Review

35 years of computerized cognitive assessments of aging – where are we now?

Avital Sternin\* 1, Adrian M. Owen 1,2 and Alistair Burns 3

1 Brain and Mind Institute, Department of Psychology, University of Western Ontario

2 Department of Physiology and Pharmacology, University of Western Ontario

3 Affiliation 3

**\*** Correspondence: avital.sternin@uwo.ca

Received: date; Accepted: date; Published: date

**Abstract:** 200 words

**Keywords (3-10):** Computerized cognitive assessment; aging; dementia;

1. Introduction

The goal of this paper is to provide an overview of how computerized cognitive assessments started and how they have changed since they were first designed in the 80s with the advent of the internet and technologies like small, portable touch screens (iPads). The differences between the existing and commonly used paper-pencil test will be discussed with an emphasis on why computerized tests are particularly advantageous for assessing cognitive changes associated with aging.

Specify difference between cognitive tests for diagnostic purposes, and assessments for monitoring changes over time. In this paper we’re talking about the latter.

2. Cognitive Assessment – Historically

The computerization of cognitive assessments began in the 1980s with the development of personal computers. New cognitive assessments were developed with the new emerging technologies (such as the mouse or touchscreen) in mind.

*There’s more to computerizing cognitive assessments than the digitization of paper-and-pencil tests, although that might be where it started*. *The idea of a well designed computerized cognitive assessment is the harnessing of the many advantages of computers in order to ultimately do better than the existing tests (more efficient, score on the fly, etc).*

*Basic practical advantages to testing on a computer/iPAD but also there are things you can’t get from paper: large N, efficiency at getting to a score (don’t have to go through all the steps, can jump to an appropriate level), guided practice is controlled, etc.). Interpretation of scores is more objective – based on stats without experimenter bias*

Experimentally, computerized testing procedures have many practical and theoretical advantages. For example, response latencies can be measured to millisecond accuracy and are also more flexible than traditional neuropsychological techniques with data being recorded, processed, and scored automatically during the testing session. Moreover, computer systems have greatly improved the accuracy with which the testing situation can be controlled reducing experimental noise such as experimenter bias as well as the consistency across testing sessions and subjects. Computerized tests can accurately record and report on many aspects of performance simultaneously providing a deeper look into the behaviour being assessed.

Another advantage of computerized assessments is that they can be designed so that each testing session is administered in a more systematic and objective way than can be done in an in-person assessment. For example, task difficulty can be automatically adapted to a subject’s abilities tailoring the test and shortening its length. In a similar way, predefined criteria can dictate the maximum number of a subject’s successes and failures such that the subjective experience of being test is equivalent across subjects. Ideally, a computerized assessment is so well controlled that it does not require the presence of a trained individual and instead can be self-adminstered.

With these advantages, computerized neuropsychological assessments are a valuable tool for investigating small behavioural changes which may not be detected using conventional methods. In particular, such computerized tests can be used in assessing and following subtle cognitive changes in aging over the long term with the hopes of catching evidence of mild cognitive impairments as early as possible. *(Citing Blackwell et al, Chapter 5 from Dementia textbook)*

The best example of the earliest computerized cognitive testing batteries is the Cambridge Neuropsychological Test Automated Battery (CANTAB). The CANTAB battery was designed to capitalize on the developments of computing technology for the neuropsychological assessment of neurodegenerative diseases. This test battery was standardized in nearly 800 older adult participants (Robbines et al, 1994) and early experiments indicated the comprehensive nature of the CANTAB tests were sensitive to deficits and the progressive decline in patients over the course of both Alzheimer’s and Parkisonian dementia (Shakian & Owen, 1992, Sahakian 1988,1990,Downes 1989,Morris, 1988, Swainson et al 2001). And results from this battery were highly successful in predicting the development of dementia in preclinical populations while also differentiating between disorders such as Alzheimer’s, Frontotemporal, and Parkinsonian dementia (Swainson et al 2001, Blackwell et al 2003, Lee et al 2003).

