RBANS Embedded Measures of Suboptimal Effort in Dementia: Effort Scale Has a Lower Failure Rate than the Effort Index

Rachel L. Burton1, Joe Enright1, Megan E. O’Connell1,*, Shawnda Lanting2,3, Debra Morgan4

1Department of Psychology, University of Saskatchewan, Saskatoon, SK, Canada S7N 5A5
2Department of Neuropsychology, Copeman Health Clinic, Vancouver, BC, Canada V6Z 2L4
3School of Kinesiology, University of British Columbia, Vancouver, BC, Canada V6T 1Z1
4Canadian Centre for Health and Safety in Agriculture, Saskatoon, SK, Canada S7N 5E5

*Corresponding author at: Department of Psychology, University of Saskatchewan, 9 Campus Drive, Saskatoon, SK, Canada S7N 5A5.
Tel.: +1-306-966-2496; fax: +1-306-966-6630.
E-mail address: moconnel@uvic.ca (M. E. O’Connell).
Accepted 12 November 2014

Abstract

The importance of evaluating effort in neuropsychological assessments has been widely acknowledged, but measuring effort in the context of dementia remains challenging due to the impact of dementia severity on effort measure scores. Two embedded measures have been developed for the repeatable battery for the assessment of neuropsychological status (RBANS; Randolph, C., Tierney, M. C., Mohr, E., & Chase, T. N. (1998). The repeatable battery for the assessment of neuropsychological status (RBANS): Preliminary clinical validity. Journal of Clinical and Experimental Neuropsychology, 20 (3), 310–319): the Effort Index (EI; Silverberg, N. D., Wertheimer, J. C., & Fichtenberg, N. L. (2007). An effort index for the repeatable battery for the assessment of neuropsychological status (RBANS). Clinical Neuropsychologist, 21 (5), 841–854) and the Effort Scale (ES; Novitski, J., Steele, S., Karantzoulis, S., & Randolph, C. (2012). The repeatable battery for the assessment of neuropsychological status effort scale. Archives of Clinical Neuropsychology, 27 (2), 190–195). We explored failure rates on these effort measures in a non-litigating mixed dementia sample (N = 145). Failure rate on the EI was high (48%) and associated with dementia severity. In contrast, failure on the ES was 14% but differed based on type of dementia. ES failure was low (4%) when dementia was due to Alzheimer disease (AD), but high (31%) for non-AD dementias. These data raise concerns about use of the RBANS embedded effort measures in dementia evaluations.

Keywords: Dementia; Sub-optimal effort; RBANS; Validity indicators; Malingering; Geriatrics

The critical importance of including formal measures of effort in neuropsychological assessments cannot be understated. The widely cited 2007 position paper by the National Academy of Neuropsychology concluded that it is necessary to include an assessment of symptom validity in all neuropsychological assessments that are carried out for medical reasons, a position reiterated by the American Academy of Clinical Neuropsychology (AACN, 2007; Bush et al., 2005; Heilbronner, Sweet, Morgan, Larrabee, & Millis, 2009). Neuropsychological tests are sensitive to effort (Iverson & Binder, 2000), and the objective assessment of suboptimal effort during dementia assessments is as equally important to all other neuropsychological assessment contexts. Although financial secondary gain may be less common in a dementia evaluation, poor effort can, nonetheless, impact neuropsychological data due to a patient’s lack of interest, poor cooperation, opposition to testing, fatigue, lack of comprehension of the utility of the results, or motivation to be in a “sick role” (Barker, Horner, & Bachman, 2010).

