OXFORD
UNIVERSITY PRESS

Archives of Clinical Neuropsychology 27 (2012) 114–118

Archives of CLINICAL NEUROPSYCHOLOGY

Brief Empirical Report

# Brief Report: The Temporal Stability of the Repeatable Battery for the Assessment of Neuropsychological Status Effort Index in Geriatric Samples

Kerry M. O'Mahar[1], Kevin Duff[2], James G. Scott[1,*], John F. Linck[3], Russell L. Adams[1], James W. Mold[4]

[1]*Department of Psychiatry and Behavioral Sciences, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA*
[2]*Department of Neurology, University of Utah, Salt Lake City, UT, USA*
[3]*VA Medical Center, Oklahoma City, OK, USA*
[4]*Department of Family and Preventative Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA*

*Corresponding author at: Department of Psychiatry and Behavioral Sciences, University of Oklahoma Health Sciences Center, 920 Stanton L Young Blvd, Oklahoma City, OK 73104, USA. Tel.: +1-405-271-8001; fax: +1-405-271-8802.
*E-mail address:* jim-scott@ouhsc.edu (J.G. Scott).

Accepted 29 July 2011

**Abstract**

The Effort Index (EI) of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) was developed to identify inadequate effort. Although researchers have examined its validity, the reliability of the EI has not been evaluated. The current study examined the temporal stability of the EI across 1 year in two independent samples of older adults. One sample consisted of 445 cognitively intact older adults (mean age = 72.89; 59% having 12–15 years of education) and the second sample consisted of 51 individuals diagnosed with amnestic Mild Cognitive Impairment (mean age = 82.41; 41% having 12–15 years of education). For both samples, the EI was found to have low stability (Spearman's $\rho = .32–.36$). When participants were divided into those whose EI stayed stable or improved versus those whose EI worsened (i.e., declining effort) on retesting, it was observed that individuals with lower baseline RBANS Total scores tended to worsen on the EI across time. Overall, the findings suggest low temporal stability of the EI in two geriatric samples. In particular, individuals with poorer cognition at baseline could present with poorer effort across time. These findings also suggest the need to further examine the temporal stability of other effort measures.

*Keywords:* Malingering/symptom validity testing; Elderly/geriatrics/aging; Mild cognitive impairment

## Introduction

The Effort Index (EI) of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998) was developed to identify inadequate effort (Silverberg, Wertheimer, & Fichtenberg, 2007). In the EI development study, extremely poor performance on two subtests of the RBANS, Digit Span and List Recognition, was very rare in a mixed clinical sample of patients with neurological disorders. Higher weighted values were assigned to less frequently occurring raw scores on these two subtests, and the weighted scores were summed to calculate the EI. The EI was validated in separate samples of patients with traumatic brain injuries, a clinical malingering sample, and three simulation groups. Classification accuracy rates indicated that the EI effectively identified inadequate effort in these groups (Silverberg et al., 2007).

Despite this initial development and validation study, researchers have recently questioned the validity of the EI, especially in geriatric samples. Hook, Marquine, and Hoelzle (2009) found that 31% of their clinically referred, non-litigating older patients were identified with suspect effort on the EI. Barker, Horner, and Bachman (2010) found only modest sensitivity for the EI in their large sample of clinically referred geriatric cases. Lastly, Duff and colleagues (2011) concluded that the EI is valid for use among cognitively intact and mildly cognitively impaired older adults (i.e., low base rates of elevated EI scores), but that among more cognitively impaired older adults, the EI found inflated rates of poor effort.


doi:10.1093/arclin/acr072 Advance Access publication on 9 November 2011

Although reliability data for the scales that make up the EI have been published, to date there are no reliability studies for the EI itself. Temporal stability of the EI is an important psychometric property to evaluate. First, to date, there has been limited empirical attention on the reliability of effort measures. Second, it has been proposed that effort level is not necessarily stable, even within a single battery or test (Heilbronner, Sweet, Morgan, Larrabee, & Millis, 2009). Third, the scales that make up the EI (i.e., Digit Span and List Recognition) have been demonstrated to have age-related change over time (see Duff et al., 2005). Therefore, some degree of test–retest variability for the EI seems likely and clinicians will be expected to make conclusions about the amount and significance of change in performance (Lezak, Howieson, Loring, & Fischer, 2004).

