Computers in Human Behavior 30 (2014) 199–205

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Reliability and usability of an internet-based computerized cognitive testing battery in community-dwelling older people

D.G. Darby [a,b,c,*], J. Fredrickson [b], R.H. Pietrzak [d], P. Maruff [b,c], M. Woodward [e], A. Brodtmann [a]

[a] Florey Institute of Neuroscience and Mental Health, University of Melbourne, Parkville, Vic. 3010, Australia
[b] CogState Ltd, Level 2, 255 Bourke St, Melbourne, Vic. 3000, Australia
[c] Centre for Neuroscience, University of Melbourne, Parkville, Vic. 3010, Australia
[d] Department of Psychiatry, Yale University School of Medicine, 300 George St, Suite 901, New Haven, CT 06511, USA
[e] Aged and Residential Care/Neurosciences Clinical Service Unit, Austin Health, University of Melbourne, West Heidelberg, Vic. 3081, Australia

ARTICLE INFO

Article history:
Available online 11 September 2013

Keywords:
Computerized testing
Elderly
Cognition
Usability
Screening
Community based

ABSTRACT

Cognitive decline is an early feature of neurodegenerative conditions. CogState has developed a game-like computerized test battery with demonstrated acceptability, validity, reliability, stability, efficiency and sensitivity to detecting cognitive decline in older people under supervised conditions. This study aimed to evaluate an internet-based version of this test when used remotely and self-administered in a cohort of healthy, community-dwelling older adults aged 55 and above over a 12 month period at 1–3 monthly intervals. Test usability and reliability was examined in terms of acceptability, stability and reliability. Of 150 participants (age: 63.6 ± 5.6, range 55–83 years), 143 (95%) successfully completed a valid baseline test. Of these, 67% completed 3 month and 43% 12 months of testing. Technical difficulties were reported by 9% of participants. For those participants who completed 12 months tests, all tasks showed moderate to high stability and test–retest reliability.

This brief computerized test battery was shown to have high acceptability for baseline self-administered testing and moderate to high stability for repeated assessments over 12 months. Attrition was high between baseline and 3 months. These data suggest that this tool may be useful for high frequency monitoring of cognitive function over 6–12 months, and deserves further evaluation.

1. Introduction

Simple, inexpensive methods for detecting early neurodegenerative disease are urgently required (Sperling et al., 2011). In addition to the worldwide incidence of dementia increasing dramatically (2009 Alzheimer's disease facts and figures, 2009; Ferri et al., 2005; Mayeux, 2003), research efforts to develop disease-modifying therapies have been ineffective despite promising benefit in animal models (Gilman et al., 2005; Salloway et al., 2009). Deployment of such therapies in prodromal Alzheimer's disease (AD) are more likely to be beneficial, with current research approaches utilizing either groups at high risk including deterministic genetic mutation carriers (Bateman et al., 2011, 2012; Reiman & Tariot, 2010), older adults (Sperling & Johnson, 2012; Sperling et al., 2011) and those with apolipoprotein E4 status (Fleisher et al., 2012), or biomarkers sensitive to early AD pathology (Blennow et al., 2007; Fagan et al., 2007; Morris et al., 2009; Pike et al., 2007; Villemagne et al., 2010). The latter are not suitable for wide-scale screening because of: cost; exposure to ionizing radiation or level of invasiveness (Darby, Brodtmann, Woodward, Budge, & Maruff, 2011). More cost-effective instruments are required.

Screening for subtle cognitive impairment is a possible alternative (Landau et al., 2012). However, conventional cognitive testing usually requires trained supervisors overseeing testing at specified times and testing locations (Fredrickson et al., 2010). In addition, testing of large numbers of individuals introduces additional compromises in length of assessments, test selection, and staffing resources. These issues are compounded for prospective studies requiring repeated assessments (Fredrickson et al., 2010). Furthermore, metric properties of such tests are not necessarily ideal for repeated testing at short retest intervals due to range restriction, practice effects or limited sensitivity to detect subtle cognitive changes (Bartels, Wegrzyn, Wiedl, Ackermann, & Ehrenreich, 2010; Benedict & Zgaljardic, 1998; Fredrickson et al., 2010). Computerized cognitive testing can help circumvent these limitations, particularly if such tests are designed to be appropriate for repeated administration and for use in older people (Falleti, Maruff, Collie, & Darby, 2006; Fredrickson et al., 2010), though none have

* Corresponding author. Address: The Florey Institute of Neuroscience and Mental Health, Level 1, Howard Florey Building, University of Melbourne, Parkville, Vic. 3010, Australia. Tel.: +61 3 8344 1856, mobile: +61 (0)419 136 847; fax: +61 3 93495105.
E-mail address: ddarby@unimelb.edu.au (D.G. Darby).

0747-5632/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.chb.2013.08.009

been evaluated comprehensively for primary care practice (Tierney & Lermer, 2010). In addition, although techniques are described to utilize biofeedback in unsupervised brain training games (Mandryk et al., 2012), we are not aware of any prior attempts to utilize remote unsupervised internet-based remote and self-administered computer testing specifically for the detection of cognitive decline.