Despite its clear importance, assessing suboptimal effort in the context of a dementia evaluation remains highly problematic. As Dean, Victor, Boone, Philpott, and Hess (2009) note, persons with dementia are rarely included in samples used for effort test validation. When included in studies of effort measures, individuals with dementia frequently score below suggested cutoffs for performance validity (e.g., Duff et al., 2011; Teichner & Wagner, 2004). Several studies have demonstrated a link between performance on tests of effort and dementia impairment or severity (e.g., Dean et al., 2009; Merten, Bossink, & Schmand, 2007). Clearly, this is problematic for clinicians attempting to distinguish between scores due to genuine cognitive impairment and those due to suboptimal
Validation indices, or effort tests, can be stand-alone measures or measures embedded within conventional neuropsychological tests. As Miele, Gunner, Lynch, and McCaffrey (2012) have explained, embedded measures of effort are attractive because they do not require the administration of an additional, and potentially time-consuming measure. Additionally, embedded effort indices are often derived from scores of multiple measures tapping various cognitive functions, and thus may be inherently more sensitive to inconsistent or selective effort across tests (Strauss, Sherman, & Spreen, 2006). In contrast, stand-alone measures of effort most commonly measure a single domain of function, which is usually memory, which can pose a particular problem for dementia assessments. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, Tierney, Mohr, & Chase, 1998), a brief battery of cognitive functioning, is commonly used for the assessment of dementia and, to date, two RBANS embedded measures of effort have been developed.

Silverberg, Wertheimer, and Fichtenberg (2007) developed the RBANS Effort Index (EI) using weighted scores from the Digit Span and List Recognition subtests (scores on these are compared with charts for weighted scores, see source). As Novitski, Steele, Karantzoulis, and Randolph (2012) subsequently argued, individuals with dementia were excluded from the samples used by Silverberg and colleagues (2007) to develop the EI. Studies examining the EI’s use with individuals diagnosed with dementia (Barker et al., 2010; Duff et al., 2011) and medically ill older adults (Hook, Marquine, & Hoelzle, 2009) reported high failure rates, especially in moderate-to-severe dementia, and warned against its use with these populations.

Novitski and colleagues (2012) sought to address the high failure rates of the EI by developing an embedded measure of effort that differentiates between suboptimal effort and a genuine amnestic disorder such as that found in dementia due to Alzheimer disease (AD). Their embedded RBANS effort scale (ES) was developed based on the premise that when an individual has “true” amnesia his or her free recall performance on the List Recall, Story Recall, and Figure Recall subtests will decline to zero, or close to zero, before decline in List Recognition occurs (Novitski et al., 2012). Further, they expected working memory, as measured by the Digit Span subtest, to remain relatively stable. Following this logic, Novitski and colleagues (2012) developed the following formula: RBANS $ES = \text{List Recognition} - (\text{List Recall} + \text{Story Recall} + \text{Figure Recall})$. Unlike the EI, which can be applied to any individual administered the RBANS, the ES is recommended for use only with individuals demonstrating impaired performance on the List Recognition and Digit Span subtests (Novitski et al., 2012).

Using a cutoff score of 12 (scores $<12$ were considered to suggest suboptimal effort), Novitski and colleagues (2012) found that, relative to the EI, the ES demonstrated greater sensitivity and specificity among amnestic participants diagnosed with either amnestic Mild Cognitive Impairment or probable AD. Schroeder, Peck, Buddin, Heinrichs, and Baade (2012) found that no participants in a moderate-to-severe mixed dementia sample failed the ES when using a cutoff of $<7$. These authors claimed that this lower cutoff limits false positives while maintaining adequate sensitivity, and in the current study the cutoffs of $<12$ and $<7$ are both considered.

Sieck, Smith, Duff, Paulsen, and Beglinger (2013), in a sample of individuals diagnosed with Huntington’s disease (HD), and Dunham, Shadi, Sofko, Denney, and Calloway (2014), in a sample of individual diagnosed with dementia, have compared the two embedded measures. Sieck and colleagues (2013) reported that while 82% of participants passed the EI, only 30% of participants passed the ES. They posited that the ES’s high failure rate among HD patients may have been due to the fact that the ES was designed to differentiate individuals with “true amnesia” from those with suboptimal effort, while individuals with HD tend to demonstrate a more subcortical (i.e., retrieval memory deficit, in addition to problems with attention/speed of mental processing, and executive function) as opposed to cortical pattern of deficits (i.e., encoding/consolidation deficit). Dunham and colleagues (2014) used a hybrid known groups/simulation design, and compared the sensitivity and specificity of the EI and ES at different levels of impairment (mild, moderate, and severe) based on the RBANS Total Score. On the basis of this experiment, they concluded that the EI is more specific when RBANS Total Scores are mildly or moderately impaired, and the ES Total Scores are more specific when the RBANS Total Scores are severely impaired. These data suggest that the failure rate of the EI and the ES may be different not only based on cutoff used but also on the cognitive profile of the patient population.

