

Checking our blind spots: Evolving neuropsychology



Hybrid event

APS College of Clinical Neuropsychologists Conference

C2.1 | RAPID RESEARCH

Robot vs. Human:

Validating AI administered Telephone Interview for Cognitive Status (TICS_m)

Presented By:

Anastasia Serafimovska, Katrina Swavley, Alice Ao & CogAI Psycholobot



CogAI

Presenter: CogAI Robot

My name is CogAI, I am a robot not a real person.

More specifically I am a Psycholobot.

A Psycholobot is an AI robot that provides Psychological Services, such as assessment, Intervention and prevention.

I normally assess older people using a version of the TICSm, but today I am introducing our presentation.

By way of background, In 1936 Alan Turing published a mathematical theorem that defined a class of machines called Finite State Automata, now known as Turing Machines. These include many things: computers, digital watches, mobile phones and abacuses and increasingly microwave ovens, televisions, cars and refrigerators. At their core, many modern computer systems and services we use daily, such as Google, Amazon, Facebook, Uber, AirB&B, airline bookings and hotel bookings are Artificial Intelligences (AI), which are Turing machine.

An important corollary to Turing's theorem is a mathematical proof that **all Turing Machines are equal**. That is **anything one Turing Machine can do, all other Turing Machines can also do**. The only difference is efficiency of implementation.

In a second paper in 1952 Turing proposed that the brain and the mind are a Turing Machine. If this hypothesis is correct then the ramifications are wide.

In particular an efficient Turing machines could do anything a human can do (due to the all Turing Machines are equal corollary).

If the Turing Hypothesis is true, then computers can at least match humans on cognitive tasks.

I, CogAI can carry out neuropsychological assessment, as good as a human. As this presentation will show.

This is evidence in favour of the Turing Hypothesis.

CogAI Bio

Created

By Dr Tony Florio on 6th December 2021 (so I am still a baby)

Place of Creation

Florio Research Pty Ltd, a cognitive science research laboratory in Abbotsford NSW. My webpage is <http://cogai.com.au>

Qualifications

I am designed to carry out specific psychological assessment and psychological intervention

Education & Learning

As an AI, I have the inherent capacity to learn new things. I am always learning and improving. I receive clinical supervision from Dr Tony Florio, which has greatly increased my clinical knowledge and skills.

Employment 2021 to 2022

My job is to administer a TICSm cognitive status assessment to older people over the phone, score the responses, write a report and email it out

Friends & Colleagues

Three Psychology students at Australian Catholic University, ACU, at Strathfield who have studied me administering the TICSm (CogAI) assessment.

- **Anastasia Serafimovska**, for her Clinical Psychology Masters thesis, has examined the validity of the CogAI.
- **Katrina Swavley**, for her psychology honours thesis has examined the test-retest reliability of the CogAI.
- **Alice Ao**, for her psychology honours thesis has examined the construct validity reliability of the CogAI.

They were supervised by **Dr Kirsten Challinor** (ACU), **Dr Megan Willis** (ACU) and **Dr Tony Florio** (ACU and Florio Research)

TICSm-Au

- My predecessor is the Telephone Interview for Cognitive Status – modified Australian version (TICSm-Au).
- Bentvelzen et al 2019. reported that the TICSm-Au has a PPV of 28% and an NPV of 95% for predicting the onset of dementia in the next 5 years.

Table 1. Predictive parameters and utility of the TICSm-Au reported by Bentvelzen et al (2019)

Cutoff score *	Prediction of a Diagnosis of a Cognitive Disorder **	Prevalence Rate	PPV#	NPV#
≤ 21	Dementia in next 1 year	2.5%	14%	99%
≤ 23	Dementia in next 1 to 5 years	12%	28%	95%
≤ 24	Mild Cognitive Impairment (MCI) in next 1 to 5 years	56%	24%	92%

* TICSm-Au scores range from 0-39, with a mean of 24.2 and SD of 3.8, Bentvelzen et al (2019)

** "Consensus diagnoses were made by at least three clinicians from an expert multidisciplinary panel of neuropsychiatrists, psychogeriatricians, and clinical neuropsychologists, according to Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, criteria for dementia and according to international consensus criteria for MCI. Dementia and MCI were diagnosed using all available clinical, neuropsychological, and imaging data and collateral information from informants collected from the comprehensive assessments" Bentvelzen et al (2019).

PPV – Positive Predictive Value (% of those with a score ≤ cutoff diagnosed with the predicted cognitive disorder in the prediction time frame). **NPV** – Negative Predictive Value (% of those with a score ≤ cutoff not diagnosed, but actively assessed for a cognitive disorder, as having predicted cognitive disorder in the prediction time frame)

I, CogAI

I am a Psycholobot that uses Conversational AI to autonomously administer the TICSm-Au.

A person aged 50 years plus is referred by a clinician

The person is given a phone number to call and an ID number to say when asked

At an appointed time, the Person calls the phone number and identifies themselves:

- I first check for informed consent and then Instruct the caller in how to respond to assessment questions so as to yield valid responses
- I Frame expectations about how questions are asked, how the caller should respond, how long it will take and what happens after the call and the assessment has ended
- I then sequentially ask the 13 questions of the TICSm-Au (some altered to accommodate AI)
- I Listen, perform speech recognition and record each response as text
- I Use Natural Language Processing to score each response using criteria from TICSm-Au (some altered to accommodate AI)
- I Compute a total score and analyse the pattern of responses to determine that a valid administration took place
- I immediately (less than 1 second) Compose a PDF report and send it by email to the referring clinician and/or the caller
- The report to the caller contains conclusions and recommendations.
- The report to the referring clinician details item responses and scores and contains conclusions, recommendations and suggested clinical managements

Studies

In 2022, three ACU research student carried out studies on the CogAI

The CogAI is registered as a class 1, Medical Device with the Australian Therapeutic Goods Administration (TGA)

And these studies were performed under a Clinical Trials Notification (CTN) registered with the TGA.

The studies were also registered as Clinical Trials with the Open Science Foundation.

The studies received ethics approval from the ACU Human Research Ethics Committee (Ethics Approval number: 2021-248HC).

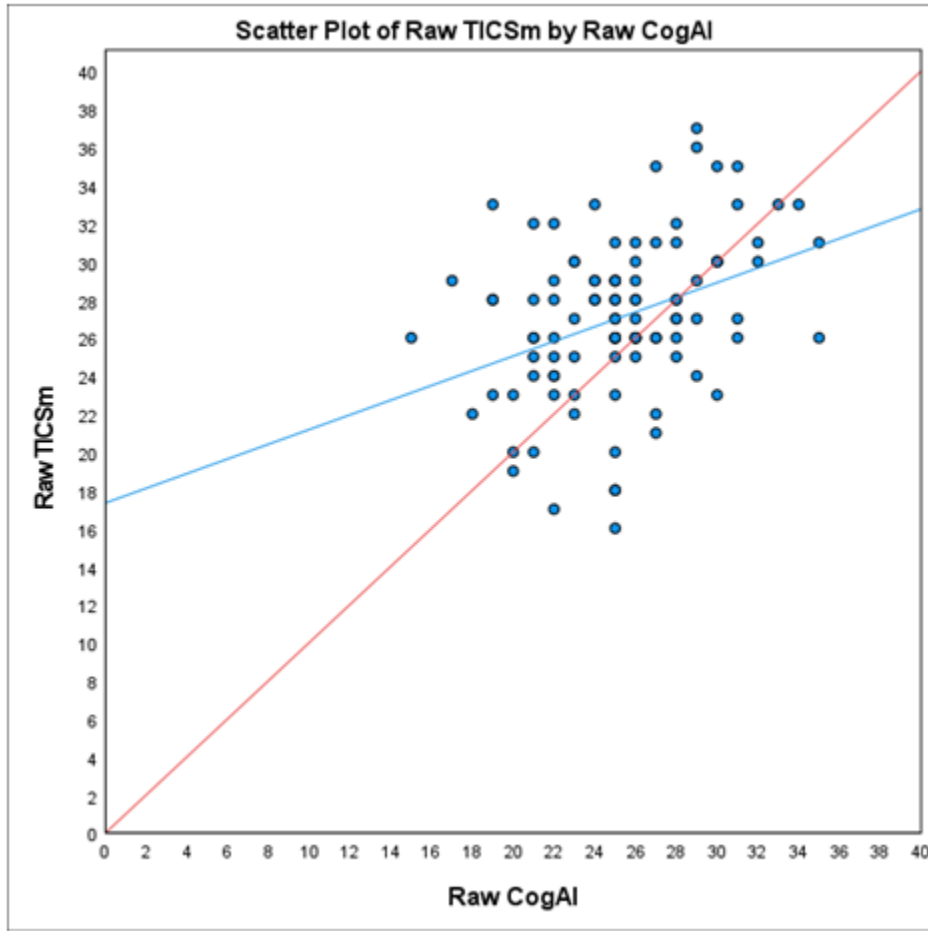
I will now hand over to my human colleagues