Interpretation of variability in the EI would be enhanced if clinicians were provided data about temporal stability within older populations for whom adequate and consistent effort over time was presumed. Toward that end, the present study involved the examination of the temporal stability of the EI in two samples of older adults for whom adequate and stable effort was assumed. The available data on the reliability of the subtests that make up the EI (i.e., Digit Span, $r = 0.59$; List Recognition, $r = 0.53$; Duff et al., 2005) indicate that the temporal stability of this index would be low. Next, using a change score in EI performance, participants were identified as having either stable/improving effort or worsening effort on the EI. Older age was expected to be associated with worsening EI. Similarly, lower education was expected to be associated with worsening EI. As men tend to do better on Digit Span and women tend to do better on List Recognition (Duff, Schoenberg, Mold, Scott, & Adams, 2011), gender was not expected to affect EI. Finally, lower RBANS Total score was expected to be associated with worsening EI.

## Methods

### *Participants*

Data from two separate and independent samples were used in the analyses. Analyses were conducted separately for the two samples. One sample consisted of 445 cognitively intact, healthy older adults who participated in a longitudinal study of health outcomes. Additional details about this sample are reported in Duff and colleagues (2005). The second sample consisted of 51 older adults diagnosed with amnestic Mild Cognitive Impairment (MCI) participating in a longitudinal study of cognitive progression. Additional details about this sample are reported in Duff and colleagues (2010). Demographic data for each sample are presented in Table 1. In both samples, RBANS data from baseline (i.e., EI1) and 1-year follow-up (i.e., EI2) visits were used. Since both samples were collected as part of research projects and participation in these research projects was voluntary, effort was presumed to be adequate. All participants provided written informed consent, and all procedures were approved by the local Institutional Review Boards.

### *Measures*

The RBANS (Randolph, 1998) is a screening measure for cognitive status, and it is comprised of 12 subtests that yield five Index scores (Immediate Memory, Visuospatial/Constructional, Language, Attention, and Delayed Memory) and a Total Scale

**Table 1.** Demographic information and temporal reliability analyses

| | Cognitively intact | Amnestic MCI |
|---|---|---|
| *N* | 445 | 51 |
| Age (years) | 72.89 (5.52) | 82.41 (6.58) |
| Education (%) | | |
| <9 years | 3 | 0 |
| 9–11 years | 7 | 0 |
| 12 years | 26 | 17 |
| 13–15 years | 33 | 24 |
| >15 years | 31 | 59 |
| Gender (% female) | 57 | 80 |
| RBANS Total Score | 101.42 (14.99) | 96.18 (11.19) |
| EI1 | 0.34 (0.90) (range = 0–6) | 0.43 (0.90) (range = 0–3) |
| EI2 | 0.33 (0.89) (range = 0–6) | 0.65 (1.11) (range = 0–4) |
| Retest interval (days) | 379.83 (49.47) | 371.36 (31.69) |
| Spearman's $\rho$ EI1 and EI2 | .363** | .319* |

*Notes:* MCI = Mild Cognitive Impairment; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; EI = Effort Index;
*$p < .05$.
**$p < .01$.

score. Test–retest reliability data are presented in the manual and elsewhere (Duff et al., 2005). To calculate the EI, the procedures described by Silverberg and colleagues (2007) were used. EI scores ranged from 0 to 12, with higher scores indicating poorer effort. EI scores were calculated for each individual at each time point (i.e., EI1 and EI2).

### *Data Analyses*

The temporal stability of the EI was examined separately for each group by comparing EI1 and EI2 scores using Spearman's $\rho$. Next, within each sample separately, individuals were classified as having either stable or improving EI scores (i.e., same or better score on follow-up) or worsening EI scores (i.e., worse score on follow-up). The participants with stable ($n = 347$) or improving ($n = 52$) scores were grouped together because these scores appeared to be more reflective of adequate effort than worsening EI scores. Non-parametric between-group analyses (i.e., Mann–Whitney $U$-tests, $\chi^2$ test) were used to compare the two groups (stable/improving vs. worsening; independent variable) on age, education, gender, retest interval, and baseline RBANS Total Scale score (dependent variables). Nonparametric statistics were used because the EI was expected to violate assumptions of the central limit theorem due to non-normally distributed scores (i.e., high mean and restricted *SD* on both Digit Span and List Recognition; Kiess, 1996). Additionally, many of the dependent variables were not normally distributed. Alpha was set at 0.05 throughout these analyses.