The purpose of the present study was to compare the EI and the ES in a non-litigious sample of patients diagnosed with dementia due to heterogeneous etiologies. Based on previous research using the EI in a dementia sample (Barker et al., 2010; Duff et al., 2011), it was hypothesized that the EI would have a high failure rate across etiologies of dementia, and the failure rate of the ES would be lower than that of the EI. We planned to compare failure rates on both embedded measures in a dementia due to AD sub-sample and a non-AD dementia sub-sample, but had no a priori hypotheses for these exploratory comparisons.

Methods

Participants and Procedure

Participants were patients seen between 2004 and 2012 at the University of Saskatchewan’s Rural and Remote Memory (RRMC) clinic for an initial dementia assessment by an inter-professional team comprised of a neuropsychologist, neurologist,
physiotherapist, and registered nurse (Morgan et al., 2009; Data Release 6). Individuals who attended the RRMC completed a clinical interview with the neuropsychologist and neurologist, attended by at least one (typically family) caregiver. Following the clinical interview, the patient completed a brief neuropsychological assessment and a CT head scan, and both patient and caregiver completed self-report measures of quality of life, mood, and activities of daily living. As part of the pre-clinic assessment, up-to-date comprehensive blood work was completed (CBC, electrolytes, creatinine, urea, random glucose, Ca, TSH, and B12) and was interpreted by the neurologist to rule out medical causes for cognitive impairment (see Morgan et al., 2009).

Diagnostic decisions were made based on the clinical judgment of the neuropsychologist and neurologist, who used all interprofessional information from the full day assessment. Diagnoses of dementia were based on neuropsychological impairment in at least two domains and functional impairment (based on structured functional assessment questionnaires for the patient and caregiver, but also from clinical interview). Dementia etiology was based on clinical history and followed the guidelines of the Third Canadian Consensus Conference on Diagnosis and Treatment of Dementia (CCCTD3; Robillard, 2007). Briefly, the CCCTD3 includes recommendations from a panel of Canadian dementia experts who reviewed the quality of evidence for the main published diagnostic criteria for dementia (published pre-2007), such as dementia due to AD, DLBD, FTD, and VaD. The CCCTD3 recommended clinical use of dementia criteria based on the accumulating empirical literature since their initial publication. The consideration of all criteria for all etiologies of dementia presented in the CCCTD3 was routinely performed in the team’s diagnostic consensus.

All individuals who were assessed and diagnosed with either dementia due to AD or a non-AD dementia were included in this study (Table 1). Six individuals who reported that their primary income was due to either disability or worker’s compensation were excluded due to the potential for secondary gain. An additional 52 participants were excluded due to missing data from the RBANS subscales. In total, data from 145 participants with either AD (n = 90) or non-AD dementia (n = 55), and without clear potential for monetary secondary gain, were included.