We have previously shown that the tasks of the CogState Brief Battery (CBB) can be used in a community setting to serially assess older people efficiently without substantial supervision (Darby, Brodtmann, Pietrzak, et al., 2011; Darby et al., 2012; Fredrickson et al., 2010). The CBB tasks assess psychomotor processing speed (DET), visual attention (IDN), working memory (OBK) and memory (OCL) (Bakker et al., 2012; Fredrickson et al., 2010). These provide a broad screening of cognition to determine whether decline affects multiple or specific cognitive domains, particularly episodic memory which is typically impaired early in Alzheimer's disease (Bakker et al., 2012; Fredrickson et al., 2010). Such evaluations were well tolerated, producing high rates of valid tests, low attrition over 12 months of 3 monthly testing, and good psychometric test–retest properties (Fredrickson et al., 2010). In addition, decline in memory on the brief memory test from the CBB was found to be predictive of accumulation of cortical amyloid (odds ratio 6–8) suggestive of prodromal AD in otherwise healthy older adults (Darby, Brodtmann, Pietrzak, et al., 2011; Darby et al., 2012). With more frequent testing (e.g. 1–3 times monthly), such decline can also be detected within 6 months with sufficient power to be considered useful in clinical trials of potential disease-modifying therapies (Lim, Pietrzak, et al., 2013). These findings suggest that serial cognitive screening in the community may be an appropriate strategy to detect persons with memory decline indicative of prodromal AD to allow them to participate in research trials (Darby, Brodtmann, Woodward, et al., 2011).

However, such supervised testing is still labor-intensive and disruptive to the routines of individuals who are required to attend at a specified testing time and site. Fixed appointment scheduling can also be at inopportune times during the day, peak traffic times, and after unexpected life events compromising mood, motivation or fatigue. Anxiety levels at initiation of such screening is also predictive of later decline (Pietrzak et al., 2012), suggesting complex relationships between optimal testing circumstances and test performance. Hence, it is possible that fixed scheduling and site testing introduces further variability in test performance that may obscure even more subtle changes in cognition.

The recent advent of an internet-based version of the CBB, which is equivalent in presentation to prior supervised versions, raises the prospect of home-based remote self-administered testing (Cockayne et al., 2011). This might allow more flexible scheduling at a time the individual deems optimal, and in familiar settings may provoke less anxiety and more reliable performance. Although it might alleviate the requirement of trained supervisors and site attendance, as well as provide the opportunity for even higher frequency testing, these advantages may be outweighed by increased attrition, failure to complete or perform valid tests, or greater test–retest variability.

In this study, we aimed to evaluate the acceptability, validity, and stability of remote self-administered computerized cognitive testing using the CBB in older people who were part of our prior supervised testing cohort (Fredrickson et al., 2010). We chose participants who were experienced with the CBB, having already performed it multiple times over the prior 12 months (Fredrickson et al., 2010). The first aim was to determine whether experienced participants in an unsupervised setting could complete a baseline test that was satisfactory for future comparisons. This was determined by the proportion of participants who completed baseline tests and satisfied previously described performance integrity criteria (Fredrickson et al., 2010; Moriarity et al., 2012). The second aim was to determine the acceptability of high frequency testing over the minimum likely time required to detect declining memory. This was evaluated by the attrition rate over 6 and 12 months of serial testing. The third aim was to determine the stability of testing. To determine this, we compared performance between visits and estimated the within-individual variation for each task across the study period. Hence, we proposed that baseline, 3, 6 and 12 month acceptability and test–retest stability would be high indicating that the tests were potentially useful serial screening measures using remote self-administration in older people at risk for neurodegenerative disease.

2. Materials and methods

2.1. Participants and recruitment

Participants were recruited from the previously described Melbourne community screening study (Fredrickson et al., 2010), and offered participation if they could provide a contact email address and were willing to attempt monthly unsupervised testing using a web-based version of the same battery for a year. Inclusion criteria from the original study were: (1) age 50 years or over; (2) willingness to nominate a current medical practitioner to be kept informed of study participation and results; (3) willingness to be informed of their results during the study and to accept the potential risk that participation in the study may demonstrate their cognitive performance is impaired or declining, and (4) ability and willingness to provide informed consent. The exclusion criteria were based on self-report and included: (1) any known significant cognitive impairment due to neurological or medical disease; (2) any other condition that might make it difficult for them to complete 12 months of testing; (3) uncorrectable impaired visual acuity preventing discrimination of visual changes on a computer screen; (4) physical handicap or condition preventing effective use of a computer keyboard or mouse, and (5) unwillingness to undergo testing using a computer. The study design was approved by the institutional ethics committee of the University of Melbourne and all participants gave written informed consent to participate.

2.2. Study design

This was a prospective community-based longitudinal, observational study (Darby, Brodtmann, Pietrzak, et al., 2011; Darby et al., 2012; Fredrickson et al., 2010; Pietrzak et al., 2012). Each participant was given a unique study id and testing web site address. No other measures were used to check the identity of each participant each time they logged in. On the first day of each month, an email was sent to all participants requesting they repeat the test. For the first test, participants were encouraged to do a "Practice" test to remind them of the task requirements. The practice test was exactly the same as the full test. If they failed to complete a monthly testing, they were encouraged to rejoin in subsequent testing sessions. There were 13 requests for testing between November 2010 and November 2011. Participants could individually contact the study coordinator via email with specific questions (e.g. if they forgot their login information). Reasons for having trouble with testing were reported directly to the study coordinator who would follow-up by email or telephone. A single coordinator managed the study including email notifications, responses and phone calls. No feedback was given to participants unless requested by a participant or their medical practitioner. If they did request a report, it was generated automatically by the computer system and compared the performance of the participant to normative data. If a participant expressed concern about their memory or results to the study coordinator, they were offered initial