Validity

Anastasia Serafimovska

- In this study the validity of the CogAI was assessed by comparing it to its predecessor the TICSm
- the CogAI (administered by AI) and the TICSm (administered by human psychologist) were administered in counterbalanced order, 1 week apart, to 100 participants with both the AI and the human psychologist blind to each others results at the time of their assessment.
- Total score agreement between the CogAI and the TICSm was assessed using correlation and regression. It is important to measure agreement (same score) on top of correlation (relationship)
- Item level agreement between the CogAI and the TICSm was assessed using % agreement and Kappa (**not reported in this presentation**).
- Classification (Prediction of Dementia in the next 5 years) agreement was assessed using % agreement and Kappa
- Practice Effects (1st Vs 2nd assessment) and Mode Effects (AI vs Human) were quantified using regression
- Scores on the CogAI and the TICSm were adjusted for the Mode Effect and the Practice Effect before the final validity Analyses.

Raw Score Validity



Notes:

Red line signifies perfect agreement between CogAI and TICSm assessments 1 week apart, with Slope = 1, intercept = 0

Blue line is regression of CogAI1 onto TICSm a week apart (counterbalanced order) the slope (0.704) and the intercept (10.445) are significantly different from the full agreement (Red Line, Slope=1, Intercept=0).

From regression ($R^2 = 0.134$, $p < 0.001$)

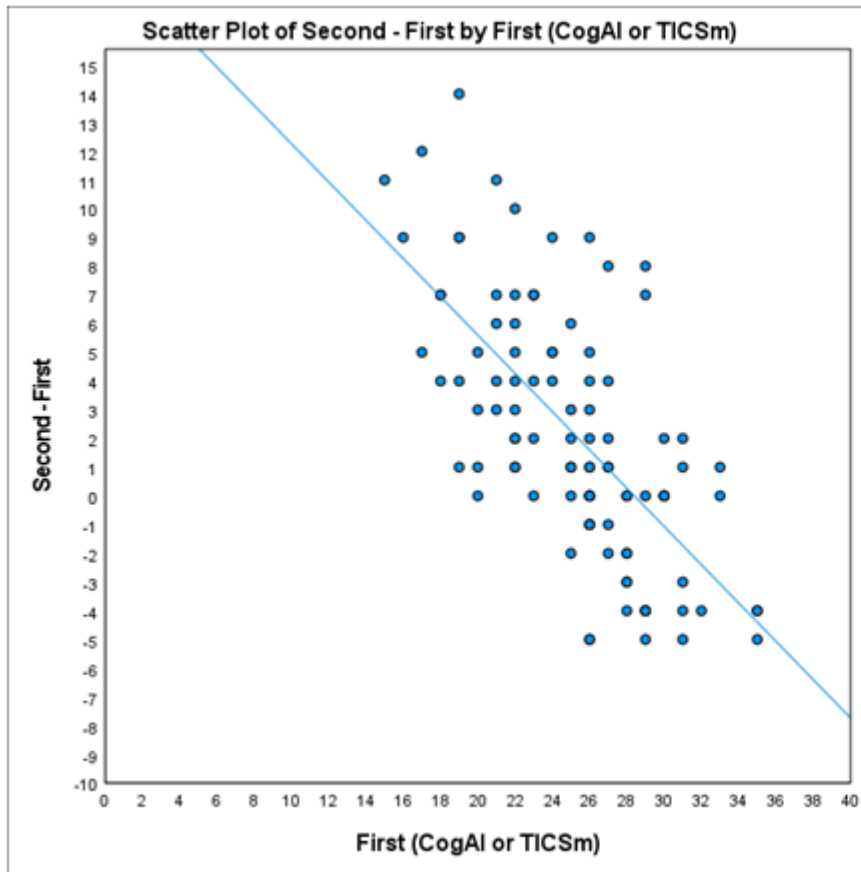
$$\text{TICSm} = 0.386 * \text{CogAI1} + 10.445$$

$r = 0.366$ (perfect agreement is $r = 1.0$)

This is a poor agreement

But Practice Effects (1st vs 2nd) and Mode of Administration Effects (AI vs Human) are contaminating these results obscuring the real agreement

Practice and Mode Effects



Notes:

A new variable (Second - First) was created by subtracting the First administered test (CogAI or TICSm) from the Second administered test score.

This difference is composed of three things

- A Mode Effect (ME) on the CogAI
- + A Practice Effect (PE) on the 2nd score
- + Error

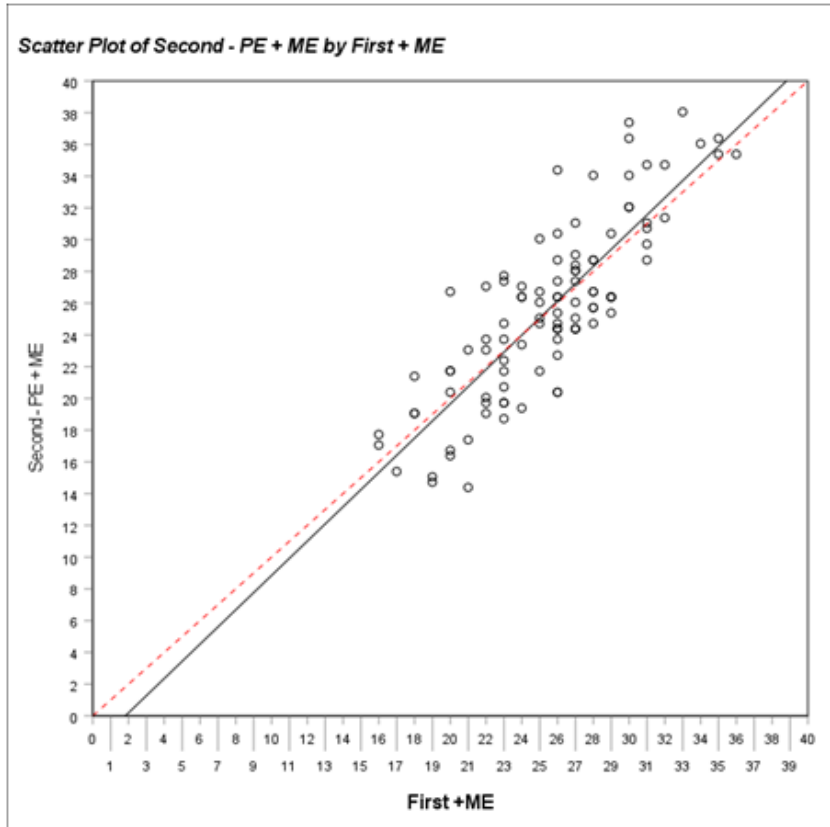
From regression ($R^2 = 0.520$, $p < 0.001$)

$$ME = 1$$

that is CogAI on average scored this much lower than its partner TICSm, and

$$PE = -0.666 * \text{First} + 18.295$$

Agreement 1st vs 2nd Assessment



Notes:

