Article

Assessment
2019, Vol. 26(6) 976–983

DOI: 10.1177/1073191118763726
journals.sagepub.com/home/asm
SAGE

Factor Structure of the CVLT-II Short Form: Evidence From a Trauma-Exposed Sample

Anne R. Carlew[1], Keke L. Schuler[1], Camilo J. Ruggero[1], Jennifer L. Callahan[1], Benjamin J. Luft[2], and Roman Kotov[2]

Abstract

The current study sought to investigate the factor structure of the California Verbal Learning Test–Second Edition (CVLT-II) Short Form in a trauma-exposed sample. We used confirmatory factor analysis to test four competing models proposed by Donders in a study investigating the CVLT-II Standard Form. Consistent with Donders, a four-factor model consisting of Attention Span, Learning Efficiency, Delayed Memory, and Inaccurate Memory was supported. These results confirm the latent structure of the CVLT-II holds for the CVLT-II in its Short Form as well as in a trauma-exposed sample. Findings are particularly important, given previous research indicating attention span and learning efficiency may underpin memory complaints in trauma-exposed individuals.

Keywords

posttraumatic stress disorder, learning, memory, CVLT-II SF, trauma

The California Verbal Learning Test–Second Edition Short Form (CVLT-II SF; Delis, Kramer, Kaplan, & Ober, 2000) is a nine-item version of the CVLT-II, a highly regarded measure of both qualitative (e.g., encoding strategy) and quantitative (e.g., count of words learned/recalled) aspects of learning and memory. The CVLT-II SF was introduced along with the revised CVLT-II to enhance its clinical utility for use as a screening measure or for individuals with more severe cognitive deficits (Delis et al., 2000). Because of the relatively large amount of information it yields regarding the processes of learning and memory functioning, the CVLT-II is commonly used in assessment of clinical populations. Furthermore, distinct profiles of learning and memory performance on this measure have been associated with specific clinical conditions (DeJong & Donders, 2010; Greenaway et al., 2006). Relatively few studies, however, have reported findings using the CVLT-II SF, and no studies to our knowledge have examined its internal factor structure. Furthermore, it is important to understand the latent structure of the CVLT-II SF with clinical populations and confirm that the assumptions and psychometric properties of the CVLT-II extend to the CVLT-II SF and its use in clinical populations. For example, there is no confirmation in the literature that the semantic clustering variable actually captures a "learning strategy" in clinical populations.

Donders (2008) investigated the latent structure of the CVLT-II at three age levels reanalyzing the original normative sample. This study found a four-factor model consisting of Attention Span, Learning Efficiency, Delayed Memory, and Inaccurate Memory best fit the data for participants aged 16 to 60 years old, noting the model demonstrated less fit for older participants. For that reason, the current study sought to extend Donders' four-factor model to the CVLT-II SF in an adult sample, responders to the World Trade Center (WTC; average age of 54 years). This sample is of special interest because of prior research indicating that an unexpectedly high proportion of WTC responders have cognitive impairment (Clouston et al., 2016) as well as chronic posttraumatic stress disorder (PTSD; Bromet et al., 2016). Of specific interest, we investigated whether a similar factor structure would emerge in this trauma-exposed sample given evidence of disturbance in the attentional and executive components of learning associated with PTSD in other populations (Johnsen & Asbjørnsen, 2008).

While the CVLT-II has gained popularity among clinicians because of the information it yields, it takes nearly an hour to complete in its entirety, which may be prohibitive in some contexts and patient populations, including the WTC Health Program which provides annual monitoring

[1]University of North Texas, Denton, TX, USA
[2]Stony Brook University, Stony Brook, NY, USA

Corresponding Author:
Camilo J. Ruggero, Department of Psychology, University of North Texas, 1155 Union Circle #311280, Denton, TX 76203, USA.
Email: camilo.ruggero@unt.edu

