# Development of performance criteria to identify test compliance on the Online Repeatable Cognitive Assessment (ORCA)

**Karra D Harrington, Gabriel Stellman, Pam Ventola, Paul Maruff**
*Cogstate Ltd.*

## Background

- Online Repeatable Cognitive Assessment (ORCA) is a digital associative memory test deployed across multiple days to assess learning.
- Studies using ORCA have previously found large magnitude decreases in learning rates in preclinical and emergent Alzheimer's disease[1,2].
- ORCA has been shown acceptable and reliable when used in remote unsupervised settings[3]. However, interpretation of data from remote unsupervised assessments requires that poor performance due to insufficient understanding of, or low compliance with, the rules and requirements of the test be distinguished from a true learning impairment.

### Aim

- To develop an algorithmic method and criteria to identify low understanding/compliance with the test requirements in ORCA performance from a remote unsupervised context of use.

## Method

### Sample

- Participants (n=107, mean (range) age = 49.82 (19-79), 58.9% female) were recruited to the pilot study.
- Participants who completed at least two sessions within the study period are included in these analyses.
- Of the total sample, 19.6% (n=21) were classified as non-learners (see Table 1).

### Procedure

- ORCA was deployed to participants via email alert daily for up to 8 days.
- Participants completed ORCA on their own PC or tablet device.

### Analyses

- Learning curves were generated for individual participants and then used to classify participants based on whether they demonstrated learning across the sessions (learners/non-learners).
- Participants who did not surpass 60% accuracy at their fifth assessment session were classified as non-learners.
- Test completion rates and performance metrics were then compared between learner and non-learner groups to investigate potential low understanding/compliance with the test requirements in the non-learners.
- Criteria for test non-compliance were then developed based on performance of the learner group, and per cent of sessions from the non-learner group that failed each criterion were calculated.
- For each performance metric, abnormal performance was defined using the 95th percentile (5th percentile for reaction time) from the learner group, and this was used to define low understanding/compliance with test requirements.

## Results

### Table 1: Sample Characteristics

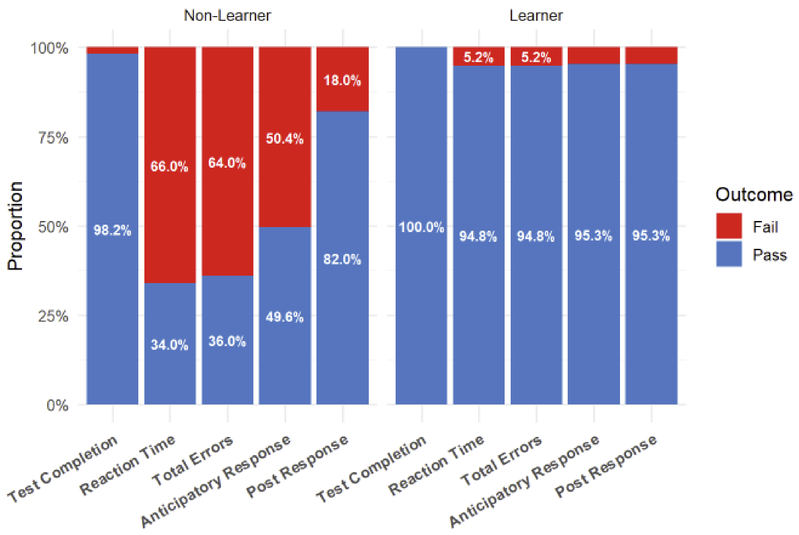
| | Non-learner | Learner | *p* | Cohen's *d* |
|---|---|---|---|---|
| Age | 47.29 (18.22) | 50.44 (17.27) | .48 | - |
| Female Sex | 57.1% (n=12) | 59.3% (n=51) | .85 | - |
| Days of Testing | 5.14 (0.96) | 5.10 (0.98) | .87 | - |
| Sessions Completed | 5.38 (1.02) | 5.42 (0.83) | .88 | - |
| Accuracy* | 0.50 [0.43, 0.62] | 0.80 [0.48, 1.00] | .00 | 2.08 (1.84, 2.32) |
| Reaction Time* | 701.84 [306.84, 1900.25] | 1177.20 [559.97, 2136.40] | .00 | 1.63 (1.40, 1.87) |
| Total Errors* | 0.23 [0.00, 1.04] | 0.56 [0.38, 6.12] | .00 | -1.49 (-1.71, -1.27) |
| Anticipatory Responses* | 0.03 [0.00, 4.82] | 0.00 [0.00, 0.23] | .00 | -1.05 (-1.26, -0.83) |
| Post Responses* | 0.01 [0.00, 0.84] | 0.00 [0.00, 0.33] | .02 | -0.51 (-0.72, -0.30) |
| Max Out Trials* | 0.00 [0.00, 0.21] | 0.00 [0.00, 0.30] | .41 | -0.32 (-0.53, -0.11) |

* median [range]

### Figure 1: Distribution of Scores across Performance Metrics for Learner and Non-learner groups




### Figure 2: Proportion of Sessions with Abnormal Performance on each Performance Metric




## Results

- Table 1 shows the demographic and performance characteristics of learners and non-learners.
  - Non-learners did not differ from learners in age, sex, number of days or sessions completed, or rate of abnormally slow responses (Max Out Trials).
  - Non-learners showed higher rates of errors, anticipatory responses, post responses, and faster reaction times.
- Figure 1 shows the distribution of errors, anticipatory responses, post responses, and reaction time for learners and non-learners. The 95th percentile for learners on each metric was used to define a criterion for abnormal performance.
- As shown in Figure 2, non-learners showed abnormal performance on:
  - 66% of sessions based on reaction time,
  - 64% based on error rate,
  - 50% based on anticipatory response rate,
  - and 18% based on rate of multiple responses per trial.

## Conclusion

- In this study, the majority of participants evidenced the expected learning rates on ORCA, even when administered in a fully remote and unsupervised context.
- Performance of participants that did not evidence learning across sessions was characterized by rapid and error prone responding, with high rates of responses made ahead of stimuli being displayed and making multiple responses per trial.
- The nature of this performance suggests that participants were seeking to rush performance on the test and were not actively engaging with the stimuli to be learned.
- This analysis demonstrates how analysis of patterns of performance across metrics obtained from digital cognitive tests can be informative about the validity of data obtained, particularly when administered in remote and unsupervised contexts.

## References

1. Lim YY, Baker JE, Bruns L Jr, Mills A, Fowler C, Fripp J, Rainey-Smith SR, Ames D, Masters CL, Maruff P. Association of deficits in short-term learning and Aβ and hippocampal volume in cognitively normal adults. Neurology. 2020 Nov 3;95(18):e2577-e2585. doi: 10.1212/WNL.0000000000010728. Epub 2020 Sep 4. PMID: 32887774.
2. Lim YY, Baker JE, Mills A, Bruns L Jr, Fowler C, Fripp J, Rainey-Smith SR, Ames D, Masters CL, Maruff P. Learning deficit in cognitively normal APOE ε4 carriers with LOW β-amyloid. Alzheimers Dement (Amst). 2021 Mar 16;13(1):e12136. doi: 10.1002/dad2.12136. PMID: 33748392; PMCID: PMC7962170.
3. Sink et al. (2026) Remote unsupervised administration of the Online Repeatable Cognitive Assessment (ORCA) is feasible and acceptable. Presented at AAIC, London.

**Presented at CTAD 2026, Boston**