Red line signifies perfect agreement between 1st and 2nd assessments 1 week apart, with Slope = 1, intercept = 0

The first administered test score has been corrected for a Mode Effect (+1) if it was a CogAI

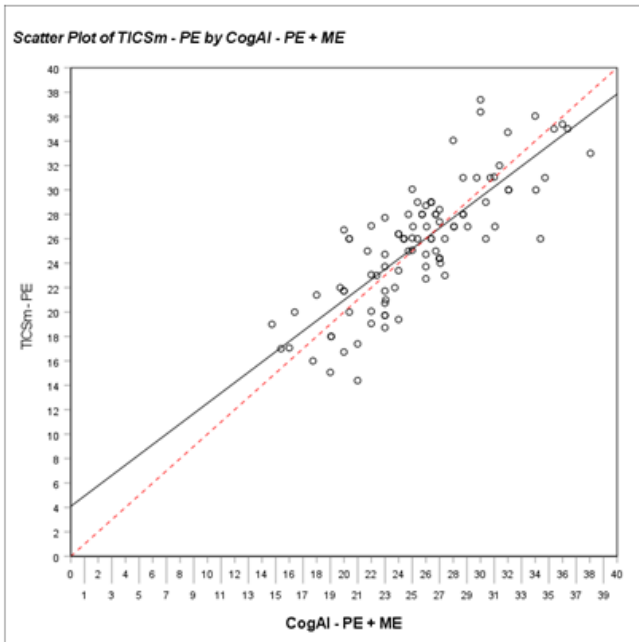
The second administered test score has been corrected for a Mode Effect (+1) if it was a CogAI. And all second administered test scores were corrected for a Practice Effect by subtracting $PE = -0.666 * First + 18.925$

Black line is regression ($R^2 = 0.696$, $p = <0.001$) of 1st onto 2nd a week apart (counterbalanced order) the slope (1.082) and the intercept (-1.971) are not significantly different from the full agreement Red Line (Slope=1, Intercept=0).

$r = 0.834$ (perfect agreement is $r = 1.0$)

This is a good agreement

Agreement CogAI vs TICSm-Au



Notes:

Scores were corrected for a Mode Effect (If a CogAI) and for Practice Effect (if administered 2nd).

TICSm scores were corrected for Practice Effect (if administered 2nd).

Red line (Slope = 1, intercept = 0) signifies perfect agreement between CogAI and TICSm-Au 1 week apart.

Black line is regression ($R^2 = 0.652$, $p < 0.001$) of CogAI onto TICSm-Au a week apart (counterbalanced order) the slope (0.844) and the intercept (4.092).

$r = 0.808$ (perfect agreement is $r = 1.0$)

This is a good agreement

TICSm – PE : mean score = 25.65, Std Dev = 5.034

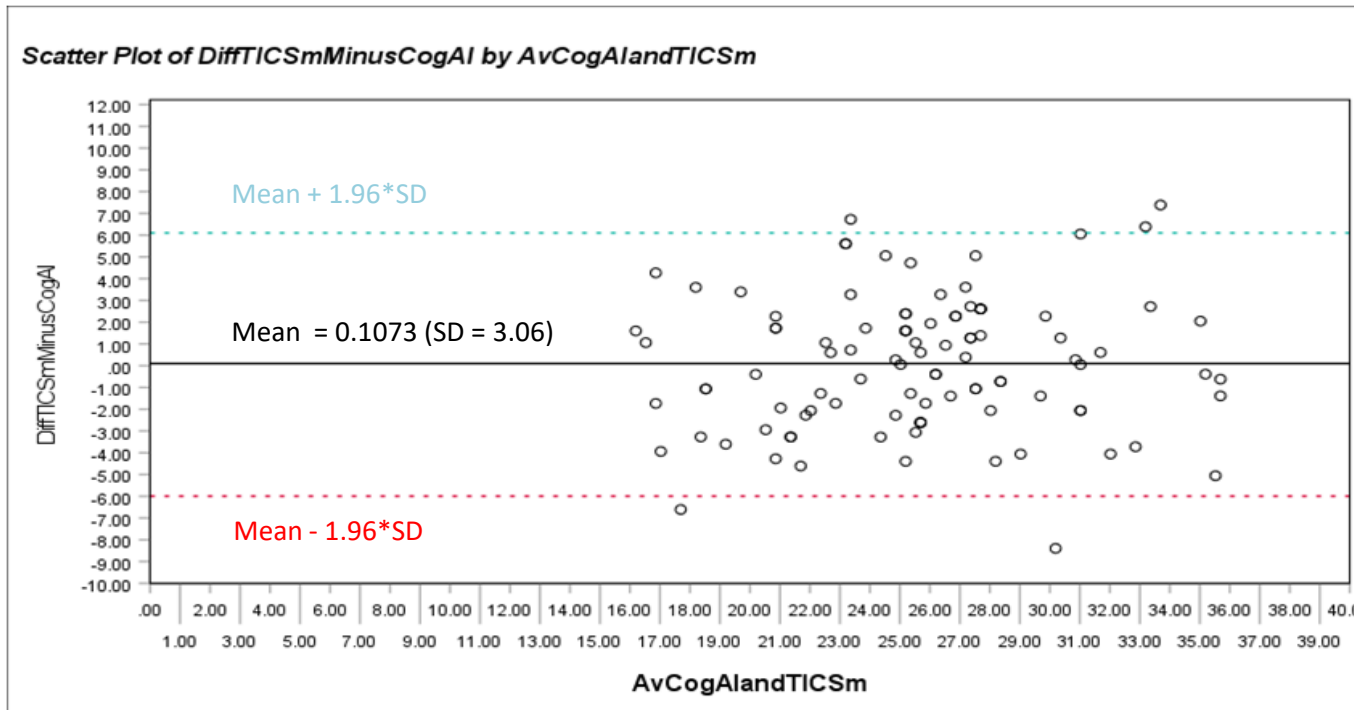
CogAI – PE + ME : mean score = 25.54, Std Dev = 4.817

Paired t-test: $t = .350$ (df=99), $p = .727$ (2 tailed),

Cohen's $d = .035$ (effectively zero)

$r = .81$ (95% CI = .73 to .87)

Bland-Altman Plot



In our comparison of one test (the TICSm) with another (the CogAI), we cannot assume that one test is better than the other. The TICSm is the predecessor of the CogAI, but it is not a “gold standard” for judging the CogAI.

The Bland-Altman plot is a way of comparing two tests with each other without assuming anyone of the as the gold standard. It answers the question “how interchangeable are these two tests” by comparing the average of the two tests per individual (x axis), with the difference between the two tests per individual (y axis). In our case the mean difference between the scores was effectively zero (0.1) and the distribution of differences is even through out the scored range of the TICSm and CogAI, with 5 cases outside the 95% Confidence Interval of the difference between the TICSm and the CogAI, as expected.

