HHS Public Access

Author manuscript

J Neurol Sci. Author manuscript; available in PMC 2020 May 15.

Published in final edited form as:

J Neurol Sci. 2019 May 15; 400: 104–109. doi:10.1016/j.jns.2019.03.016.

Brief and Cost-Effective Tool for Assessing Verbal Learning in Multiple Sclerosis: Comparison of the Rey Auditory Verbal Learning Test (RAVLT) to the California Verbal Learning Test – II (CVLT-II)

Meghan Beier, PhD[a], **Abbey J Hughes, PhD**[a], **Michael W. Williams, PhD**[b], and **Elizabeth S. Gromisch, PhD**[c,d]

[a]Johns Hopkins University School of Medicine, Department of Physical Medicine and Rehabilitation, Division of Rehabilitation Psychology and Neuropsychology, Suite 413, Church Professional Building, 98 North Broadway, Baltimore, MD 21231, USA

[b]University of Houston, Department of Psychology, 3695 Cullen Blvd, Heyne Building Rm 126, Houston, TX 77204, USA, MWwilliams2@uh.edu

[c]Mandell Center for Multiple Sclerosis, Mount Sinai Rehabilitation Hospital, Trinity Health Of New England, 490 Blue Hills Avenue, Hartford, CT 06112, USA.

[d]Department of Neurology, University of Connecticut School of Medicine, Farmington, CT, USA, elizabeth.gromisch@stfranciscare.org

Abstract

Background.—The Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) is a common cognitive screening tool. However, administration and scoring can be time-consuming, and its use of proprietary subtests like the California Verbal Learning Test – II (CVLT-II) is financially limiting. Use of the non-proprietary Rey Auditory Verbal Learning Test (RAVLT) may be provide a valid alternative.

Objectives.—To compare the RAVLT and CVLT-II in terms of diagnostic accuracy for detecting cognitive impairment, and to determine optimal cut-scores for the RAVLT.

Methods.—100 participants with MS completed the five learning trials from the RAVLT and CVLT-II. Receiver operating characteristic analyses were used to compare the measures' sensitivities, specificities, positive predictive values (PPV) and negative predictive values (NPV), and to identify optimal cut-scores.

Results.—Using a criterion of 1.5 *SD* below the normative sample mean, the RAVLT showed fair to good (κs= .21-.41) agreement with the CVLT-II. A cut-score of 12 on Trials 1+2 of the

Corresponding Author: Meghan Beier, PhD, Johns Hopkins University School of Medicine, Department of Physical Medicine and Rehabilitation, Division of Rehabilitation Psychology and Neuropsychology, Suite 413, Church Professional Building, 98 North Broadway, Baltimore, MD 21231; Phone: 410-502-2441; mbeier1@jhu.edu.

RAVLT showed fair sensitivity (75%) and specificity (76%) and did not differ significantly from the CVLT-II (p>.05).

Conclusions.—Performance on initial learning trials of the RAVLT may provide a brief, valid, and cost-effective alternative to the CVLT-II for screening verbal learning impairments in MS.

Keywords

Multiple sclerosis; neuropsychological tests; cognitive dysfunction; cut scores

1. Introduction

Multiple sclerosis (MS) is a chronic, neurologic condition characterized by inflammatory demyelination and neurodegeneration. Forty to 65% of persons with MS experience cognitive impairment, which limits full participation in daily activities such as employment, management of household finances, and social relationships [1,2]. Although MS can impact performance across a number of cognitive domains, research indicates that deficits are most pronounced in the areas of processing speed and initial learning [1,3,4]. Given these findings, it is imperative that these cognitive domains be accurately and routinely screened as part of standard MS care to: (1) determine the need for comprehensive cognitive assessment; (2) monitor cognitive status over time; and (3) identify patients who would benefit from referrals to cognitive rehabilitation treatment [5].

Historically, neurology and primary care settings have employed brief cognitive screening tools such as the Montreal Cognitive Assessment (MOCA) for a variety of clinical populations [6]. Although one study showed correlations between MOCA scores and performance on objective tests of processing speed and learning in MS, the sensitivity, specificity, and other psychometric properties of the MOCA (e.g., cut-scores) have not been established, rendering the MOCA less optimal for use in this population [7]. To address these limitations, a growing body of literature has emerged in support of the Brief International Cognitive Assessment for MS (BICAMS), a brief, sensitive, and reliable screening battery that specifically assesses processing speed and initial learning deficits in individuals with MS [8,9]. The BICAMS includes the Symbol Digit Modalities Test (SDMT) [10], the five learning trials of the California Verbal Learning Test – II (CVLT-II) [11], and the three initial learning trials of the Brief Visuospatial Memory Test – Revised (BVMT-R) [12].