## Results

When examining the temporal stability of the EI, significant and positive correlations were observed for both samples. For the cognitively intact sample, the Spearman's $\rho$ correlation between EI1 and EI2 was .36, $df = 443$, $p < .001$. In the MCI sample, EI1 and EI2 correlated at .32, $df = 49$, $p = .02$ (Table 1).

In the cognitively intact sample, 399 participants had EI scores that remained stable or improved on follow-up testing and 46 participants had EI scores that worsened. When these two groups were compared, there were no differences ($p > .05$) for age, education, gender, or retest interval between the stable/improving EI and the worsening EI groups. However, there was a statistically significant difference on RBANS Total Scale score. The worsening EI group had significantly lower RBANS Total Score at baseline than the stable/improving group ($U = 6{,}481.5$, $p = .005$).

In the MCI sample, 38 participants had EI scores that remained stable or improved across 1 year. In this sample, 13 participants showed worsening EI scores. For the MCI sample, there were no group differences for age, education, gender, retest interval, or RBANS Total score between the stable/improving EI and the worsening EI groups (Table 2).

## Discussion

This study examined the temporal stability of the EI in two geriatric samples: one that was cognitively intact and the other classified with amnestic MCI. Results indicated that for both samples, the EI had low stability (Spearman's $\rho = .32–.36$) over a 1-year period. To our knowledge, this is first study to examine the temporal stability of the EI. In fact, studies of the stability of

**Table 2.** Descriptive statistics for the stable/improved and worsened EI groups

| | Cognitively Intact | | Amnestic MCI | |
|---|---|---|---|---|
| | Stable/improved EI | Worsened EI | Stable/improved EI | Worsened EI |
| *N* | 399 | 46 | 38 | 13 |
| Age | 72.79 (5.42) | 73.74 (6.29) | 82.45 (6.52) | 82.31 (7.04) |
| Education (%) | | | | |
| <9 years | 2 | 6 | 0 | 0 |
| 9–11 years | 6 | 11 | 0 | 0 |
| 12 years | 28 | 15 | 21 | 8 |
| 13–15 years | 33 | 33 | 24 | 23 |
| >15 years | 31 | 35 | 55 | 69 |
| Gender (% female) | 58 | 46 | 84 | 69 |
| RBANS Total | 96.48 (14.36)* | 91.55 (12.70) | 96.13 (9.90) | 98.77 (12.56) |
| Retest interval (days) | 379.10 (49.85) | 386.18 (46.05) | 371.55 (30.23) | 370.89 (32.17) |
| EI change | −0.22 (0.63) (range = −4 to 0) | 1.74 (1.01) (range = 1–6) | −0.26 (0.72) (range = −3 to 0) | 1.62 (0.87) (range = 1–3) |

*Notes:* MCI = Mild Cognitive Impairment; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; EI = Effort Index
*Note that for the cognitively intact sample, the RBANS Total score is significantly higher for the stable/improved EI group; $p = .005$.

most symptom validity tests appear to be lacking. Hilsabeck and Gouvier (2005) reported test–retest reliabilities of 0.62–0.94 (mean $r = 0.86$) on a range of cognitively based effort measures (e.g., Word Completion Memory Test, Rey Fifteen Item Test, Dot Counting Test) across a 2-week interval for healthy controls, simulators, and memory-impaired patients. Merckelbach and Smith (2003) observed test–retest reliability of 0.72 on a self-report scale of psychiatric and cognitive symptoms (Structure Inventory of Malingered Symptomatology) across 3 weeks for undergraduate students. Although the retest correlations observed in the current study are well below than those reported in the other studies, there are some notable differences. First, our retest interval was considerably longer (i.e., 1 year vs. 2–3 weeks), and retest intervals are known to affect reliability. Second, our study used geriatric subjects, where most participants in the other studies were well under age 65. Lastly, and perhaps most importantly, the other studies examined measures specifically designed to assess effort, whereas our study used an embedded effort measure. Embedded effort measures are often originally and primarily developed to assess cognitive functions, but they are later realized to also assess effort. One challenge of these embedded effort measures is trying to determine if poor performance is due to low effort or impaired cognition. In the case of the EI, the two RBANS subtests that make up this measure (i.e., Digit Span and List Recognition) only have low reliability ($r = 0.59$ and 0.53, respectively) across 1 year in geriatric samples (Duff et al., 2005), so it was expected that the EI's reliability would be limited. Clinicians are encouraged to interpret changes in the EI very cautiously. Low stability is expected.