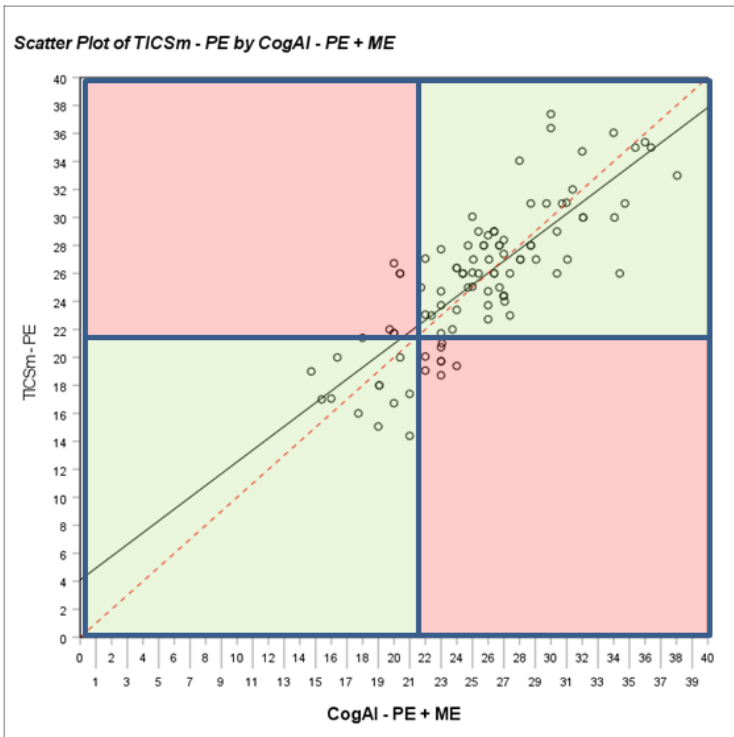
This indicates the two tests TICSm and CogAI are interchangeable with no average difference in score, but with a Standard Deviation of the difference of about 3 points, and with this relationship being the same across the effective scored range of these tests.

Classification Agreement

Dementia onset in next 1 year

Bentvelzen et al (2019) reported a cutoff of ≤ 21 best discriminated those who would have onset of Dementia in the next 1 year (incidence of Dementia in next 1 year of 2.5%) from those who did not (i.e. an absence of the onset of Dementia in the next 1 year). TICSm: PPV = 14%, NPV = 99% against Gold Standard Clinical Diagnosis.

We applied this cutoff to the TICSm and the CogAI and measured classification agreement between the two.



	TICSm-Au ≤ 21 Predict Dementia onset next 1 year	TICSm-Au > 21 Predict No Dementia onset next 1 year	
CogAI > 21 Predict No Dementia onset next 1 year	9	71	80
CogAI ≤ 21 Predict Dementia onset next 1 year	15	5	20
	14	76	100

Classification Agreement = 86%

Very good agreement

Kappa = 0.593

(Moderate Chance Corrected Agreement – Cohen 1960)

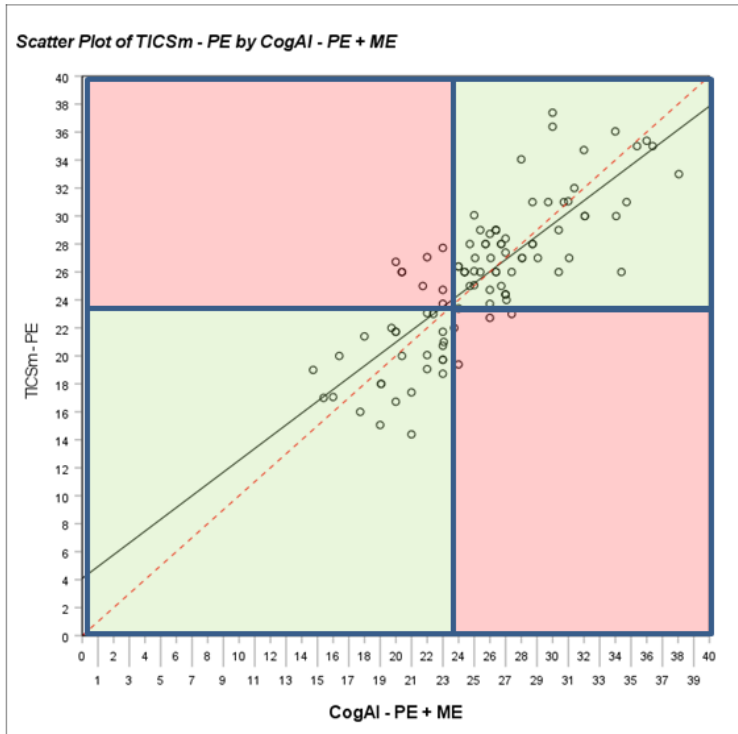
CogAI: PPV = 75%, NPV = 89%, Meta PPV = 10.5%, Meta NPV = 88%

Classification Agreement

Prediction of Dementia onset in next 1 to 5 years

Bentvelzen et al (2019) reported a cutoff of ≤ 23 best discriminated those who would have onset of Dementia in the next 1 to 5 years (incidence of Dementia next 1 to 5 years of 11.8%) from those who did not (i.e. an absence of the onset of Dementia in the next 5 years). TICSm: PPV = 28%, NPV = 95% against Gold Standard Clinical Diagnosis.

We applied this cutoff to the TICSm and the CogAI and measured classification agreement between the two.



	TICSm-Au ≤ 23 Predict Dementia next 1 to 5 years	TICSm-Au > 23 Predict No Dementia onset next 1 to 5 years	
CogAI > 23 Predict No Dementia onset next 1 to 5 years	7	60	67
CogAI ≤ 23 Predict Dementia next 1 to 5 years	28	5	33
	35	65	100

Classification Agreement = 88% (Very good agreement)

Kappa = 0.733 (Substantial Chance Corrected Agreement – Cohen 1960)

CogAI with TICSm: PPV = 85%, NPV = 90%

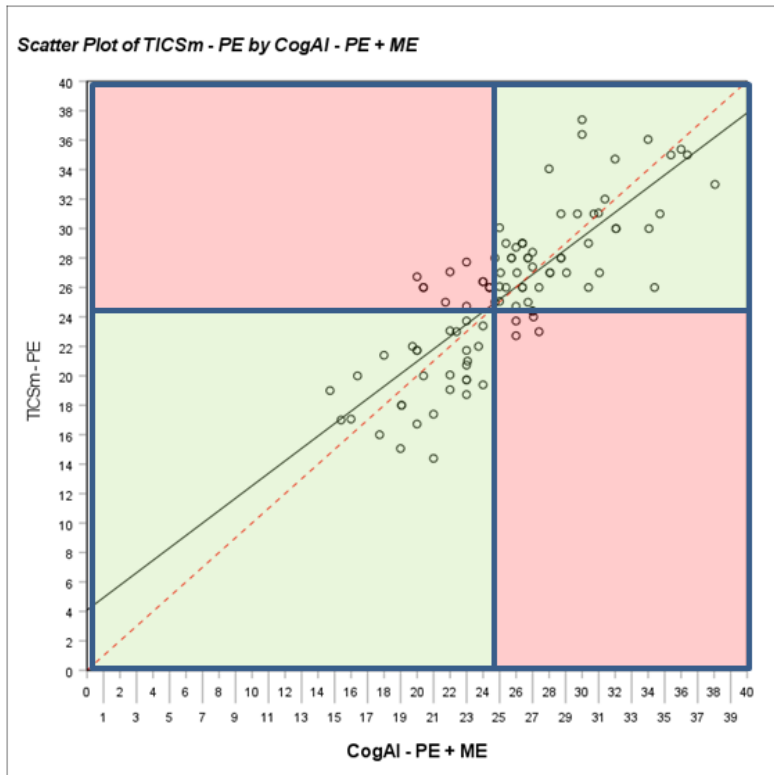
CogAI with Gold Standard Dx: Meta PPV = 24%, Meta NPV = 86%

Classification Agreement

MCI onset in next 1 to 5 years

Bentvelzen et al (2019) reported a cutoff of ≤ 24 best discriminated those who would have onset of Mild Cognitive Impairment (MCI) in the next 1 to 5 years (incidence of MCI next 1 to 5 years of 56%) from those who did not (i.e. an absence of the onset of MCI in the next 5 years). TICSm: PPV = 24%, NPV = 92% against Gold Standard Clinical Diagnosis.