Like the MOCA and other brief cognitive screening tools, the BICAMS takes less than 30 minutes to administer and score, and results can be used to determine whether a full neuropsychological evaluation may be warranted. However, the MOCA and similar counterparts (e.g., Mini Mental Status Exam, MMSE) have a few advantages over the BICAMS. Specifically, they are non-proprietary and can be scored and interpreted within a few minutes by a non-psychologist or non-physician. Conversely, each of the tests within the BICAMS are owned and sold by different companies, which can create administrative and financial burdens on busy clinics that may have limited resources. For example, despite the excellent psychometric properties and significant clinical utility of the CVLT-II for conducting comprehensive neuropsychological evaluations, the costs of scoring software and

record forms may not be justified for clinics using the measure for brief screening. Second, administration, scoring, and interpretation of the CVLT-II can be time-intensive, especially if the administrator is using hand-scoring to determine standardized scores (z or T) for each trial.

Recent studies have proposed strategies for making administration and scoring of the BICAMS more efficient. One study examined an abbreviated version of the CVLT-II and found that scores generated from the first two learning trials (Trial 1 and Trials 1+2) provided 97.5% accuracy when compared with all five learning trials [13]. Another study proposed raw cut-scores for the BICAMS so that hand-scoring time or use of expensive scoring software could be minimized [14]. Despite availability of the abbreviated versions, however, the CVLT-II remains a costly test given the need for purchasing proprietary record forms.

A potential alternative to the CVLT-II is the Rey Auditory Verbal Learning Test (RAVLT), a widely-used, reliable, and valid assessment of auditory verbal learning and memory [15]. The RAVLT is available in the public domain and has been used in previous MS research, including a recent study that used the RAVLT in lieu of the CVLT-II for a German version of the BICAMS [16]. Although prior research in a healthy sample demonstrated equivalence of the two measures [17], a similar study of patients with head injury showed lower normative scores for the CVLT-II relative to the RAVLT. To our knowledge, the RAVLT has not been compared to the CVLT-II in an MS sample. Additionally, raw cut-scores have not yet been identified for the RAVLT. Therefore, the aims of the present study were to: (1) compare the RAVLT to the CVLT-II in terms of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for detecting verbal learning in an MS sample; and (2) determine optimal raw cut-scores for the RAVLT to increase clinical utility. We hypothesized that performance on learning trials of the RAVLT would show acceptable and comparable diagnostic accuracy to performance on the CVLT-II.

2. Methods

2.1 Participants

One hundred participants (see Table 1) were recruited from a MS clinic in an academic medical center located in the Pacific Northwest. Patients in the center were approached during the course of their regularly scheduled appointment with an IRB-approved questionnaire assessing interest in research participation and basic demographic information. Interested patients were then screened by research staff to determine eligibility. If eligible, the participant completed a onetime brief cognitive evaluation. For inclusion, participants were required to be between the ages of 18 and 79 years old and able to read, speak, and write in English. All procedures were approved by the university's institutional review board, and participants provided written informed consent prior to enrolling in the study.

2.2 Procedures

The present study featured a secondary analysis of data collected as part of a cognition measure reliability study [18]. Demographic (e.g., age, gender, race, and ethnicity) and

disease-related variables (e.g., MS type, date of diagnosis, co-morbid medical conditions, and medications) were collected by research staff using chart review or a short demographic questionnaire. Following consent and enrollment, participants completed a single 30- to 45-minute study appointment. Each participant completed four cognitive tests in the following order: SDMT, learning trials of the RAVLT, learning trials of the BVMT-R, and learning trials of the CVLT-II. The measures were administered orally by trained research assistants using standardized administration instructions.