Although the literature indicates the usefulness of computerized cognitive tests in the early detecting and monitoring of neurodegenerative disorders, there are many different testing batteries and little consensus on which are the most effective and suitable for this task. Two recent reviews of computerized cognitive tests suitable for use with aging populations described 17 of these batteries (Wild et al 2008, Zygouris et al, 2014). The consensus across these reviews is that although the tests were judged to be valid for testing aging populations many of these batteries were faced with a number of shortcomings. For example, many of them were tested in small samples sizes or samples that lacked data specific to older adults, some batteries had only small amounts of normative data, weak reports of reliability, the ‘quality assurance’ (i.e. are computerized versions of paper-pencil tests measuring the same psychometric properties as expected). Ultimately, both reviews suggest that the usefulness of each battery must be assessed on a case-by-case basis and that it is not yet possible to select a single test or battery for the reliable screening and monitoring of cognitive impairment.

Computerized cognitive assessments have come a long way in the last 40 years, but their adoption into health care systems has been slow. A lack of normative data plagues many of these testing batteries and may be responsible for the hesitancy of clinicians to adopt any single set of tests as a screening and/or monitoring tool (Barnett et al, 2015). In order for progress to be made, the issue of normative data must be resolved. The only way to do this is to greatly increase the number of subjects who complete the computerized tests. This may not be feasible when these assessments must be taken within a lab testing environment, however the advancement of the self-administration of computerized testing via the internet has opened up many new possibilities.

*Bigger Ns needed – clinicians aren’t adopting the more efficient testing strategies*

3. Cognitive Assessment – In the internet age

*Internet age can give us bigger Ns (get info on data like age, SES, etc) calculate things like meaningful change*

The internet and proliferation of portable computers provides new opportunities for computerized cognitive assessments.

An illustrative example of such an online tool is the Cambridge Brain Sciences (CBS) battery. The tests in this battery are based on well validated neuropsychological tests but have been adapted and designed to capitalize on the numerous advantages that internet-based testing can offer. Additionally, the relationship of the tests in this battery to independent cognitive functions have been assessed using a factor analysis based on the data of nearly 45 000 participants (Hampshire et al, 2012) validating the use of this battery to assess such functions.

The CBS battery has been used in several large-scale population-based studies involving tens of thousands of participants (Owen et al., Nature, 2010; Hampshire et al., Neuron, 2012), as well as more than 300 bespoke studies by researchers from all over the world (e.g. Metzler-Baddeley et al., 2016; Pausova et al., 2015; Esopenko et al., 2017). The tests are deployed through a custom-built research platform for hosting and conducting scientific trials. In total, more than 8 million tests have been taken and normative data from 75,000 healthy participants is available. As a result of this battery’s wide reach and easy-to-use interface, there are also normative data from approximately 5000 adults over the age of 65. Figure from Conor showing hundreds of older adult scores on 12 different tests.

The CBS battery is also designed to address many of the issues raised by Wild et al 2008 and Zygouris et al, 2014. It can cover a wide range of assessments in a short period of time, has large banks of normative data across the entire life-span (youngest-oldest), strong reliability scores (cite), corroborating neuroimaging evidence, and has built in mechanisms for providing meaningful feedback to the participant or physician (see example CBS score report). This score report includes a Meaningful Change indicator, which objectively determines whether the current assessment result represents a potentially significant change relative to two time-points: the patient’s previous assessment results and the patient’s baseline (e.g., initial) assessment results. The simplification and interpretability of results addresses the issue raised by ONE OF THE REVIEWS regarding physician’s not being willing to wade through the complicated data output of computerized testing batteries to interpret their meaning. Off-loading the task of interpreting test results into meaningful change to the test battery increases the likelihood of physicians being amenable to adopting it for monitoring/screening.

*Objective interpretation of scores (in relation to a human’s interpretation) – based on numbers/stats*

The added benefit of the battery translating score results into meaningful change measures is that they can be understood by non-physicians increasing the likelihood of the tasks being adopted into the health care system. In this way, any health care provider or family member can monitor an aging patient’s cognitive changes over time, the effect of drugs, or even cognitive changes post-surgically. When a significant meaningful change is seen, this alerts the caregivers or health care providers that more in-depth testing may be necessary to assess the individuals cognitive status. This has relevance in home care, assisted living facilities, and in hospital settings for reducing the administration burden of monitoring cognitive changes while increasing the sensitivity of testing and catching important changes early enough to be appropriately addressed.