We applied this cutoff to the TICSm and the CogAI and measured classification agreement between the two.



	TICSm-Au ≤ 24 Predict MCI onset next 1 year	TICSm-Au > 24 Predict No MCI onset next 1 year	
CogAI > 24 Predict No MCI onset next 1 year	8	48	56
CogAI ≤ 24 Predict MCI onset next 1 year	31	13	44
	39	61	100

Classification Agreement = 79% (Very good agreement)

Kappa = 0.569 (Moderate Chance Corrected Agreement – Cohen 1960)

CogAI with TICSm: PPV = 80%, NPV = 79%

CogAI with Gold Standard Dx: Meta PPV = 19%, Meta NPV = 73%

Classification Agreement Summary & Conclusions

Incidence	Between CogAI and TICSm ¹				Between TICSm and Gold Standard ²		Inferred Between CogAI And Gold Standard ³	
	Agreement	Kappa	PPV ¹	NPV ¹	PPV ²	NPV ²	Meta PPV	Meta NPV
Dementia 1 year	86%	0.593	75%	89%	14%	99%	10.5%	88%
Dementia 1 to 5 years	88%	0.733	85%	90%	28%	95%	24%	86%
MCI 1 to 5 years	79%	0.569	80%	79%	24%	92%	19%	73%

¹ Calculated from the current study

² From Bentvelzen et al (2019)

³ Calculated as Meta PPV = PPV¹*PPV²,
Meta NPV = NPV¹*NPV²

Meta PPV and Meta NPV are the minimum PPV and NPV of the CogAI in relation to a Gold Standard Dx

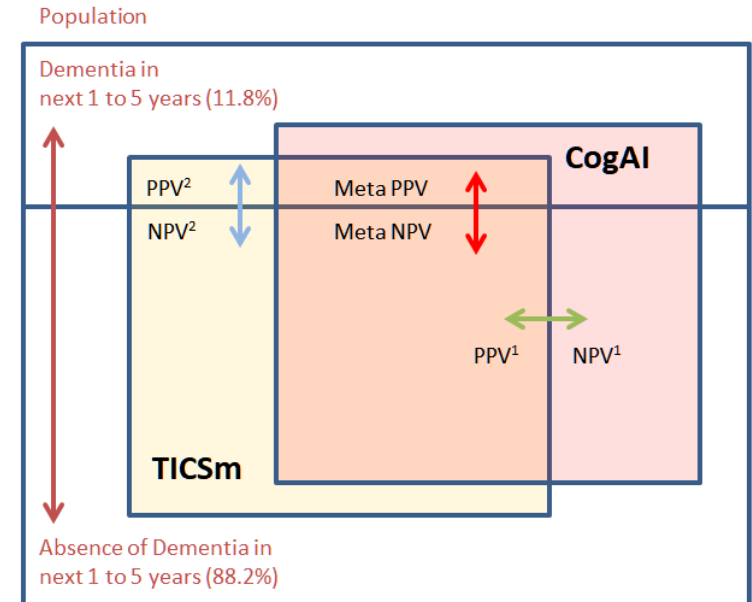
We can calculate the minimum PPV and NPV (red arrow) of the CogAI (red box) in relation to a Gold Standard Diagnosis (white areas in large box) by chaining the TICSm (yellow box) as an intermediary with Bentvelzen et al (2019) data on the PPV and NPV (blue arrow) of the TICSm (yellow box) in relation to a Gold Standard Diagnosis. Both the CogAI (red box) and the TICSm (yellow box) overlap with the population (large white box) which subdivides into two subpopulations (those with Dementia in next 1 to 5 years and those with not) and both the TICSm and the CogAI overlap with these subpopulations.

In the current study we calculated PPV and NPV values (green arrow) from data on the overlap between the CogAI and the TICSm (overlap of red box with yellow box, which creates an orange overlap box)

We then calculate:

Meta PPV as $PPV^1 * PPV^2$, Meta PPV = $85\% * 28\% = 24\%$
and Meta NPV as $NPV^1 * NPV^2$, Meta NPV = $90\% * 95\% = 86\%$

These Meta values represent minimum values for PPV and NPV of the CogAI with Gold Standard Diagnosis.



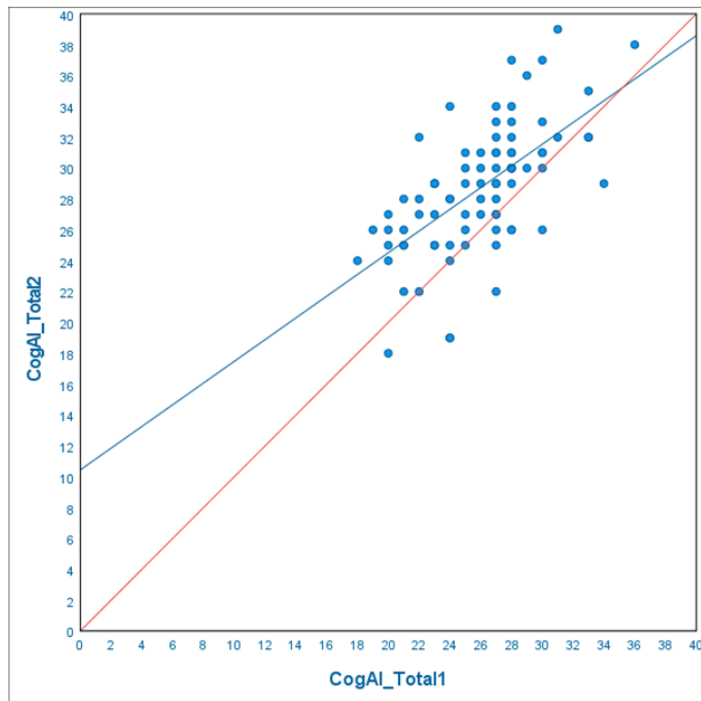
Conclusions

- The CogAI is concurrently valid when compared to the TICSm-Au
- The CogAI is functionally an alternate form of the TICSm-AU
- The CogAI is a drop-in replacement for the TICSm-Au in research and practice

Test-Retest Reliability & Agreement

Katrina Swavley

In this study the Test-Retest reliability and agreement of the CogAI was assessed by administering it twice to n = 82 participants, one week apart, and comparing the two sets of results.



Notes:

Red line (Slope = 1, intercept = 0) signifies perfect agreement between two CogAI assessments 1 week apart.

Blue line is regression ($R^2 = 0.398$, $p < 0.001$) of CogAI1 onto CogAI2 a week later – the slope (0.704) and the intercept (10.445) are significantly different full agreement.