2.3 Measures

RAVLT.—The standard RAVLT is a verbal list-learning and memory test that has been validated in several neurologic populations [19]. After hearing a verbally presented list of 15 unrelated words, examinees are asked to recite as many words as they can recall. This learning procedure is repeated for a total of five trials. Scores include the sum of words recalled within each trial, as well as the sum of words recalled across the five trials. Consistent with previous studies of the abbreviated CVLT-II, raw scores were generated for Trial 1, Trials 1+2, and the sum of Trials 1 through 5 referred to as Total Learning [13,14,20]. A standardized (z) score was also calculated based on manual norms [15] for the Total Learning score.

CVLT-II.—The CVLT-II is a verbal list-learning and memory test that is used in the BICAMS. The CVLT-II uses a list of 16 words that are read aloud to examinees. The 16 words can be grouped into four semantic categories although not presented in grouped format. In each trial, the participant is asked to recall as many words as they can, and the procedure is repeated for a total of five trials. Scores include the sum of words recalled within each trial, as well as the sum of words recalled across the five trials. Raw scores were generated for Trial 1, Trials 1+2, and Total Learning. A standardized (T) score was also calculated based on manual norms [11] for the Total Learning score.

2.4 Statistical Analyses

Preliminary descriptive statistics were calculated with regard to demographic characteristics, as well as the percentage of participants exhibiting cognitive impairment on the RAVLT and CVLT-II. A missing values analysis was run ($\chi^2(111) = 81.57$, $p = .98$), which indicated that the data were missing completely at random and thus justified the use of the expectation maximization (EM) method to impute values for missing data points [21]. See Table 2 for detailed information about missing data from the individual trials of both measures.

Two sets of receiver operative characteristic (ROC) analyses were run to calculate diagnostic classification accuracy (i.e., the area under the curve; AUC). For the first set of analyses, raw RAVLT and CVLT-II scores (Trial 1, Trials 1+2, and Total Learning) were plotted against the criterion variable – impairment on the standardized Total Learning score of their respective measure. Impairment was defined as either 1.5 or 2.0 standard deviations below the normative sample mean (i.e., $z \leq -1.5$ or 2.0 for the RAVLT; $T \leq 35$ or 30 for the CVLT-II). Sensitivity, specificity, PPV, and NPV values were then calculated based on the optimal cut-off score, per the Youden index, for each test [22]. In addition to comparing the AUCs using the R package pROC [23], the level of agreement between the measures was assessed

using Cohen's kappa (κ), which has been previously used in other validation studies of cognitive assessment in MS [20,24]. Interpretive labels (and ranges) for kappa values are: less than chance (<.01); slight (.01-.20); fair (.21-.40); moderate (.41-.60); substantial (.61-.80); and almost perfect (.81-.99).

Given that the first set of analyses identified cognitive impairment based on different normative samples from each test's manual, the proposed cut-off scores and resulting sensitivities, specificities, and predictive values could not be directly compared between the RAVLT and CVLT-II. Thus, for the second set of analyses, the criterion variable – cognitive impairment – was defined as 1.5 or 2.0 *SD* below the normative sample mean on both measures. Cut-off scores for each measure were then plotted against the criterion variable and compared with regard to AUCs, sensitivities, specificities, PPVs, and NPVs using the R package DTComPair [25].

3. Results

Demographic and clinical characteristics of the study sample are presented in Table 1. The sample was predominately female (74%) with relapsing remitting MS (77%). On average, they were 46.24 years old (*SD* = 12.91), had 15.46 years (*SD* = 2.46) of education, and had MS for 10.69 years (*SD* = 8.37). The percentages of impaired participants were numerically similar for the RAVLT and CVLT-II (see Table 3). Individually, 8 participants were impaired on both measures, 8 on just the RAVLT, 7 on only the CVLT-II, and 77 evidenced no impairment for - 1.5 SD below the normative mean. At 2 SD below the normative mean 5 individuals were impaired on both measures, 4 on just the RAVLT, 4 on only the CVLT-II, and 87 evidenced no impairment.