examinations for ~40,000 responders (Dasaro et al., 2015). Abbreviated testing, such as the CVLT-II SF, may be preferable for several reasons. First, neuropsychologists are increasingly pressured to administer fewer/less time-consuming tests but deliver the same results by insurance companies who are reimbursing services at lower rates. Several short form measures have been shown to produce the same results as their lengthier counterparts (Greve, 2001; Keefe et al., 2004). The CVLT-II SF improves over other abbreviated measures of learning and memory (e.g., Hopkins Verbal Learning Test–Revised; Brandt & Benedict, 2001) with its inclusion of the same process variables that make the CVLT-II attractive in its Standard Form. The CVLT-II SF purportedly has the same abilities of the CVLT-II to investigate the attentional and executive components of learning while also parsing apart encoding, memory, and retrieval deficits. The CVLT-II SF also includes a performance validity measure, which has been shown to be of increasing importance to neuropsychological testing (e.g., Wisdom et al., 2014). Furthermore, lengthy testing may be inappropriate for populations like WTC responders due to testing fatigue, cognitive, psychiatric or physical impairments, or both. In some cases, CVLT-II SF is clinically preferable to mitigate testing fatigue and to reduce overpathologizing due to test burden, a problem relevant to several clinical populations (e.g., Bar-On Kalfon, Gal, Shorer, & Ablin, 2016). Alternatively, it may be used as a quick measure to rule out learning and memory deficits which in WTC responders could potentially explain PTSD symptom reports and poor performance on tests like the Montreal Cognitive Assessment (Clouston et al., 2016). Similarly, it may be preferable in research batteries due to time constraints. Furthermore, researchers have found using abbreviated measures in research batteries may limit missing data (Keefe et al., 2007) and contribute to completion of the battery (Keefe et al., 2004).

In the current study, we conducted a confirmatory factor analysis (CFA) on the CVLT-II SF with models drawn directly from Donders (2008) in a sample of WTC responders being monitored by the Stony Brook WTC Health Program (Bromet et al., 2016). Model 1 hypothesized the CVLT-II SF could be represented by a single, general memory factor, similar to the Full Scale Intelligence Quotient of the Wechsler Adult Intelligence Scale. Model 2 was composed of two factors, an Accurate Memory and an Inaccurate Memory factor. This model included the important distinction between accurate recall (i.e., hits) and inaccurate recall (e.g., intrusions and repetitions). Model 3 was a three-factor model that consisted of Immediate Memory, Delayed Memory, and Inaccurate Memory. This model carried the advantage of distinguishing between immediate memory and delayed memory, which seems to be important in some clinical samples (Lange, Chelune, Taylor, Woodward, & Heaton, 2006). Finally, Model 4 was a four-factor model, which hypothesized the presence of Attention Span, Learning Efficiency, Delayed Memory, and Inaccurate Memory factors. As mentioned earlier, this is the model found to best fit the CVLT-II normative data for participants aged 16 to 60 years. This model's addition of the Attention Span latent construct reflects the theory that span of attention is distinct from consolidated information (Donders, 2008). As noted, no studies have reported on the use of the CVLT-II SF in a trauma-exposed sample. Given the reports of attentional and executive components of learning as impaired in trauma-exposed individuals (e.g., Johnsen & Asbjørnsen, 2009), the WTC responder population provided an ideal group with which to investigate Donders' (2008) four-factor model, including the Attention Span subfactor. We hypothesized Model 4 would best fit our data given the implication of variability in attention among individuals exposed to trauma.

Method

Participants

The sample was composed of WTC responders from the Long Island site of WTC Health Program recruited for an ecological momentary assessment study on PTSD and respiratory problems between October 2014 and February 2016. Data used in the present study were collected at the baseline assessment, prior to the ecological momentary assessment portion of the study. The responders worked or volunteered as a part of rescue, recovery, restoration, or cleanup of the WTC sites (Dasaro et al., 2015). Two participants withdrew from the study after the baseline assessment and therefore were excluded, leaving a final sample of 205 participants in the present study. The final sample had a mean age of 54.39 (SD = 9.65) years at the time of assessment. The sample was primarily male (82.4%, n = 169), White (87.3%, n = 179), and non-Hispanic (81.0%, n = 166), with an average 14.79 (SD = 2.26) years of education. The sample was oversampled for PTSD symptoms. The score of the PTSD checklist for *Diagnostic and Statistical Manual of Mental Disorders–5* (*DSM-5*; Weathers et al., 2013) ranged from 20 to 88 with an average of 39.76 (SD = 17.07), and 18.5% (n = 38) of the sample carried a current diagnosis of PTSD at the time of baseline assessment. All participants provided written informed consent, and the study was approved by the Stony Brook University Institutional Review Board.

Assessments

The structured clinical interview for *DSM-IV* (First, Spitzer, Gibbon, & Williams, 1997) was used for PTSD diagnosis at the baseline assessment. Participants were instructed to refer to the WTC disaster during the interviews. The interviews were administrated by trained master-level interviewers, who were closely supervised by clinical psychologists (C.R. and R.K.). Previous assessments of interrater reliability in this population showed very good interrater agreement (κ = 0.82; Bromet et al., 2016).

The PTSD checklist for *DSM-5* (Weathers et al., 2013), a 20-item self-report measure, was used at the baseline assessment to assess *DSM-5* symptoms of PTSD. Participants were instructed to rate "How much you have been bothered by that problem in the past month" on a 5-point Likert-type scale from 1 (*not at all*) to 5 (*extremely*). The Cronbach's α of the item scores was .95.