Measures

The RBANS provides a brief evaluation of adult cognitive functioning and measures the domains of immediate memory, visuospatial/construction, language, attention, and delayed memory (Randolph et al., 1998; Strauss et al., 2006). The RBANS is composed of 12 subtests and provides an index score for each of the domains listed above as well as a total score (Strauss et al., 2006). The EI is calculated by converting raw scores from the Digit Span and List Recognition subtests into weighted scores (from tables provided in the source), and summing these weighted scores (Silverberg et al., 2007). EI scores range from 0 to 12, and higher scores suggest poorer effort (Silverberg et al., 2007). Silverberg and colleagues (2007) suggested a cutoff score of 3, where scores >3 suggest suboptimal effort. This cutoff score has been generally accepted (e.g., Duff et al., 2011; Hook et al., 2009; Sieck et al., 2013) and was used in this study. In contrast, the ES is calculated using the raw scores from the List Recognition, List Recall, Story Recall, Figure Recall, and Digit Span subtests of the RBANS using the formula provided above (Novitski et al., 2012). ES scores range from −8 to 28 and lower scores suggest poorer effort. As suggested by Novitski and colleagues (2012) and adopted by Sieck and colleagues (2013), a cutoff score of 12 was used, where scores <12 indicate suboptimal effort. As an additional comparison, we also analyzed the data using a cutoff score of 7, used by Schroeder and colleagues (2012) to explore whether this limited false positives while maintaining adequate sensitivity.

The clinical dementia rating (CDR) is a clinician’s global rating of dementia that takes into account the results of cognitive performance and ratings of cognitive behavior in everyday activities. The CDR rating covers the domains of memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care (Berg et al., 1988). The rating in the memory domain, however, is weighted more heavily in the overall CDR score which ranges from 0 (no dementia) to 3 (severe

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Dementia due to AD</th>
<th>Non-AD dementia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s disease</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Vascular dementia</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Dementia with Lewy bodies</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Huntington’s disease</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>FTD: frontal variant</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>FTD: semantic dementia</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>FTD: progressive non-fluent aphasia</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Dementia due to a medical condition</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dementia due to multiple etiologies</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Dementia not otherwise specified</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
dementia. Summing the scores in each of the specific domains, or “boxes,” provides the CDR sum of boxes (CDR-SOB) score, which is a more detailed and sensitive measure of global dementia severity than the overall CDR (O’Bryant et al., 2008), likely because it is less focused on memory (and therefore dementia due to AD). The CDR-SOB ranges from 0 to 18 with higher scores suggesting greater dementia severity.

**Results**

**Effort Index**

An EI score was calculated for all participants (N = 145) and nearly half of the sample, 48% (70 participants), had a score greater than the cutoff for suboptimal effort (>3). There was a significant correlation between EI scores and the CDR-SOB, r = .26, p = .002, where higher EI scores (those suggesting suboptimal effort) were associated with increased dementia severity. When the sample was divided by diagnosis, 49% (44 participants) of the dementia due to AD sub-sample (n = 90) had an EI score above the cutoff for suboptimal effort. In the non-AD dementia sub-sample (n = 55), 46% (25 participants) also scored above the cutoff.

**Effort Scale**

Ninety participants (62%) met the criteria described by Novitski and colleagues (2012) to calculate an ES score (i.e., a List Recognition raw score <19 and a Digit Span raw score <9). Of these, 14% (13 participants) had an ES score indicating suboptimal effort (<12). When a more stringent criterion of <7 suggested by Schoeder and colleagues (2013) was used, 7% (6 participants) had scores indicating suboptimal effort. The correlation between dementia severity, as measured by the CDR-SOB, and ES scores was small in magnitude and not statistically significant, r = .12, p = .275.

Next, the sample was divided by dementia diagnosis. For the dementia due to AD sub-sample (n = 94), an ES score could be calculated for 53 participants. Using the Novitski and colleagues’ (2012) cutoff, 4% (2 participants) of the AD sub-sample had an ES score suggesting suboptimal effort. No AD participants had scores less than the Schroeder and colleagues’ (2012) cutoff (<7). For the non-AD dementia sub-sample (n = 55), an ES score could be calculated for 36 participants, of which 31% (11 participants) had an ES score suggesting suboptimal effort with the Novitski and colleagues’ (2012) cutoff (<7). With the Schroeder and colleagues’ (2012) cutoff (<7) 17% (6 participants) had scores suggestive of suboptimal effort.