Results from the first set of ROC analyses are presented in Table 4. For participants who were at least 1.5 *SD* below the normative sample mean, raw RAVLT scores of 5, 12, and 42 were the optimal cut-offs for Trial 1, Trials 1+2, and Total Learning scores, respectively. Cut-offs of 4, 12, and 39 were optimal for the CVLT-II. Classification accuracies (AUCs) were not significantly different between the two measures (Trial 1: $p = .85$; Trial 1+2: $p = .16$; Total Learning: $p = .15$), and agreement levels were respectively fair, moderate, and fair for Trial 1, Trials 1+2, and Total Learning. For participants who were at least 2.0 *SD* below the normative sample mean, raw RAVLT scores of 5, 13, and 42, and raw CVLT-II scores of 4, 11, and 35 were optimal cut-offs for Trial 1, Trials 1+2, and Total Learning, respectively. The AUCs did not differ significantly between the two measures (Trial 1: $p = .68$; Trial 1+2: $p = .31$; Total Learning: $p = .09$), and agreement levels were respectively fair, slight, and fair for Trial 1, Trials 1+2, and Total Learning.

Results from the second set of ROC analyses are presented in Table 5. For participants who were at least 1.5 *SD* below the normative sample mean on both tests, AUCs were not significantly different between the RAVLT and CVLT-II (all ps > .05). Additionally, sensitivity, specificity, and predictive values did not significantly differ between the measures (all ps > .05); however, the RAVLT exhibited lower specificity for Trial 1 that approached statistical significance (64% vs 76%; $p = .06$). Consistent with results from the 1.5 *SD* criterion, AUCs were not significantly different between the RAVLT and CVLT-II

(all *p*s > .05) for participants who were at least 2.0 *SD* below the normative sample mean on both tests. However, several differences emerged for other ROC metrics at the 2.0 *SD* criterion. For Trial 1, Trials 1+2, and Total Learning, the RAVLT exhibited significantly lower specificity than the CVLT-II (Trial 1: 64% vs 77%, $p = .04$; Trials 1+2: 65% vs 81%, $p < .01$; Total Learning: 76% vs 93%, $p < .01$). The RAVLT also exhibited lower PPV on Total Learning (18% vs 42%, $p = .02$). The RAVLT and CVLT-II exhibited no significant differences with regard to sensitivity or NPV.

4. Discussion

Overall, as a non-proprietary screening measure of verbal learning, results from the present study support the use of the RAVLT as a comparable measure to the CVLT-II for detecting cognitive impairment in individuals with MS. Given the financial barriers often inherent to conducting evidence-based neuropsychological assessment, the present findings offer the potential for using the RAVLT as a cost-effective alternative to the CVLT-II, without compromising clinical care. Regular screening for cognitive impairment is imperative for comprehensive MS care [26] and has clinical implications for monitoring disease progression and cognitive intervention outcomes. These findings are also consistent with prior literature showing equivalence between RAVLT and CVLT raw scores in a healthy sample [17]. Additionally, this study proposed useful raw cut-scores for the RAVLT, which further support ease and efficiency of administration.

Although overall study findings supported the use of the RAVLT, a few notable differences emerged between the measures. When using a criterion of 2.0 *SD* below the normative sample mean to define cognitive impairment, the RAVLT exhibited poorer specificity than the CVLT-II on all initial learning trial scores, and poorer PPV on the Total Learning score. In contrast, these differences were not observed when using a criterion of 1.5 *SD* below the normative sample mean. Thus, clinicians who elect to use the RAVLT should be mindful to use the recommended cut-off scores of 5 for Trial 1 and 12 for Trials 1+2. Use of a more stringent cut-off of 13 for Trials 1+2 will not improve specificity for identifying more impaired patients. Overall, a cut-off score of 12 for Trials 1+2 on the RAVLT provided the highest level of agreement with the CVLT-II, suggesting that the combined score from the first two trials of the RAVLT may provide the best alternative to the abbreviated CVLT-II when administering the BICAMS. Of note, the optimal cut-offs for the CVLT-II in the present study were the same as those identified in a prior study [13].

The significant findings of this study must be contextualized with the limitations of this study. As a secondary analysis, the original study design did not counterbalance the RAVLT and CVLT-II order of administration; however order of administration was not found to be a significant factor in an earlier study comparing the RAVLT and CVLT [17]. Additionally, the RAVLT and CVLT-II were not co-normed using the same sample, which could account at least in part for the specificity differences observed between the measures at the 2.0 *SD* criterion. It is also important to note that the use of raw scores on the abbreviated RAVLT or CVLT are not intended to replace full neuropsychological evaluations. Based on the available data only initial learning was evaluated. Future studies should examine delayed recall. Of note, however, many studies in MS have found learning as the primary deficit

[1,3,4]. Using these measures as screeners does not identify impairment in other cognitive domains, or factors that may impact cognition such as effort or emotional state. However, using either of these as screeners of verbal learning, potentially along with the SDMT or alternate measure of processing speed, may help identify which patients would benefit from a more comprehensive evaluation. Sample homogeneity should also be acknowledged. The sample from this study was largely female, diagnosed with relapsing-remitting MS, and of high education. Therefore, additional research is needed to determine if similar results are found in males, individuals with low education, and those with progressive forms of the disease. Finally, additional data would be needed to determine if the proposed cut-off scores for the RAVLT predict overall performance on a full neuropsychological battery.