The CVLT-II SF (Delis et al., 2000) was used at the baseline assessment to assess learning and memory functioning. The CVLT-II SF is a nine-item version of the CVLT-II, which was designed to measure learning strategies (i.e., semantic and serial clustering), free and cued recall, serial position effects (i.e., primacy and recency effects), intrusions, and interference. In the CVLT-II SF, there are four learning trials composed of nine words each in three semantic categories (i.e., fruits, clothing, and tools). Following the four learning trials, there is a 30-second delayed free recall trial, followed by a 10-minute delayed free recall, 10-minute delayed cued recall (using the semantic categories), and 10-minute delayed recognition (Delis et al., 2000). The CVLT-II SF was scored using the CVLT-II scoring assistant, which draws from the normative data in the published manual (Delis et al., 2000). The CVLT-II SF is sensitive to serial and semantic clustering during free recall. Thus, if a participant recalls several words from the same category sequentially, the participant is likely engaging in semantic organization. In contrast, if a participant recalls words in the same order as they were presented in the target list, the patient is likely engaging in serial organization. In the recognition section, distractor words may be semantically related, phonemically related, or completely unrelated. Improved performance on the recognition task subsequent to poor performance on free recall tasks indicated a deficit in retrieval rather than learning or encoding (Delis, Kramer, Kaplan, & Thompkins, 1987).

On the arrival for testing, participants were instructed to complete a battery of self-report questionnaires, then proceeded to the administration of the CVLT-II SF. The CVLT-II SF was administrated individually using standardized procedures. In brief, the examiner read the target list of words aloud in a fixed order over four learning trials at the rate of one word per second. After each trial, the participants were asked to repeat as many words as possible (i.e., free recall), in any order, including words they remembered from previous trials. After the fourth free recall trial, the examiner asked the participants to count backward from 30 to keep participants from reciting the target words subvocally.

Free recall of the target list was assessed immediately following the distraction task. Cued recall was subsequently tested, and the examiner asked the participants to recall words specifically from the four semantic categories included in the target list (e.g., "Name all of the words in the first list that are animals"). Free recall and cued recall were again tested after a delay of 10 minutes (hereafter referred to as: long delay). Next, the participants were presented with a recognition task in which the participant was asked to indicate whether a word is a target or distractor (yes/no recognition).

Table 1. Hypothetical Testing Models for CVLT- II SF.

CVLT-II SF variables	Model 1	Model 2	Model 3	Model 4
Trial 1 free recall	1	1	1	1
Trial 4 free recall	1	1	1	2
Short-delay free recall	1	1	2	3
Long-delay free recall	1	1	2	3
Long-delay cued recall	1	1	2	3
Semantic clustering	1	1	1	2
Middle region recall	1	1	1	1
Total intrusions	1	2	3	4
Recognition hits	1	1	2	3
Recognition false-positives	1	2	3	4

Note. CVLT-II SF = California Verbal Learning Test–Second Edition Short Form; Model 1 = a single general memory factor; Model 2 = 1—Accurate Memory and 2—Inaccurate Memory; Model 3 = 1—Immediate Memory, 2—Delayed Memory, and 3—Inaccurate Memory; Model 4 = 1—Attention Span, 2—Learning Efficiency, 3—Delayed Recall, and 4—Inaccurate Memory.

Analytic Plan

Four CFAs were conducted to examine the competing models presented in Table 1. Prior to testing the CFA models, data were inspected for multivariate normality using *MVN*, an R package (Korkmaz, Goksuluk, & Zararsiz, 2014). Results from the Mardia's multivariate normality test showed that data were not multivariate normal. Following the recommendations by Finney and DiStefano (2013), the maximum likelihood estimation with robust standard errors and a Satorra–Bentler scaled test statistic were used for each CFA. Factor variance was fixed to 1 in all the CFA models, which allowed free estimation of factor loadings. All analyses were conducted using *lavaan* package (Rosseel, 2012) in R.

In comparing these models, the following fit indices were considered: the χ^2 goodness-of-fit statistic, the comparative fit index (CFI), the Tucker–Lewis index (TLI), the root mean square error of approximation (RMSEA), the 90% confidence interval of the RMSEA, standardized root mean square residual, the Akaike information criterion, the Bayesian information criterion, and adjusted Bayesian information criterion (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004). Although recommended cutoff values for the model fit indices were not uniform in the broad literature,

Table 2. Descriptive Statistics and Correlation among the CVLT-II SF Indicators.