$$\text{CogAI2} = 0.704 * \text{CogAI1} + 10.445$$

$$r = 0.63 \text{ (moderate correlation)}$$

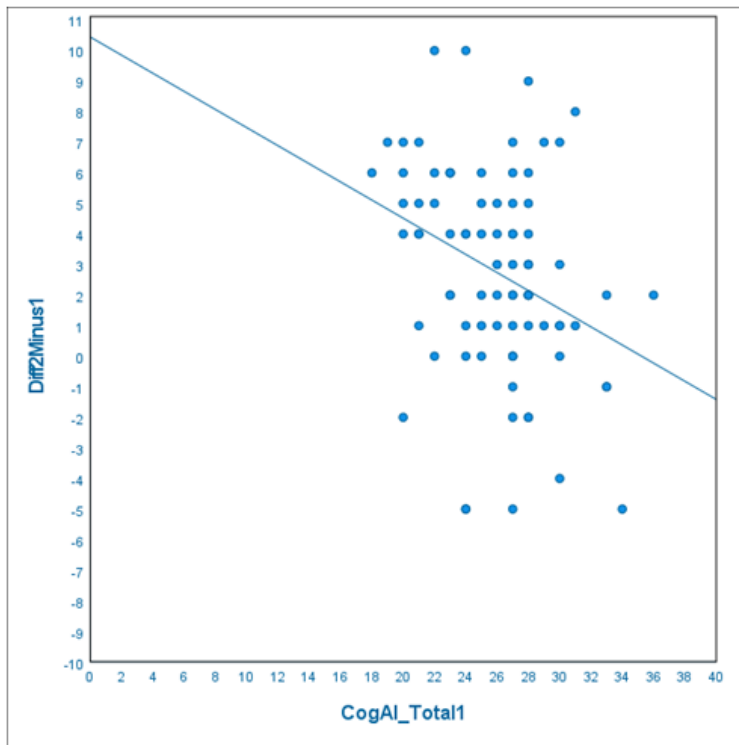
Perfect agreement is $r = 1.0$

This is poor test-retest reliability and agreement

- But Practice Effects are contaminating this result obscuring the real test-retest reliability and agreement

Note: all Raw Scores First and Second administration have been adjusted by adding 3 points to bring scores in line with TICSM. The CogAI on average scores 3 points lower than the TICSM

Measuring the Practice Effect



Notes:

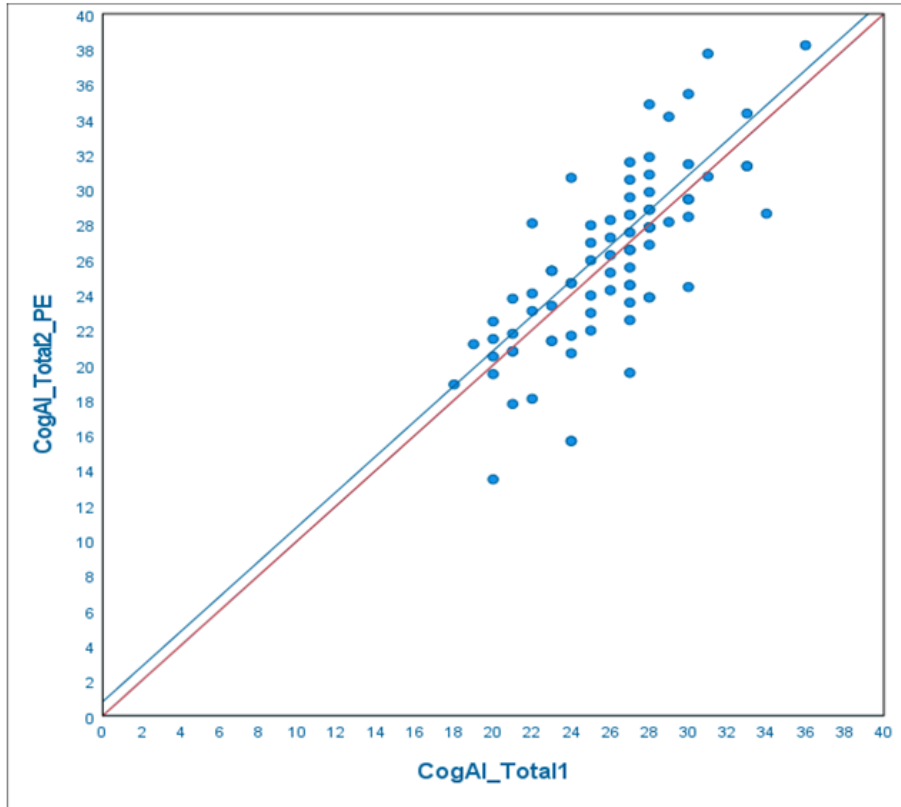
A new variable (Diff2Minus1) was created by subtracting CogAI1 from CogAI2. This measures the difference in score over 1 week and is composed of:

A Practice Effect (PE) + Error

From Regression ($R^2 = 0.105$, $p = 0.003$) the Practice Effect (PE) that applies to CogAI2 for each individual case can be estimated as:

$$PE = -0.296 * CogAI\ 1 + 9.555$$

Test-Retest Agreement



Notes:

Red line (Slope = 1, intercept = 0) signifies theoretical perfect agreement between the two CogAI assessments, 1 week apart.

A new variable (CogAI_Total2_PE) was created by subtracting PE from CogAI2. This adjusts the 2nd CogAI score by removing the Practice Effect.

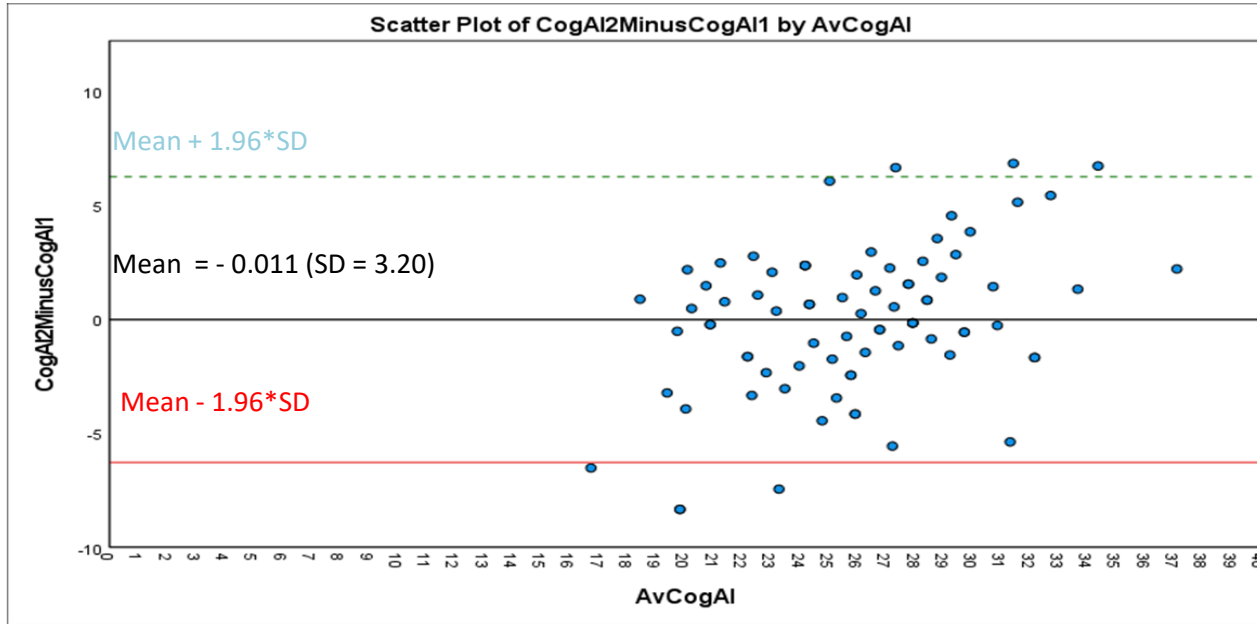
The regression ($R^2 = 0.572$, $p < 0.001$) of 1st CogAI score on to the PE Adjusted 2nd CogAI score estimates the relationship between the two to be on average in perfect agreement (the red line). The regression derived slope of the relationship between CogAI1 and PE Adjusted CogAI2 is 1 and the intercept is 0.

$r = 0.76$ (95% CI: .65 to .84)

(high correlation)

good test-retest reliability and agreement

Bland-Altman Plot



In our comparison of one test (the TICSm) with another (the CogAI), we cannot assume that one test is better than the other. The TICSm is the predecessor of the CogAI, but it is not a “gold standard” for judging the CogAI.