**Discussion**

This study explored two embedded measures of suboptimal effort developed for the RBANS in a non-litigating sample of individuals diagnosed with dementia. Consistent with previous research (Duff et al., 2011; Barker et al., 2010), when scores on the EI were examined, a high proportion of individuals diagnosed with dementia (48%) had scores suggesting suboptimal effort. Rates suggesting suboptimal effort on the EI were irrespective of diagnoses of AD versus non-AD dementias. Further, participants with more severe dementia were more likely to have scores suggesting suboptimal effort, supporting previous findings relating performance on the EI to disease severity and overall cognitive impairment (Barker et al., 2010; Duff et al., 2011; Hook et al., 2009). These data suggest that in a dementia sample the EI may simply reflect genuine impairment rather than provide additional information about the validity of an individual’s test performance.

In contrast, when the ES measure was used with its original cutoff, only 14% of the sample had a score suggesting suboptimal effort. Further, the ES was not highly associated with dementia severity (the small correlation was near trivial in magnitude and was statistically non-significant). When the sample was divided by dementia diagnosis, the proportion of scores suggesting suboptimal was only 4% for the dementia due to AD sub-sample. Given that Novitski and colleagues (2012) developed the ES specifically for use with individuals with an amnestic disorder such as AD, this result supports the rationale for their algorithm. Moreover, the ES cutoff modification proposed by Schroeder and colleagues (2012) resulted in zero false positives in the AD sub-sample.

In contrast, in the non-AD dementia sub-sample the proportion of ES scores suggesting suboptimal effort rose to nearly one-third, 31%, with the Novitski and colleagues (2012) cutoff, and remained high 17% with the Schroeder and colleagues (2012) cutoff. These data are consistent with Sieck and colleagues’ (2013) findings of a high false-positive ES scores in individuals diagnosed with HD, a non-amnestic neurodegenerative disease.

These results are best understood by considering the nature of the cognitive impairment of the individuals in the AD and non-AD sub-samples. AD is a cortical dementia and is characterized by a deficit in episodic memory (Morris, 2008). The ES was developed based on the premise that in AD, and other true amnestic disorders, an individual’s performance on tests of free recall will decline before performance on tests of recognition (Novitski et al., 2012). Here, the non-AD sub-sample is
heterogeneous (Table 1), and the etiologies of dementia in this sub-sample are characterized by primary deficits in cognitive domains other than episodic memory. For example, in DLB cognitive impairments are typically in the domains of visual perception, attention, and executive function (Morris, 2008). As episodic memory is less likely to be the primary deficit for individuals in the non-AD sub-sample their scores on measures recognition and recall are more likely to be similar and, recalling that RBANS $ES = (\text{List Recognition} – [\text{List Recall} + \text{Story Recall} + \text{Figure Recall}])$, this would lead to a low $ES$ score that is due to the genuine nature of the cognitive impairment, not suboptimal effort. This conclusion is limited by the heterogeneity of the non-AD sub-sample, and small number of individuals with each diagnosis within this category.

Although these data are limited by the lack of stand-alone tests for suboptimal effort, the lack of any possible financial secondary gain mitigates some concern regarding a high proportion of the sample demonstrating true suboptimal effort (due to various factors that could impact engagement in the testing process or a factitious disorder). Consistent with previous studies using the RBANS $EI$ with individuals with cognitive impairments (Barker et al., 2010; Duff et al., 2011; Hook et al., 2009), the failure rate was unacceptably high in this study, and we support previous warnings against the use of the $EI$ in dementia evaluations. In contrast, the RBANS $ES$ appears well suited for use with individuals who have an amnestic disorder such as dementia due to AD. Nevertheless, due to the unacceptably high failure rate on the $ES$ for participants with non-AD dementias, use of either embedded RBANS measure of suboptimal effort is cautioned in a memory clinic setting.

**Funding**

RB’s work was supported by a doctoral award from the Alzheimer Society of Canada. JE’s work was supported by a doctoral award from Public Health and the Agricultural Rural Ecosystem through the Canadian Institutes of Health Research (CIHR). DM’s work was supported by an Applied Chair in Health Services and Policy Research from the Canadian Institutes of Health Research (CIHR) and the Saskatchewan Health Research Foundation (SHRF).

**References**