Although there are limitations to this study, the present findings may benefit and encourage screening of verbal learning abilities in clinics with limited resources. Impairments in initial learning are common among those with MS, and can have deleterious effects on quality of life and daily activities. Given the emerging research in support of cognitive rehabilitation for learning and memory impairments in MS [27,28], routine screening may help identify those most in need of intervention. Despite neuropsychological assessment, including use of the BICAMS, having a number of financial and administrative barriers, the present study provides evidence for inclusion of a free learning measure that, when using specified cut-offs, is equivalent to its proprietary counterpart. Findings support consideration for additional prospective research utilizing the RAVLT in persons with MS, and exploration for other accessible, cost-effective measures for screening cognitive impairment in MS.

Acknowledgements:

The content and data used for this manuscript were collected and developed under a pilot grant from the Consortium of Multiple Sclerosis Centers, Beier (PI). Data collection was also supported in part by a grant from the National Multiple Sclerosis Society, grant number MB 0008, Ehde (PI). A corresponding author is supported by a grant from the National Institutes of Health (grant no. K23HD086154), Hughes (PI).

References

[1]. Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. Lancet Neurol 2008;7:1139–51. doi:10.1016/S1474-4422(08)70259-X. [PubMed: 19007738]

[2]. Kordovski VM, Frndak SE, Fisher CS, Rodgers J, Weinstock-Guttman B, Benedict RHB. Identifying employed multiple sclerosis patients at-risk for job loss: When do negative work events pose a threat? Mult Scler Relat Disord 2015;4:409–13. doi:10.1016/j.msard.2015.07.005. [PubMed: 26346789]

[3]. Deluca J, Leavitt VM, Chiaravalloti N, Wylie G. Memory impairment in multiple sclerosis is due to a core deficit in initial learning. J Neurol 2013;260:2491–6. doi:10.1007/s00415-013-6990-3. [PubMed: 23832311]

[4]. Rao SM, Grafman J, DiGiulio D, Mittenberg W, Bernardin L, Leo GJ, et al. Memory dysfunction in multiple sclerosis: Its relation to working memory, semantic encoding, and implicit learning. Neuropsychology 1993;7:364–74. doi:10.1037/0894-4105.7.3.364.

[5]. Dardiotis E, Nousia A, Siokas V, Tsouris Z, Andravizou A, Mentis A-FA, et al. Efficacy of computer-based cognitive training in neuropsychological performance of patients with multiple sclerosis: A systematic review and meta-analysis. Mult Scler Relat Disord 2018;20:58–66. doi: 10.1016/j.msard.2017.12.017. [PubMed: 29306740]

[6]. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005;53:695–9. doi:10.1111/j.1532-5415.2005.53221.x. [PubMed: 15817019]

[7]. Dagenais E, Rouleau I, Demers M, Jobin C, Roger E, Chamelian L, et al. Value of the MoCA test as a screening instrument in multiple sclerosis. Can J Neurol Sci J Can Sci Neurol 2013;40:410–5.

[8]. Benedict RHB, Amato MP, Boringa J, Brochet B, Foley F, Fredrikson S, et al. Brief International Cognitive Assessment for MS (BICAMS): international standards for validation. BMC Neurol 2012;12:55. doi:10.1186/1471-2377-12-55. [PubMed: 22799620]

[9]. Langdon DW, Amato MP, Boringa J, Brochet B, Foley F, Fredrikson S, et al. Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). Mult Scler Houndmills Basingstoke Engl 2012;18:891–8. doi:10.1177/1352458511431076.

[10]. Bever CT, Grattan L, Panitch HS, Johnson KP. The Brief Repeatable Battery of Neuropsychological Tests for Multiple Sclerosis: a preliminary serial study. Mult Scler Houndmills Basingstoke Engl 1995;1:165–9. doi:10.1177/135245859500100306.