The Bland-Altman plot is a way of comparing two tests with each other without assuming anyone of the as the gold standard. It answers the question “how interchangeable are these two tests” by comparing the average of the two tests per individual (x axis), with the difference between the two tests per individual (y axis). In our case the mean difference between the scores was effectively zero (0.1) and the distribution of differences is even through out the scored range of the TICSm and CogAI, with 5 cases outside the 95% Confidence Interval of the difference between the TICSm and the CogAI, as expected.

This indicates the two tests TICSm and CogAI are interchangeable with no average difference in score, but with a Standard Deviation of the difference of about 3 points, and with this relationship being the same across the effective scored range of these tests.

Classification Agreement

Cog1_Classification * Cog2_Classification Crosstabulation

		Cog2_Classification		Total
		0	1	
Cog1_Classification	0	Count: 40	7	47
		% of Total: 48.8%	8.5%	57.3%
1		Count: 7	28	35
		% of Total: 8.5%	34.1%	42.7%
Total		Count: 47	35	82
		% of Total: 57.3%	42.7%	100.0%

Percentage Agreement is
 $48.8\% + 34.1\% = 82.9\%$

Kappa = 0.65 ($p < .001$)

Which is substantial agreement.

Conclusions

- The CogAI has good Test-Retest Agreement

$r = 0.76$ Test-Retest regression: slope = 1, intercept = 0

- Classification Agreement using TICSm-Au cut-offs (Bentvelzen et al 2019) for Dementia in Next 5 Years is substantial, they agree on 83% of cases
- Using the formula given by Jacobson and Truax (1991), with $r = 0.76$, $SD = 4$ and 95% confidence:

$$S_{\text{diff}} = 2.77 \quad \text{and} \quad RCI = 5.4$$

Therefore, a change of 6 or more points is statistically significant change on the CogAI

Construct Validity

Alice Ao

In this set of studies, the Construct Validity of the CogAI was examined by replicating several previous studies of the TICSm and comparing the findings with those generated using the CogAI.

The dataset (n= 262) for this investigation was created by concatenating the CogAI records collected in the Validity Study (n=100) and the Test-Retest Reliability study (n=162).

The four studies were replicated in five different ways:

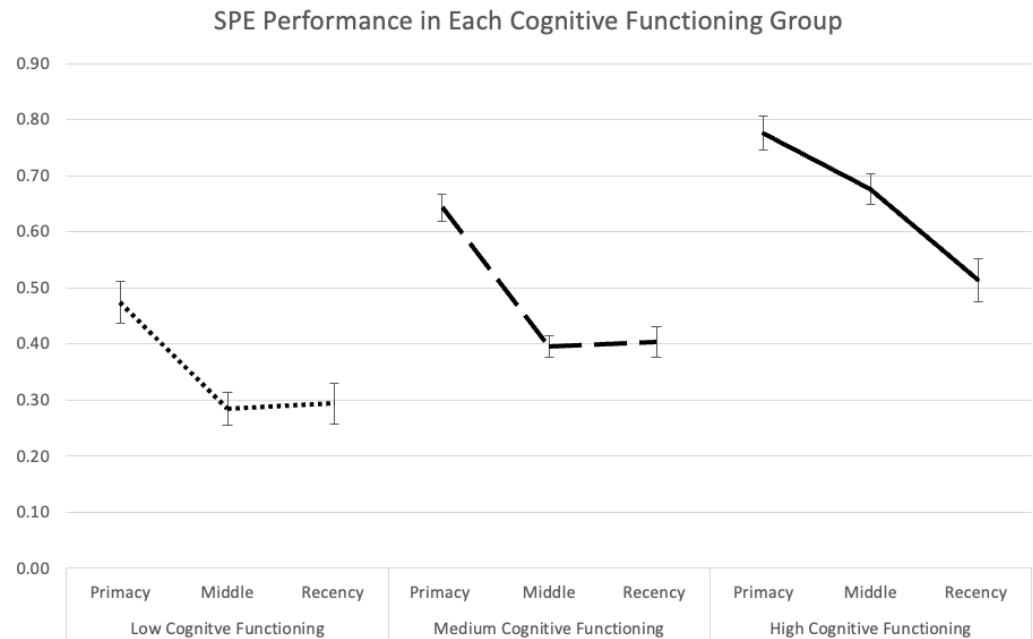
1. Factor Analysis of the TICSm (Brandt et al 1993)
2. Serial position effect in immediate recall of the word list (Harris and Dowson 1982)
3. The percentage of participants recalling each individual word in immediate recall (Harris and Dowson 1982)
4. The distribution curves for the TICSm total scores, immediate recall scores, and scores for non-word list items (Lindgren et al 2021)
5. The relationships of Age, Sex and Years of Education with TICSm-Au total score (Bentvelzen et al 2019)

Factor Analysis

Reference Study and Replication	Findings	Comment
<p>Brandt et al. (1993)</p> <p>Replicated a Principal Components Analysis with Quartimax rotation, using the CogAI in place of the TICSm.</p>	<p>Brandt et al. (1993) found 4 factors.</p> <p>The current study found 6 factors.</p> <p>Some factors were the same, but others were not.</p>	<p>The version of the TICSm used by Brandt et al. (1993) contained more items than the CogAI.</p> <p>The study was conducted as part of a large twin study and included people with cognitive disorders, which our study excluded.</p> <p>The differences could be due to differences in items between the CogAI and TICSm, or due to differences in the samples studied.</p> <p>The similarities found could be due to commonalities in items and samples where they crossed over.</p>

Serial Position Effect in Immediate Recall

Reference Study and Replication	Findings	Comment
<p>Harris & Dowson (1982)</p> <p>Replication of finding that people with cognitive disorders show impaired primacy but intact recency.</p>	<p>There were pattern differences, but impaired primacy was not found in participants categorised as lower cognitive functioning.</p>	<p>Harris and Dowson had a sample that included people with dementia and cognitive impairment, but the current study sample was largely a normal sample. The impaired recency performance observed by Harris and Dowson is due to cognitive pathology present in their sample, which was not present in the samples used in the current study.</p>



Percentage of Individuals Recalling Each Word in Immediate Recall

Reference Study and Replication	Findings	Comment
<p>Harris & Dowson (1982)</p> <p>Replication of finding that the pattern of word recall for each word was distinct – participants from all cognitive functioning levels consistently performed worse at recalling specific words.</p>	<p>Our findings showed similar patterns in individual word recall performance.</p>	<p>Despite differences in samples between Harris and Dowson (1982) and the current study, the same pattern of word recall was observed → this pattern may be normative and not altered by cognitive pathology.</p>