Indicators	*M* (*SD*)	1	2	3	4	5	6	7	8	9	10
1. Trial 1 free recall	−0.51 (1.29)	1.00									
2. Trial 4 free recall	−0.23 (0.94)	.26***	1.00								
3. Short-delay free recall	0.15 (1.13)	.28***	.57***	1.00							
4. Long-delay free recall	0.21 (0.93)	.30***	.53***	.67***	1.00						
5. Long-delay cued recall	0.01 (0.90)	.27***	.48***	.51***	.80***	1.00					
6. Semantic clustering	0.52 (1.38)	.21**	.27***	.25***	.26***	.24***	1.00				
7. Middle region recall	−0.18 (1.08)	.18*	.25***	.18**	.30***	.26***	.18*	1.00			
8. Total intrusions	0.46 (0.73)	−.06	−.06	−.24**	−.14*	−.08	.06	−.05	1.00		
9. Recognition hits	−0.25 (1.08)	.16*	.31***	.40***	.60***	.62***	.11	.21**	−.03	1.00	
10. Recognition false-positives	−0.16 (0.83)	−.01	−.14*	−.26***	−.32***	−.24**	.00	−.07	.29***	−.11	1.00

Note. CVLT-II SF = California Verbal Learning Test–Second Edition Short Form.
$*p < .05$. $**p < .01$. $***p < .001$.

Table 3. Goodness of Fit Indices for the Four CFA Models.

	df	χ^2	CFI	TLI	RMSEA	90% CI of RMSEA	SRMR	AIC	BIC	Adjusted BIC
Model 1	35	86.99	0.91	0.88	0.09	[0.07, 0.11]	0.06	5332.90	5432.62	5337.57
Model 2	34	73.88	0.93	0.91	0.08	[0.05, 0.10]	0.05	5321.11	5424.12	5325.90
Model 3	32	63.66	0.95	0.92	0.07	[0.05, 0.10]	0.04	5312.30	5421.95	5317.40
Model 4	29	62.00	0.94	0.91	0.08	[0.05, 0.10]	0.04	5317.91	5437.54	5323.48

Note. *N* = 205. *df* = degrees of freedom; CFA = confirmatory factor analysis; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; CI = confidence internal; SRMR = standardized root mean square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion. Model 1 = a single general memory factor; Model 2 = 1—Accurate Memory and 2—Inaccurate Memory; Model 3 = 1—Immediate Memory, 2—Delayed Memory, and 3—Inaccurate Memory; Model 4 = 1—Attention Span, 2—Learning Efficiency, 3—Delayed Recall, and 4—Inaccurate Memory.

the best available evidence support that CFI and TLI ≥ 0.90 indicates adequate fit and ≥ 0.95 indicate excellent fit; RMSEA ≤ 0.08 indicates adequate fit and ≤ 0.06 indicates excellent fit, when the value falls within the 90% confidence interval (Marsh et al., 2004).

Results

The descriptive statistics and zero-order correlations among the CVLT indicators are summarized in Table 2. The model fit indices for each CFA model are presented in Table 3. Standardized factor loadings and the correlations among factors in Model 4 are presented in Table 4. As seen in Table 3, Model 3 and Model 4 fit the data better than the other two models (i.e., Model 1 and Model 2) and had very similar model fit indices. Although Model 3 yielded a slightly better fit to the data based on three model fit indices (i.e., CFI, TLI, and RMSEA) and Attention Span and Learning Efficiency had a high correlation arguing against separation, the differences in model fit indices were not large enough to recommend Model 3 over Model 4. Importantly, Model 4 is consistent with findings from the CVLT-II standard form studies lending support for using Model 4 over Model 3 given no clear superiority in fit.

Table 4. Standardized Factor Loadings and Correlations Among Factors in Model 4.

Indicators	1	2	3	4
Trial 1 free recall	.44			
Middle region recall	.41			
Trial 4 free recall		.70		
Semantic clustering		.38		
Short-delay free recall			.70	
Long-delay free recall			.95	
Long-delay cued recall			.84	
Recognition hits			.64	
Total intrusions				.36
Recognition false-positives				.80
Factor correlations				
1. Attention Span	1.00			
2. Learning Efficiency	.90	1.00		
3. Delayed Recall	.74	.80	1.00	
4. Inaccurate Memory	−.13	−.21	−.40	1.00

The difference between the two models involves parsing the immediate memory factor into attention span and learning efficiency, which is essential considering the high rate of PTSD effects in the WTC sample. Given this difference, the factor loadings of these four items (i.e., Trial 1 free recall, Trial 4 free recall, semantic clustering, and middle region recall) between the two models were inspected. The four variables loaded slightly more strongly onto attention span (i.e., standardized factor loadings were .44 for Trial 1 free recall and .41 for middle region recall) and learning efficiency (i.e., standardized loadings were .70 for Trial 4 free recall and .38 for semantic clustering) in Model 4 than they did on Model 3 (i.e., standardized factor loadings were .41 for Trial 1 free recall, .69 for Trial 4 free recall, .38 for semantic clustering, and .38 for middle region recall), providing further support to the recommendation of Model 4.