[11]. Delis Dean C., Joel H Kramer Edith Kaplan, Ober Beth A.. California Verbal Learning Test, Second Ediction (CVLT-II) 2000.

[12]. Benedict Ralph H. B.. Brief Visuospatial Memory Test–Revised. Lutz, FL: Psychological Assessment Resources, Inc; n.d.

[13]. Gromisch ES, Zemon V, Benedict RHB, Chiaravalloti ND, DeLuca J, Picone MA, et al. Using a highly abbreviated California Verbal Learning Test-II to detect verbal memory deficits. Mult Scler Houndmills Basingstoke Engl 2013;19:498–501. doi:10.1177/1352458512454347.

[14]. Beier M, Gromisch ES, Hughes AJ, Alschuler KN, Madathil R, Chiaravalloti N, et al. Proposed cut scores for tests of the Brief International Cognitive Assessment of Multiple Sclerosis (BICAMS). J Neurol Sci 2017;381:110–6. doi:10.1016/j.jns.2017.08.019. [PubMed: 28991659]

[15]. Schmidt M Rey auditory verbal learning test: A handbook. Los Angeles, CA: Western Psychological Services; 1996.

[16]. Filser M, Schreiber H, Pöttgen J, Ullrich S, Lang M, Penner IK. The Brief International Cognitive Assessment in Multiple Sclerosis (BICAMS): results from the German validation study. J Neurol 2018;265:2587–93. doi:10.1007/s00415-018-9034-1. [PubMed: 30171410]

[17]. Crossen JR, Wiens AN. Comparison of the Auditory-Verbal Learning Test (AVLT) and California Verbal Learning Test (CVLT) in a sample of normal subjects. J Clin Exp Neuropsychol 1994;16:190–4. doi:10.1080/01688639408402630. [PubMed: 8021306]

[18]. Beier M, Alschuler K, Amtmann D, Rutter K, Garcia C, Reeves S, et al. BICAMS Tablet Application: A Reliable and Fast Way to Assess Cognitive Functioning, Indianapolis, Indiana: 2015.

[19]. Schoenberg MR, Dawson KA, Duff K, Patton D, Scott JG, Adams RL. Test performance and classification statistics for the Rey Auditory Verbal Learning Test in selected clinical samples. Arch Clin Neuropsychol Off J Natl Acad Neuropsychol 2006;21:693–703. doi:10.1016/j.acn.2006.06.010.

[20]. Gromisch ES, Portnoy JG, Foley FW Comparison of the abbreviated minimal assessment of cognitive function in multiple sclerosis (aMACFIMS) and the brief international cognitive assessment for multiple sclerosis (BICAMS). J Neurol Sci 2018;388:70–5. doi:10.1016/j.jns.2018.03.012. [PubMed: 29627034]

[21]. Do CB, Batzoglou S. What is the expectation maximization algorithm? Nat Biotechnol 2008;26:897–9. doi:10.1038/nbt1406. [PubMed: 18688245]

[22]. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32–5. [PubMed: 15405679]

[23]. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77. doi:10.1186/1471-2105-12-77. [PubMed: 21414208]

[24]. Niccolai C, Portaccio E, Goretti B, Hakiki B, Giannini M, Pastò L, et al. A comparison of the brief international cognitive assessment for multiple sclerosis and the brief repeatable battery in multiple sclerosis patients. BMC Neurol 2015;15:204. doi:10.1186/s12883-015-0460-8. [PubMed: 26472052]

[25]. Stock C, Hielscher T. DTComPair: comparison of binary diagnostic tests in a paired study design 2014.

[26]. Kalb R, Beier M, Benedict RH, Charvet L, Costello K, Feinstein A, et al. Recommendations for cognitive screening and management in multiple sclerosis care. Mult Scler Houndmills Basingstoke Engl 2018;24:1665–80. doi:10.1177/1352458518803785.

[27]. Goverover Y, Chiaravalloti ND, O'Brien AR, DeLuca J. Evidenced-Based Cognitive Rehabilitation for Persons With Multiple Sclerosis: An Updated Review of the Literature From 2007 to 2016. Arch Phys Med Rehabil 2018;99:390–407. doi:10.1016/j.apmr.2017.07.021. [PubMed: 28958607]

[28]. Gromisch ES, Fiszdon JM, Kurtz MM. The effects of cognitive-focused interventions on cognition and psychological well-being in persons with multiple sclerosis: A meta-analysis. Neuropsychol Rehabil 2018:1–20. doi:10.1080/09602011.2018.1491408.

