsychologist, 32(1), 16–41. https://doi.org/10.1080/13854046.2017.1337932

8.Patrick, S. D., Rapport, L. J., Kanser, R. J., Hanks, R. A., & Bashem, J. R. (2021). Performance validity assessment using response time on the Warrington Recognition Memory Test. The Clinical Neuropsychologist, 35(6), 1154–1173. https://doi.org/10.1080/13854046.2020.1716997

9.Paulo, R., & Albuquerque, P. B. (2019). Detecting memory performance validity with DETECTS: A computerized performance validity test. Applied Neuropsychology. Adult, 26(1), 48–57. https://doi.org/10.1080/23279095.2017.1359179

10. Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E., & Marshall, P. S. (2012). Reliable digit span: A systematic review and cross-validation study. Assessment, 19(1), 21–30. https://doi.org/10.1177/1073191111428764Baileyet%20al%202018

11. Scott, J. C., Moore, T. M., Roalf, D. R., Satterthwaite, T. D., Wolf, D. H., Port, A. M., Butler, E. R., Ruparel, K., Nievergelt, C. M., Risbrough, V. B., Baker, D. G., Gur, R. E., & Gur, R. C. (2023). Development and application of novel performance validity metrics for computerized neurocognitive batteries. Journal of the International Neuropsychological Society, 29(8), 789–797. https://doi.org/10.1017/S1355617722000893

12. Shura, R. D., Brearly, T. W., Rowland, J. A., Martindale, S. L., Miskey, H. M., & Duff, K. (2018). RBANS validity indices: A systematic review and meta-analysis. Neuropsychology Review, 28(3), 269–284. https://doi.org/10.1007/s11065-018-9377-5

13. Singh, S., & Germine, L. (2021). Technology meets tradition: A hybrid model for implementing digital tools in neuropsychology. International Review of Psychiatry (Abingdon, England), 33(4), 382–393. https://doi.org/10.1080/09540261.2020.1835839

14. Stevens, A., Bahlo, S., Licha, C., Liske, B., & Vossler-Thies, E. (2016). Reaction time as an indicator of insufficient effort: Development and validation of an embedded performance validity parameter. Psychiatry Research, 245, 74–82. https://doi.org/10.1016/j.psychres.2016.08.022

15. Sweet, J. J., Heilbronner, R. L., Morgan, J. E., Larrabee, G. J., Rohling, M. L., Boone, K. B., Kirkwood, M. W., Schroeder, R. W., & Suhr, J. W. (2021). American Academy of Clinical Neuropsychology (AACN) 2021 consensus statement on validity assessment: Update of the 2009 AACN consensus conference statement on neuropsychological assessment of effort, response bias, and malingering. The Clinical Neuropsychologist, 35(6), 1053–1106. https://doi.org/10.1080/13854046.2021.1896036

16. Tsoy, E., Zygouris, S., & Possin, K. L. (2021). Current state of self-administered brief computerized cognitive assessments for detection of cognitive disorders in older adults: A systematic review. The Journal of Prevention of Alzheimer's Disease, 8(3), 267–276. https://doi.org/10.14283/jpad.2021.11

QUESTIONS?