Discussion

To investigate the factor structure of the CVLT-II SF in a trauma-exposed sample, we evaluated four models drawn from Donders's (2008) reanalysis of the CVLT-II standardization sample. We analyzed data from a sample of WTC responders in their mid-50s, many of whom have had significant PTSD symptoms. Using CFA, we evaluated the fit of four models, adding theoretically important latent factors to the first model composed of a general memory factor. Our results showed that the four-factor model originally proposed by Donders (2008), namely, Attention Span, Learning Efficiency, Delayed Recall, and Inaccurate Memory, was supported. A three-factor model also fit well and was more parsimonious given high correlation among two factors, but the four-factor model was retained given its consistency with the original, standard form of the CVLT and the lack of compelling contradictory evidence. It is important to note that Attention and Learning Efficiency are distinguishable conceptually in the long form of the CVLT-II, but that they are less distinguishable in the short form, which is perhaps not unexpected given how one relies on the other. The current model lends support to the delineation of theoretically defined components of verbal learning and memory. Our finding of an Attention Span factor in the WTC sample tested with the CVLT-II SF is important because Donders (2008) cautioned that this factor may be somewhat unreliable in clinical populations (citing previous studies as well as the original CVLT; Mottram & Donders, 2005; Wiegner & Donders, 1999) and encouraged future research such as the study presented here. Overall, the findings from the WTC sample are consistent with the finding that Attention Span emerged as a factor of memory in other clinical populations (e.g., Lynch, 2004). Indeed, variations in attention have been shown to negatively affect verbal learning and memory in other clinical populations (e.g., Andersen, Hovik, Skogli, Egeland, & Øie, 2013; Henkin et al., 2005). The factors of Attention Span and Learning Efficiency are particularly salient to a trauma-exposed population such as the WTC responders, given previous reports that these may be core cognitive deficits associated with trauma exposure (Scott et al., 2015; Uddo, Vasterling, Brailey, & Sutker, 1993). The two variables that make up the Learning Efficiency factor are semantic clustering and Trial 4 recall. Semantic clustering is regarded as an important executive aspect of learning that strongly predicts later retention (Gaines, Shapiro, Alt, & Benedict, 2006; Wittrock & Carter, 1975). Four variables loaded onto the Delayed Memory factor (short delay free recall, long delay free recall, long delay cued recall, and recognition hits). This factor differs slightly from Donders (2008) Delayed Memory factor because the CVLT-II SF does not include a short delay cued recall factor. Thus, it is important to note this factor still holds despite this omission. Finally, an Inaccurate Memory factor emerged which was composed of two variables: total intrusion and recognition false positives. Again, this finding is in line with Donders' (2008) findings.

Results of the current study may be extrapolated for use of the CVLT-II SF not just in the longitudinal WTC Health Program but also in health surveillance programs for responders and populations exposed to other toxic disasters, as well as general medical populations. The CVLT-II SF measures domains of cognitive functioning that are concerning to exposed populations, particularly memory complaints (Steele et al., 2000). A previous study of WTC first responders found PTSD symptoms, particularly reexperiencing symptoms, were related to cognitive impairment on a global cognitive screening measure (Clouston et al., 2016). Also within this study's sample, longitudinal increases in PTSD and depressive symptoms in the preceding years were observed only in patients with cognitive impairment. Indeed, while the relationship between PTSD symptoms and cognitive impairment is well-established in the literature, the impact of cognitive impairment on PTSD symptoms and vice versa in exposed samples is still unclear (Clouston et al., 2016; Neria, DiGrande, & Adams, 2011). From Clouston et al.'s study, it is evident that trauma-exposed populations are at a heightened risk of cognitive impairment, and that this impairment may in turn exacerbate PTSD symptoms. Assessing cognitive functioning should be a central component of longitudinal monitoring programs. Administering full cognitive assessments is impractical, but general cognitive screening measures such as the Montreal Cognitive Assessment or Mini Mental Status Exam may be too insensitive to detect more subtle impairments that may still be exacerbating PTSD symptoms (e.g., Pendlebury, Cuthbertson, Welch, Mehta, & Rothwell, 2010). Based on our results, the CVLT-II SF is an appropriate measure to detect the components of cognitive functioning that are known to be highly affected in trauma-exposed samples. Furthermore, it provides a more in depth picture of these domains which could enhance treatment planning and intervention efforts. Previous studies have already demonstrated this shorter measure is able to capture deficits in other clinical populations with a range of cognitive impairment (Berlau, Corrada, Peltz, & Kawas, 2012; Bettcher et al., 2016; Chau, 2010; Pa et al., 2010; van Stockum, MacAskill, Anderson, & Dalrymple-Alford, 2008; Young et al., 2010).