**Experimental Section – present tense/figures**

*Issues of internet testing: 2 studies showing that it’s just as good.*

One of the obvious questions regarding at-home testing, is its validity in comparison to in-lab testing. In YEAR, 19 healthy young adult control participants completed all 12 CBS tests both at home an in-lab (order counterbalanced across participants). The mean standardized scores for each of the tests showed no significant effect of at home versus in-lab testing (p>0.1) and the tasks showed reliable correlations within subjects across the two testing environments (p<0.05). A follow-up study explored whether the stability in scores across testing environments was applicable to patients groups as well. 27 participants with Parkinson’s disease were tested on 5 of the 12 tests at home and in-lab (order counterbalanced across participants). Again, there was no significant effect of at home versus in-lab testing (p>0.1) and the tasks showed reliable correlations across the two testing environments (p<0.05).

*Validity problem – discuss CBS validity indicator (generally, without mention of CBS) – worry of testing people in unsupervises situations*

*Other mechanisms to ensure reliable data – guided learning/tutorials*

4. Online testing vs existing alternatives - MoCA/MMSE

The ability to quickly, and accurately assess changes in cognitive functioning on a regular basis has implications for quality of life, level of independence, and degree of care in the aging adult population. Currently, assessments like the Mini-Mental Status Exam (MMSE) and the Montreal Cognitive Assessment (MoCA) are used by health care providers to monitor for cognitive changes and screen for deficits. Although these tests are useful because they are short, and easy to administer, there are some downsides to using these paper-and-pencil tests: they are not adaptive to an individual’s ability level which can lead to frustration in individuals with deficits, the questions are not randomly generated each time leading to practice effects, and they must be administered by another human introducing testing bias and taking time and resources away from other health care duties. Additionally, the ability of these paper-pencil tests to detect fine grained changes in cognition are small and instead these tests group patients into broad categories (impaired, unimpaired, or borderline). Finally, the cutoff scores used in these tests are suggested to not be appropriate for aging populations (cite, cite) and result in larger numbers of patients being labeled as impaired than perhaps is necessary.

Recent studies have investigated the use of short computerized assessments for monitoring cognitive changes over time in try to better differentiate the degrees of abilities in older adult populations. Brenkel et al (2017) asked 45 older adults to complete a MoCA and a small subset of a computerized battery (CBS) to determine whether the addition of computerized tests added to the MoCA’s ability to differentiate impaired from unimpaired participants. They found that the addition of two computerized tests to the MoCA testing, better differentiated participants into impaired or unimpaired categories and provided additional details about their cognitive capacities. A follow-up study investigated whether the MoCA and MMSE could simply be replaced by computerized cognitive assessments and found that a short battery consisting of three CBS tasks was as good or better than the MoCA at differentiating impaired from unimpaired participants (Sternin et al, Under Review).

These are the first examples of exploring the ability to monitor cognitive changes in older adults using a self-administered, computerized testing battery. Although further testing in this direction is required, there is good reason to believe that a shift towards computerized cognitive testing in health care may occur in the near future.

5. Neural validation ?

*Given 35 years – time to validate these tests neutrally – from the beginning these tests were neutrally validated based on coinciding timelines*

*Paper and pencil tests were validated with brain lesion studies (Owen 1990s papers, Milner/HM)*

*The computerization of cognitive tests coincided with a growth in imaging so these tests could be validated with imaging – talk about CANTAB and CBS*

* Relationship between imaging and cognition
  + Hampshire et al 2012 – CBS in scanner
  + neural validation of CANTAB work – Owen et al 1995 *Neuropsychologia* – CANTAB and temporal lobe excisions
  + Huntley et al, 2017 *British Journal of Psychiatry –* training adaptive working memory strategies (trained on digit span) improved general cognitive function (MMSE)– with MRI correlates
  + Robbins et al 2004

6. Conclusion/Summary

*Evidence based vs intuition based (neuropsych)*

Computerized cognitive assessments have come a long way in the past 25 years. Proliferation of technology has made it cheaper and more accessible.

Provide an ‘ideal’ for what computerized cognitive assessments could look like: self-administered, self-monitoring, used by retirement homes/family physicians/other front-line health care workers to catch small changes in cognition early in order to identify early the supports that might be helpful (quality of life implications). *Using some of the keywords from previous reviews/papers*.

Working with physicians/health care providers to monitor early changes in aging/dementia will also create a large amount of data about what kinds of cognitive changes are happening during the aging process. Has broader implications for research that could impact the design of policies and health care support networks.

\*\*\* STORING OTHER THINGS HERE FOR NOW:

When cognitive assessments can be done at home without the presence of an administrator the opportunities for normative data collection, widespread use of screening tools, and ongoing health monitoring increase drastically.