Highlights

- The RAVLT is comparable to the CVLT-II for detecting learning deficits in MS.
- RAVLT and CVLT-II Trial 1, Trials 1+2, and Total Learning cut scores identified.
- Identified optimal CVLT-II cut scores replicate findings from a prior study.
- Recommended RAVLT cut scores: 5 (Trial 1), 12 (Trial 1+2), or 42 (Total Learning).

Table 1

Sample demographics

Age (years)	M (SD): 46.24 (12.91) Range: 19–72
Gender	Female: 74% Male: 24% Not Reported: 2%
Education (years)	M (SD): 15.46 (2.46) Range: 10–22
MS Duration (years)	10.69 ± 8.37 (1–37)
MS Type	Relapsing Remitting: 77% Secondary Progressive: 8% Primary Progressive: 3% Progressive Relapsing: 1% Not Reported: 11%

Abbreviations: CVLT-II = California Verbal Learning Test – II; MS = multiple sclerosis; RAVLT = Rey Auditory Verbal Learning Test.

Table 2

Missing data for all trials of the CVLT-II and RAVLT out of 100 participants

RAVLT	***N***
Trial 1	8
Trial 2	6
Trial 3	7
Trial 4	8
Trial 5	7
Total Raw Score	7
CVLT-II	
Trial 1	0
Trial 2	0
Trial 3	0
Trial 4	1
Trial 5	1
Total Raw Score	1

Table 3

Classification of cognitive impairment on the RAVLT and CVLT-II

	1.5 *SD* below the mean	**2.0 *SD* below the mean**
RAVLT Only	8%	4%
CVLT-II Only	7%	4%
Both Measures	8%	5%

Abbreviations: CVLT-II = California Verbal Learning Test – II; RAVLT = Rey Auditory Verbal Learning Test.

Table 4

Criterion validity and level of agreement between the RAVLT and CVLT-II against their respective standardized scores.

		1.5 *SD* below the mean		2.0 *SD* below the mean	
		RAVLT	**CVLT-II**	**RAVLT**	**CVLT-II**
Trial 1	AUC	0.80 (0.69, 0.90)	0.81 (0.69, 0.93)	0.85 (0.74, 0.96)	0.80 (0.63, 0.98)
	Cut-off	5	4	5	4
	Sensitivity	0.75 (0.47, 0.92)	0.60 (0.33, 0.83)	0.89 (0.51, 0.99)	0.67 (0.31, 0.91)
	Specificity	0.69 (0.58, 0.78)	0.81 (0.71, 0.89)	0.67 (0.56, 0.76)	0.79 (0.69, 0.87)
	PPV	0.32 (0.18, 0.49)	0.36 (0.19, 0.57)	0.21 (0.10, 0.38)	0.24 (0.10, 0.46)
	NPV	0.94 (0.84, 0.98)	0.92 (0.83, 0.97)	0.98 (0.90, 1.00)	0.96 (0.88, 0.99)
	κ	0.21		0.21	
Trials 1+ 2	AUC	0.87 (0.78, 0.95)	0.94 (0.88, 0.99)	0.85 (0.74, 0.96)	0.93 (0.83, 1.00)
	Cut-off	12	12	13	11
	Sensitivity	0.75 (0.47, 0.92)	0.93 (0.66, 0.97)	0.78 (0.40, 0.96)	0.78 (0.40, 0.96)
	Specificity	0.81 (0.71, 0.88)	0.84 (0.74, 0.90)	0.67 (0.56, 0.76)	0.85 (0.75, 0.91)
	PPV	0.43 (0.25, 0.63)	0.50 (0.31, 0.69)	0.19 (0.09, 0.36)	0.33 (0.15, 0.57)
	NPV	0.94 (0.86, 0.98)	0.99 (0.91, 1.00)	0.97 (0.88, 0.99)	0.98 (0.90, 1.00)
	κ	0.41		0.20	
Total Learning	AUC	0.96 (0.92, 0.99)	1.00 (0.97, 1.00)	0.95 (0.90, 1.00)	1.00 (1.00, 1.00)
	Cut-off	42	39	42	35
	Sensitivity	0.94 (0.68, 1.00)	0.93 (0.66, 1.00)	1.00 (0.63, 1.00)	1.00 (0.63, 1.00)
	Specificity	0.85 (0.75, 0.91)	0.92 (0.83, 0.96)	0.79 (0.69, 0.87)	0.97 (0.90, 0.99)
	PPV	0.54 (0.34, 0.72)	0.67 (0.43, 0.85)	0.32 (0.17, 0.52)	0.75 (0.43, 0.93)
	NPV	0.98 (0.91, 1.00)	0.99 (0.92, 1.00)	1.00 (0.94, 1.00)	1.00 (0.95, 1.00)
	κ	0.38		0.22	