Results should be interpreted in light of limitations. First, the four- and three-factor solutions had similar model fit indices, although the latter more closely corresponds to the original measure. However, the four-factor solution led to two factors with a high degree of association. As a result, their validity on the short form at least should continue to be interpreted with some caution. Second, the study did not collect information on participants' history of attention deficit hyperactivity disorder, language disorders, or learning disabilities, all of which have shown associations with CVLT performance (e.g., Schwartz et al., 2016). To characterize how CVLT-II SF performance differs among different psychiatric disorders, future work is needed to examine these associations in diverse samples. Third, the sample was primarily male with a unique trauma history and the sample size was modest. However, other studies, particularly with clinical populations, have achieved good results with similar or smaller samples sizes (e.g., DeJong & Donders, 2010). Fourth, the present study did not include a comparison sample to formally examine the measurement invariance of the CVLT-II SF. However, given the lack of measurement structure work in the short form, the present study serves as a good starting point to test the structure of this particular measure. Further work is needed to formally investigate the measurement invariance of the CVLT-II SF using different clinical samples. Fifth, three of our factors (Attention Span, Learning Efficiency, and Inaccurate Memory) are composed of two indicators due to the fewer variables yielded by the CVLT-II SF. This is fewer than recommended for latent variable identification (Kline, 2010). Thus, the stability of factor structure revealed in the present study may be attenuated. Future replication and longitudinal assessments are needed to explicitly examine the stability of factor structure underlying CVLT-II SF. Finally, in Donders's original paper, it was found that his four-factor model (our Model 4) fit "less well" with the older adults group aged 60+ years. It is important to acknowledge that our sample's age could be influencing the current results, though the mean age (i.e., 54.39) does fall below that of Donders's older adults group.

Overall, results of this study confirm the factor structure of the CVLT-II SF in WTC responders, and the results mirror the factor structure of the standard format CVLT-II in the standardization sample. Given reports that learning efficiency may be implicated in learning and memory deficits in individuals with PTSD (Johnsen & Asbjørnsen, 2008), and cognitive concerns expressed by populations exposed to neurotoxic substances, replicability of the factor structure in the short form provides support for its clinical and research use in trauma-exposed populations.

Acknowledgments

We gratefully acknowledge the support of the WTC responders for generously contributing their time and energy to this project. We also thank Evelyn Bromet, Adam Gonzalez, and Katherine Guerrera for facilitating the study, as well as study staff: Melissa Carr and Rachel Roger, Chris Ray and Peter Allen. Special thanks to Evelyn Bromet for her thoughtful comments and edits to the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study was supported by the Center for Disease Control and Prevention (CDC)/NIOSH grants 1U01 OH010712-01 and 1U01OH011321-01 awarded to Roman Kotov.

ORCID iD

Camilo J. Ruggero https://orcid.org/0000-0002-8482-6269

References

Andersen, P. N., Hovik, K. T., Skogli, E. W., Egeland, J., & Øie, M. (2013). Symptoms of ADHD in children with high-functioning autism are related to impaired verbal working memory and verbal delayed recall. *PLoS ONE*, *8*, e64842. doi:10.1371/journal.pone.0064842

Bar-On Kalfon, T., Gal, G., Shorer, R., & Ablin, J. N. (2016). Cognitive functioning in fibromyalgia: The central role of effort. *Journal of Psychosomatic Research*, *87*, 30-36. doi:10.1016/j.jpsychores.2016.06.004

Berlau, D. J., Corrada, M. M., Peltz, C. B., & Kawas, C. H. (2012). Disability in the oldest-old: Incidence and risk factors in the 90+ study. *American Journal of Geriatric Psychiatry*, *20*, 159-168. doi:10.1097/JGP.0b013e31820d9295

Bettcher, B. M., Fitch, R., Wynn, M. J., Lalli, M. A., Elofson, J., Jastrzab, L., . . . Kramer, J. H. (2016). MCP-1 and eotaxin-1 selectively and negatively associate with memory in MCI and Alzheimer's disease dementia phenotypes. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *3*, 91-97. doi:10.1016/j.dadm.2016.05.004

Brandt, J., & Benedict, R. H. (2001). *Hopkins verbal learning test, revised: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Bromet, E. J., Hobbs, M. J., Clouston, S. A. P., Gonzalez, A., Kotov, R., & Luft, B. J. (2016). DSM-IV post-traumatic stress disorder among World Trade Center responders 11-13 years after the disaster of 11 September 2001 (9/11). *Psychological Medicine*, *46*, 771-783. doi:10.1017/S0033291715002184