Corbett et al 2015 - Computerized cognitive training in older adults to combat cognitive decline that occurs as a function of age. Test groups show some improvement in IADL than controls

Ferreira et al 2015 – playing leisure games like Sudoku has benefits for cognition

There are many different testing batteries but very little consensus on effectiveness and suitability of these batteries for monitoring/assessing cognitive impairment.

* Wild, K et al 2008, *Alzheimers & Dementia* – systematic review of 11 computerized cognitive testing batteries appropriate for use in the elderly. Conclusion: batteries are not consistent and the utility of these batteries needs to be assessed on a case by case basis
* Zygouris & Tsolaki, 2015 *American Journal of Alzheimer’s Disease* – Review of computerized cognitive testing batteries appropriate for use in the elderly. Discusses inconsistencies in the research on utility of these batteries in dementia. Also discusses things that are missing from existing computerized cognitive testing batteries (e.g. Virtual reality, adaptability, length of tests)

**Supplementary Materials:** The following are available online at www.mdpi.com/xxx/s1, Figure S1: title, Table S1: title, Video S1: title.

**Author Contributions:** For research articles with several authors, a short paragraph specifying their individual contributions must be provided.

**Funding:** Please add: “This research received no external funding” or “This research was funded by NAME OF FUNDER, grant number XXX” and

**Acknowledgments:** In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

**Conflicts of Interest:** The online cognitive tests (Cambridge Brain Sciences) discussed in this review are marketed by Cambridge Brain Sciences Inc, of which Dr. Owen is the unpaid Chief Scientific Officer. Under the terms of the existing licensing agreement, Dr. Owen and his collaborators are free to use the platform at no cost for their scientific studies and such research projects neither contribute to, nor are influenced by, the activities of the company. As such, there is no overlap between the current review and the activities of Cambridge Brain Sciences Inc, nor was there any cost to the authors, funding bodies or participants who were involved in the mentioned studies.

Appendix A

The appendix is an optional section that can contain details and data supplemental to the main text.

Appendix B

All appendix sections must be cited in the main text. In the appendixes, Figures, Tables, etc. should be labeled starting with ‘A’, e.g., Figure A1, Figure A2, etc.

References

References must be numbered in order of appearance in the text (including citations in tables and legends) and listed individually at the end of the manuscript. We recommend preparing the references with a bibliography software package, such as EndNote, ReferenceManager or Zotero to avoid typing mistakes and duplicated references. Include the digital object identifier (DOI) for all references where available.

Citations and References in Supplementary files are permitted provided that they also appear in the reference list here.

In the text, reference numbers should be placed in square brackets [ ], and placed before the punctuation; for example [1], [1–3] or [1,3]. For embedded citations in the text with pagination, use both parentheses and brackets to indicate the reference number and page numbers; for example [5] (p. 10), or [6] (pp. 101–105).

1. Author 1, A.B.; Author 2, C.D. Title of the article. *Abbreviated Journal Name* **Year**, *Volume*, page range.
2. Author 1, A.; Author 2, B. Title of the chapter. In *Book Title*, 2nd ed.; Editor 1, A., Editor 2, B., Eds.; Publisher: Publisher Location, Country, 2007; Volume 3, pp. 154–196.
3. Author 1, A.; Author 2, B. *Book Title*, 3rd ed.; Publisher: Publisher Location, Country, 2008; pp. 154–196.
4. Author 1, A.B.; Author 2, C. Title of Unpublished Work. *Abbreviated Journal Name* stage of publication   
   (under review; accepted; in press).
5. Author 1, A.B. (University, City, State, Country); Author 2, C. (Institute, City, State, Country). Personal communication, 2012.
6. Author 1, A.B.; Author 2, C.D.; Author 3, E.F. Title of Presentation. In Title of the Collected Work (if available), Proceedings of the Name of the Conference, Location of Conference, Country, Date of Conference; Editor 1, Editor 2, Eds. (if available); Publisher: City, Country, Year (if available); Abstract Number (optional), Pagination (optional).
7. Author 1, A.B. Title of Thesis. Level of Thesis, Degree-Granting University, Location of University, Date of Completion.
8. Title of Site. Available online: URL (accessed on Day Month Year).

|  |  |
| --- | --- |
| copyRight | © 2019 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). |