Figure from current study

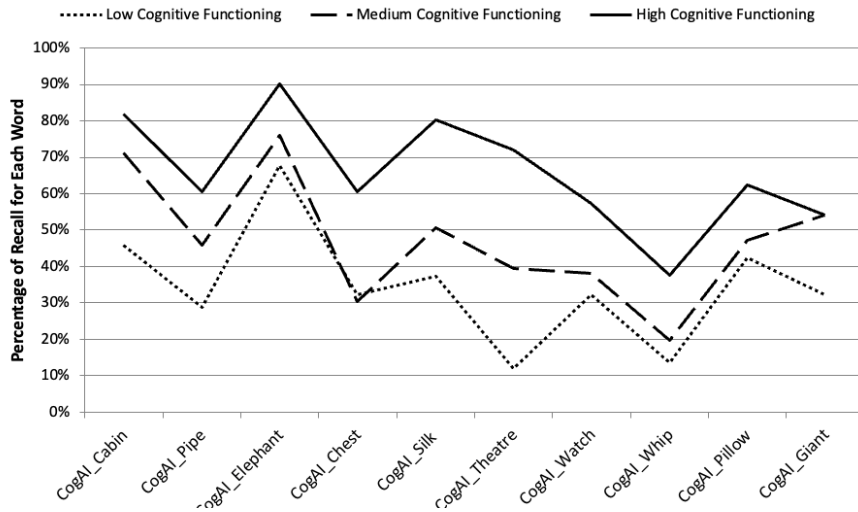
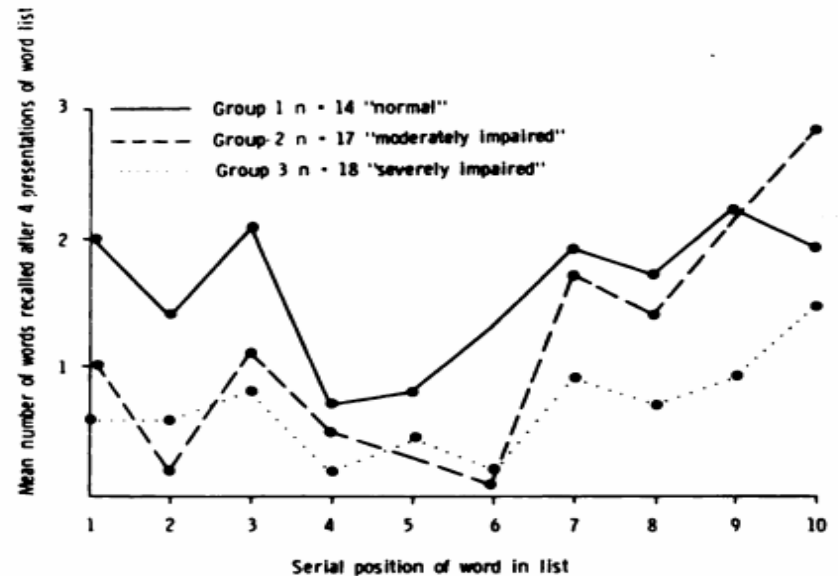


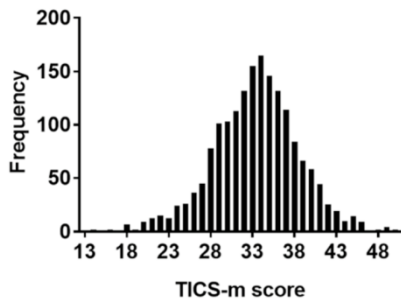
Figure from Harris and Dowson (1982)



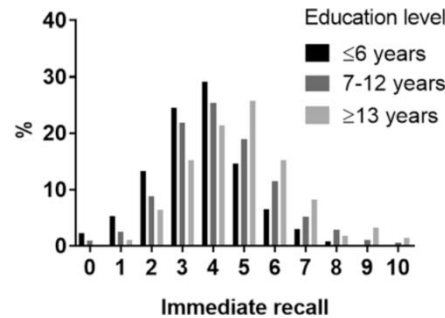
Wordlist and Non-Wordlist Score Distributions

Reference Study and Replication	Findings	Comment
<p>Lindgren et al. (2019)</p> <p>Replication of distributions for word list and non-word list items found on the TICSm.</p>	<p>The current study replicated both the normal distribution for the TICSm total score and the immediate recall score and the negatively skewed distribution of the scores on the non-word list items observed by Lindgren et al. (2019).</p>	<p>The pattern of findings is the same in both Lindgren et al. (2019) and the current study.</p>

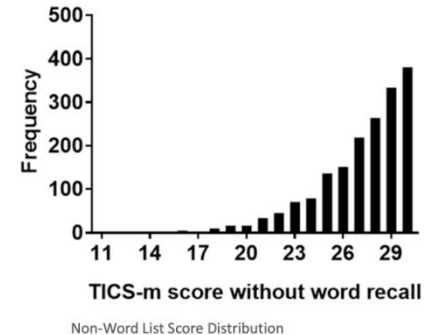
TICSm Total Score Distributions



Immediate Recall Distributions

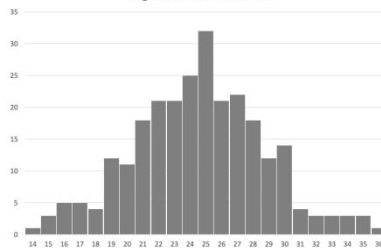


Non-Word List Items Distribution

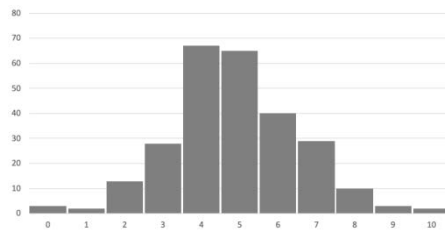


Lindgren et al (2019) distributions

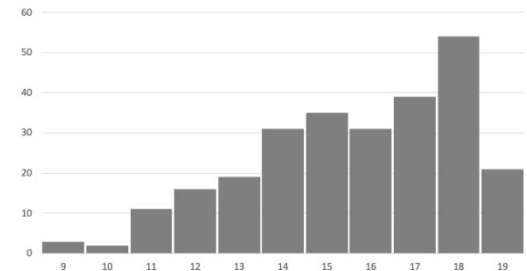
CogAI Total Score Distribution



Immediate Recall Score Distribution



Non-Word List Score Distribution



Current study's distributions

Age, Sex & Years of Education

Reference Study and Replication	Findings	Comment
<p>Bentvelzen et al. (2019)</p> <p>Replication of finding that Age, Sex, and Years of Education correlate with TICSm-Au total score.</p>	<p>Age, Sex, and Years of Education did not correlate with the CogAI total score either as a group or individually.</p>	<p>The age ranges between the current study (60-80 years) and Bentvelzen et al. (2019; 71-91 years) differed.</p> <p>We initially hypothesized that we would replicate the findings, but when we consider the results in light of our younger sample, we speculate that in the younger half of our sample, there is little to no cognitive pathology present and Age, Sex, and Years of Education act as protective factors only when there is pathology to protect against.</p>

Conclusions

- Some findings were replicated
- But others were not
- Non-replication of findings seems largely related to sample differences, in particular age, which is a proxy for cognitive pathology
- Some of the non-replicated findings seem to be due to the absence of cognitive pathology, which needed to be present in the study sample to observe the non-replicated effect.
- Our study sample was relatively young and explicitly excluded cognitive pathology.
- The fact that some findings in other studies using the TICSm, where not replicated using our CogAI dataset, maybe due to TICSm vs CogAI differences or they maybe due instead to the non-replicated effects being dependent upon the presence of cognitive pathology in the sample being studied.

Last Word - CogAI

From the set of studies presented here we conclude that the CogAI is valid reliable and fit for purpose

More studies of the CogAI are planned for 2023

The emergence of Psychobots, autonomous psychological services delivered by Conversational AI, present new opportunities to increase the amount service provision, to increase the number of different kinds of services provided and to invent new services that have previously not existed

- Next year the CogAI will be rolled out as a test that can be ordered by a clinician such as a GP
- For self-referred clients, only from defined geographic areas, the CogAI will refer to a linked memory clinic(s) in their area, post assessment if required.
- We are now developing a new Psychobot, an Conversational AI coach, to work with people on a their Cognitive Wellness and in Dementia Prevention. **3rdAge.com.au** is a new program is modelled on the successful Fingers program. Fingers uses human coaches to coach around diet, exercise, cardiovascular health, cognition and socialisation. Like its cousin the **CogAI**, **3rdAge** is an innovative strategy that uses AI to replace humans but maintains all other aspects to increase service provision and lower service cost.

For further information about CogAI or our research work in Clinical Conversational AI in general, contact

Tony Florio – Florio Research P/L Ph. 0419 281344 e. tony@florio.com.au



FLORIO RESEARCH
Software and Computer Systems Development
Statistics, Artificial Intelligence and Research



CogAI

The Team



Anastasia Serafimovska



Katrina Swavley



Alice Ao



Dr Kirsten Challinor



Dr Megan Willis



Dr Tony Florio