Secondary analyses were conducted to determine whether there were group differences between participants whose EI performance was stable or improved compared with participants whose pattern of EI declined across time (i.e., showed poorer effort on follow-up). When comparing these two groups, there were no differences on demographic variables (i.e., age, gender, education). However, the role of baseline cognitive status was equivocal. In the large cognitively intact sample, individuals who worsened on the EI across time had significantly lower scores on the RBANS Total Scale at baseline. This suggests that an individual with lower cognitive status at baseline may be more likely to demonstrate poorer effort on the EI over time. This finding extends the work of Duff and colleagues (2011) who found a relationship between cognitive status and the EI in geriatric samples. Their findings, however, observed relationships between baseline cognitive status and baseline effort measures, whereas the current findings apply to changes in EI across 1 year. One clinical implication of these findings is that "true" cognitive decline could be interpreted as worsening effort in some older and cognitively impaired patients. In our second sample of individuals with MCI, baseline cognition was not related to EI changes across time.

Another interpretation of the current findings may come from the recent observation that cognitive variability itself may be a marker of cognitive dysfunction/decline. Several investigators (Cherbuin, Sachdev, & Anstey, 2010; Gorus, De Raedt, Lambert, Lemper, & Mets, 2008; Loewenstein et al., 2006; Tractenberg & Pietrzak, 2011) have reported that poor temporal stability of a variety of cognitive measures was related to the eventual development of MCI and dementia. Those showing decline in the EI in the current study may be on a downward cognitive trajectory. Regardless of the interpretation of the findings, clinicians are encouraged to exercise caution when interpreting worsening effort on the EI among older patients.

The current study has several limitations. First, no additional measures of effort were employed beyond the EI and it is possible that some of the participants did not put forth adequate effort. However, since all individuals were voluntary participants in research studies on cognition and aging, it is unlikely that poor effort was common in these samples. Nonetheless, future investigations should include independent measures of effort. Second, we assumed that effort was consistent over time. Therefore, we concluded that the observed variability on the EI reflected cognitive (as opposed to motivational) factors. It is possible that our results simply reflect the limited stability of effort. Yet, another possibility is that the psychometric properties (e.g., restricted range of scores) of the scales that make up the EI limit the correlation coefficients, and therefore reliability, of this index. Additional examination of the stability of effort, as well as moderating variables, is needed. Third, independent measures of cognitive status should be employed. We used the RBANS Total Scale which is comprised of the subtests that make up the EI; therefore, results may reflect this overlap. However, when we calculated an RBANS Total Scale score that did not include Digit Span and List Recognition, the findings remained the same in our group comparisons. Fourth, our definition of decline in EI across time may have been too liberal. We categorized any decline in the EI (i.e., $\geq$1 point drop) as a decline, and it may be more clinically useful to examine participants who progressed from adequate EI (i.e., EI $\leq 3$) to suspect EI (i.e., EI $> 3$) using the most conservative cut-off outlined by Silverberg and colleagues (2007). Finally, the length of retest interval may be an important factor in the temporal stability of the EI. Our retest interval was restricted to 1 year, and shorter or longer retest intervals may have led to different results. Future studies should include a wider range of retest intervals to assess the role of time.

Embedded effort measures have been identified as potentially useful aides for neuropsychological evaluations (Wolfe et al., 2010). These embedded effort measures sometimes involve novel combinations of existing subtests and the psychometric properties of these new indices should be evaluated. Without such evaluation, interpretation of embedded effort measures is limited. The results of this investigation suggest that variables apart from effort (e.g., cognitive status) may explain change in EI performance. Continued investigation into the reliability and validity of embedded effort measures is encouraged.