Chau, P. M. (2010). *Laterality of motor symptom onset, disease progression, and cognition in Parkinson's disease* (Doctoral dissertation). Retrieved from https://kuscholarworks.ku.edu/handle/1808/11159

Clouston, S. A., Kotov, R., Pietrzak, R. H., Luft, B. J., Gonzalez, A., Richards, M., . . . Bromet, E. J. (2016). Cognitive impairment among World Trade Center responders: Long-term implications of re-experiencing the 9/11 terrorist attacks. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *4*, 67-75. doi:10.1016/j.dadm.2016.08.001

Dasaro, C. R., Holden, W. L., Berman, K. D., Crane, M. A., Kaplan, J. R., Lucchini, R. G., . . . Udasin, I. G. (2015). Cohort profile: World trade center health program general responder cohort. *International Journal of Epidemiology*, *46*, e9. doi:10.1093/ije/dyv099

DeJong, J., & Donders, J. (2010). Cluster subtypes on the California Verbal Learning Test–Second Edition (CVLT–II) in a traumatic brain injury sample. *Journal of Clinical and Experimental Neuropsychology*, *32*, 953-960. doi:10.1080/13803391003645640

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test* (2nd ed.). San Antonio, TX: Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Thompkins, B. A. O. (1987). *CVLT: California Verbal Learning Test-Adult Version: Manual*. San Antonio, TX: Psychological Corporation.

Donders, J. (2008). A confirmatory factor analysis of the California Verbal Learning Test—Second Edition (CVLT-II) in the standardization sample. *Assessment*, *15*, 123-131. doi:10.1177/1073191107310926

Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *A second course in structural equation modeling* (2nd ed., pp. 439-492). Charlotte, NC: Information Age.

First, M.B., Gibbon M., Spitzer R.L., Williams, J.B.W., Benjamin L.S. (1997). *Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II)*. Washington, D.C.: American Psychiatric Press, Inc.

Gaines, J. J., Shapiro, A., Alt, M., & Benedict, R. H. (2006). Semantic clustering indexes for the Hopkins Verbal Learning Test-Revised: Initial Exploration in elder control and dementia groups. *Applied Neuropsychology*, *13*, 213-222. doi:10.1207/s15324826an1304_2

Greenaway, M. C., Lacritz, L. H., Binegar, D., Weiner, M. F., Lipton, A., & Cullum, C. M. (2006). Patterns of verbal memory performance in mild cognitive impairment, Alzheimer disease, and normal aging. *Cognitive and Behavioral Neurology*, *19*, 79-84. doi:10.1097/01.wnn.0000208290.57370.a3

Greve, K. W. (2001). The WCST-64: A standardized short-form of the Wisconsin Card Sorting Test. *The Clinical Neuropsychologist*, *15*, 228-234. doi:10.1076/clin.15.2.228.1901

Henkin, Y., Sadeh, M., Kivity, S., Shabtai, E., Kishon-Rabin, L., & Gadoth, N. (2005). Cognitive function in idiopathic generalized epilepsy of childhood. *Developmental Medicine & Child Neurology*, *47*, 126-132. doi:10.1017/S0012162205000228

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1-55. doi:10.1080/10705519909540118

Johnsen, G. E., & Asbjørnsen, A. E. (2008). Consistent impaired verbal memory in PTSD: A meta-analysis. *Journal of Affective Disorders*, *111*, 74-82. doi:10.1016/j.jad.2008.02.007

Johnsen, G. E., & Asbjørnsen, A. E. (2009). Verbal learning and memory impairments in posttraumatic stress disorder: The role of encoding strategies. *Psychiatry Research*, *165*, 68-77. doi:10.1016/j.psychres.2008.01.001

Keefe, R. S., Bilder, R. M., Davis, S. M., Harvey, P. D., Palmer, B. W., Gold, J. M., . . . McEvoy, J. P. (2007). Neurocognitive effects of antipsychotic medications in patients with chronic schizophrenia in the CATIE Trial. *Archives of General Psychiatry*, *64*, 633-647. doi:10.1001/archpsyc.64.6.633

Keefe, R. S., Goldberg, T. E., Harvey, P. D., Gold, J. M., Poe, M. P., & Coughenour, L. (2004). The Brief Assessment of Cognition in Schizophrenia: Reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophrenia Research*, *68*, 283-297. doi:10.1016/j.schres.2003.09.011

Kline, R. B. (2010). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *R Journal*, *6*(2), 151-162.