Notes. 95% confidence intervals are presented in parentheses.

Abbreviations: AUC= area under the curve; CVLT-II = California Verbal Learning Test - II; NPV = negative predictive value; PPV: = positive predictive value; RAVLT = Rey Auditory Verbal Learning Test.

Table 5

Comparison of raw score cut-offs in predicting impairment on both measures of verbal learning.

		1.5 *SD* below the mean			2.0 *SD* below the mean		
		RAVLT	**CVLT-II**	**p-value**	**RAVLT**	**CVLT-II**	**p-value**
Trial 1	AUC	0.76 (0.57, 0.95)	0.67 (0.48, 0.85)	.229	0.84 (0.66, 1.00)	0.75 (0.49, 1.00)	.751
	Sensitivity	0.63 (0.26, 0.90)	0.38 (0.10, 0.74)	.157	0.80 (0.30, 0.99)	0.60 (0.17, 0.93)	.317
	Specificity	0.64 (0.53, 0.74)	0.76 (0.66, 0.84)	.056	0.64 (0.54, 0.74)	0.77 (0.67, 0.85)	.040
	PPV	0.13 (0.05, 0.29)	0.12 (0.03, 0.32)	.798	0.11 (0.03, 0.26)	0.12 (0.03, 0.32)	.697
	NPV	0.95 (0.86, 0.99)	0.93 (0.84, 0.98)	.336	0.98 (0.90, 1.00)	0.97 (0.89, 1.00)	.434
Trial 1+ 2	AUC	0.84 (0.70, 0.97)	0.86 (0.75, 0.96)	.795	0.88 (0.72, 1.00)	0.84 (0.68, 1.00)	.151
	Sensitivity	0.75 (0.36, 0.96)	0.88 (0.47, 0.99)	.317	0.80 (0.30, 0.99)	0.60 (0.17, 0.93)	.317
	Specificity	0.76 (0.66, 0.84)	0.77 (0.67, 0.85)	.835	0.65 (0.55, 0.75)	0.81 (0.71, 0.88)	.009
	PPV	0.21 (0.09, 0.41)	0.25 (0.11, 0.45)	.460	0.11 (0.04, 0.26)	0.14 (0.04, 0.37)	.454
	NPV	0.97 (0.89, 1.00)	0.99 (0.91, 1.00)	.310	0.98 (0.90, 1.00)	0.97 (0.90, 1.00)	.464
Total Learning	AUC	0.92 (0.84, 1.00)	0.95 (0.90, 1.00)	.576	0.95 (0.86, 1.00)	0.97 (0.94, 1.00)	.590
	Sensitivity	0.88 (0.47, 0.99)	0.88 (0.47, 0.99)	1	1.00 (0.46, 1.00)	1.00 (0.46, 1.00)	1
	Specificity	0.77 (0.67, 0.85)	0.85 (0.75, 0.91)	.127	0.76 (0.66, 0.84)	0.93 (0.85, 0.97)	.002
	PPV	0.25 (0.11, 0.45)	0.33 (0.15, 0.57)	.242	0.18 (0.07, 0.38)	0.42 (0.16, 0.71)	.020
	NPV	0.99 (0.91, 1.00)	0.99 (0.92, 1.00)	.947	1.00 (0.93, 1.00)	1.00 (0.95, 1.00)	1

Notes. 95% confidence intervals are presented in parentheses.

Abbreviations: AUC= area under the curve; CVLT-II = California Verbal Learning Test - II; NPV = negative predictive value; PPV: = positive predictive value; RAVLT = Rey Auditory Verbal Learning Test.