## Funding

The project described was supported research grants from the National Institutes on Aging K23 AG028417-01A2 (K.D.) and the Oklahoma Center for the Advancement of Science and Technology (OCAST) HR02-115. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.


## Conflict of interest

None declared.

## References


Barker, M. D., Horner, M. D., & Bachman, D. L. (2010). Embedded indices of effort in the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) in a geriatric sample. *The Clinical Neuropsychologist*, *24*, 1064–1077.

Cherbuin, N., Sachdev, D., & Anstey, K. J. (2010). Neuropsychological predictors of transition from health cognitive aging to mild cognitive impairment: The PATH through life study. *American Journal of Geriatric Psychiatry*, *18*, 723–733.

Duff, K., Beglinger, L., Moser, D. J., Paulsen, J. S., Schultz, S., & Arndt, S. (2010). Predicting cognitive change in older adults: The relative contribution of practice effects. *Archives of Clinical Neuropsychology*, *25*, 81–88.

Duff, K., Schoenberg, M. R., Mold, J. W., Scott, J. G., & Adams, R. L. (2011). Gender differences on the Repeatable Battery for the Assessment of Neuropsychological Status subtests in older adults: Baseline and retest data. *Journal of Clinical and Experimental Neuropsychology*, *33*, 448–455.

Duff, K., Schoenberg, M. R., Patton, D., Paulsen, J. S., Bayless, J. D., Mold, J., et al. (2005). Regression-based formulas for predicting change in RBANS subtests with older adults. *Archives of Clinical Neuropsychology*, *20*, 281–290.

Duff, K., Spering, C. C., O'Bryant, S. E., Beglinger, L. J., Moser, D. J., Bayless, J. D., Culp, R. R., Mold, J. W., Adams, R. L., & Scott, J. G. (2011). The RBANS effort index: Base rates in geriatric samples. *Journal of Applied Neuropsychology*, *18*, 11–17.

Gorus, E., De Raedt, R., Lambert, M., Lemper, J. C., & Mets, T. (2008). Reaction times and performance variability in normal aging, mild cognitive impairment, and Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology*, *21*, 204–218.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., & Millis, S. R. (2009). American Academy of Clinical Neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, *23*, 1093–1129.

Hilsabeck, R. C., & Gouvier, W. D. (2005). Detecting simulated memory impairment: Further validation of the Word Completion Memory Test (WCMT). *Archives of Clinical Neuropsychology*, *20*, 1025–1041.

Hook, J. N., Marquine, M. J., & Hoelzle, J. B. (2009). Repeatable Battery for the Assessment of Neuropsychological Status effort index performance in a medically ill geriatric sample. *Archives of Clinical Neuropsychology*, *24*, 231–235.

Kiess, H. O. (1996). *Statistical concepts for the behavioral sciences* (2nd ed.). Needham Heights, MA: Allyn and Bacon.

Lezak, M. D., Howieson, D. B., Loring, H. J., & Fischer, J. S. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Loewenstein, D. A., Acevedo, A., Agron, J., Issacson, R., Strauman, S., Crocco, E., et al. (2006). Cognitive profiles in Alzheimer's disease and in mild cognitive impairment of different etiologies. *Dementia and Geriatric Cognitive Disorders*, *21*, 309–315.

Merckelbach, H., & Smith, G. P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology*, *18*, 145–152.

Randolph, C. (1998). *Repeatable Battery for the Assessment of Neuropsychological Status*. San Antonio, TX: The Psychological Corporation.

Silverberg, N. D., Wertheimer, J. C., & Fichtenberg, N. L. (2007). An Effort Index for the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). *The Clinical Neuropsychologist*, *21*, 841–854.

Tractenberg, R. E., & Pietrzak, R. H. (2011). Intra-individual variability in Alzheimer's disease and cognitive aging: Definitions, context, and effect sizes. *PLoS One*, *6(4)*: e16973. doi:10.1371/journal.pone.0016973

Wolfe, P. L., Millis, S. R., Hanks, R., Fichtenberg, N., Larrabee, G. J., & Sweet, J. J. (2010). Effort indicators within the California Verbal Learning Test-II. *The Clinical Neuropsychologist*, *24*, 153–168.