Lange, R. T., Chelune, G. J., Taylor, M. J., Woodward, T. S., & Heaton, R. K. (2006). Development of demographic norms for four new WAIS-III/WMS-III indexes. *Psychological Assessment*, *18*, 174-181. doi:10.1037/1040-3590.18.2.174

Lynch, P. A. (2004). Factor structure of the Wide Range Assessment of Memory and Learning (WRAML) in children with insulin dependent diabetes mellitus (IDDM). *Child Neuropsychology*, *10*, 306-317. doi:10.1080/09297040490909305

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320-341. doi:10.1207/s15328007sem1103_2

Mottram, L., & Donders, J. (2005). Construct validity of the California Verbal Learning Test–Children's Version (CVLT-C) after pediatric traumatic brain injury. *Psychological Assessment*, *17*, 212-217. doi:10.1037/1040-3590.17.2.212

Neria, Y., DiGrande, L., & Adams, B. G. (2011). Posttraumatic stress disorder following the September 11, 2001, terrorist attacks: A review of the literature among highly exposed populations. *American Psychologist*, *66*, 429-446. doi:10.1037/a0024791

Pa, J., Possin, K. L., Wilson, S. M., Quitania, L. C., Kramer, J. H., Boxer, A. L., . . . Johnson, J. K. (2010). Gray matter correlates of set-shifting among neurodegenerative disease, mild cognitive impairment, and healthy older adults. *Journal of the International Neuropsychological Society*, *16*, 640-650. doi:10.1017/S1355617710000408

Pendlebury, S. T., Cuthbertson, F. C., Welch, S. J., Mehta, Z., & Rothwell, P. M. (2010). Underestimation of cognitive impairment by Mini-Mental State Examination versus the Montreal Cognitive Assessment in patients with transient ischemic attack and stroke. *Stroke*, *41*, 1290-1293. doi:10.1161/STROKEAHA.110.579888

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1-36. doi:10.18637/jss.v048.i02

Schwartz, E. S., Erdodi, L., Rodriguez, N., Ghosh, J. J., Curtain, J. R., Flashman, L. A., & Roth, R. M. (2016). CVLT-II Forced Choice Recognition trial as an embedded validity indicator: A systematic review of the evidence. *Journal of the International Neuropsychological Society*, *22*, 851-858.

Scott, J. C., Matt, G. E., Wrocklage, K. M., Crnich, C., Jordan, J., Southwick, S. M., . . . Schweinsburg, B. C. (2015). A quantitative meta-analysis of neurocognitive functioning in posttraumatic stress disorder. *Psychological Bulletin*, *141*, 105. doi:10.1037/a0038039

Steele, L. (2000). Prevalence and patterns of Gulf War illness in Kansas veterans: Association of symptoms with characteristics of person, place, and time of military service. *American Journal of Epidemiology*, *152*, 992-1002. doi:10.1093/aje/152.10.992

Uddo, M., Vasterling, J. J., Brailey, K., & Sutker, P. B. (1993). Memory and attention in combat-related post-traumatic stress disorder (PTSD). *Journal of Psychopathology and Behavioral Assessment*, *15*, 43-52. doi:10.1007/BF00964322

van Stockum, S., MacAskill, M., Anderson, T., & Dalrymple-Alford, J. (2008). Don't look now or look away: Two sources of saccadic disinhibition in Parkinson's disease? *Neuropsychologia*, *46*, 3108-3115. doi:10.1016/j.neuropsychologia.2008.07.002

Weathers, F.W., Litz, B.T., Keane, T.M., Palmieri, P.A., Marx, B.P., & Schnurr, P.P. (2013). The PTSD Checklist for DSM-5 (PCL-5). Scale available from the National Center for PTSD. Retrieved from: www.ptsd.va.gov.

Wiegner, S., & Donders, J. (1999). Performance on the California Verbal Learning Test after traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, *21*, 159-170. doi:10.1076/jcen.21.2.159.925

Wisdom, N. M., Pastorek, N. J., Miller, B. I., Booth, J. E., Romesser, J. M., Linck, J. F., & Sim, A. H. (2014). PTSD and cognitive functioning: Importance of including performance validity testing. *The Clinical Neuropsychologist*, *28*, 128-145. doi:10.1080/13854046.2013.863977

Wittrock, M. C., & Carter, J. F. (1975). Generative processing of hierarchically organized words. *American Journal of Psychology*, *88*, 489-501. doi:10.2307/1421780

Young, K., Du, A., Kramer, J., Rosen, H., Miller, B., & Schuff, N. (2010). Patterns of structural complexity in Alzheimer's disease and frontotemporal dementia. *Human Brain Mapping*, *30*, 1667-1677. doi:10.1002/hbm.20632