Psychological Assessment

Measuring Working Memory in the Spanish Population: Validation of a Multiple Shortened Complex Span Task

Mireia Felez-Nobrega, Jeffrey L. Foster, Anna Puig-Ribera, Christopher Draheim, and Charles H. Hillman

Online First Publication, April 13, 2017. http://dx.doi.org/10.1037/pas0000484

Psychological Assessment

1040-3590/17/$12.00 http://dx.doi.org/10.1037/pas0000484

BRIEF REPORT

Measuring Working Memory in the Spanish Population: Validation of a Multiple Shortened Complex Span Task

Mireia Felez-Nobrega
University of Vic–Central University of Catalonia

Jeffrey L. Foster
Western Sydney University

Anna Puig-Ribera
University of Vic–Central University of Catalonia

Christopher Draheim
Georgia Institute of Technology

Charles H. Hillman
Northeastern University

Working memory plays a key role in cognition as it is a major predictor of a wide range of higher order abilities and behaviors typical to daily life. Shorter versions of the complex span tasks (CSTs) have been recently developed, allowing for the reduction of test administration time without affecting validity and reliability in the measurement of working memory capacity (WMC). However, these short versions have not been validated for the Spanish-speaking population. The present work aimed to validate an English version of the shortened CSTs into Spanish in a sample of 325 university students (40% female; mean age = 21.04; *SD* = 2.80). Cronbach's coefficient alpha was computed for each complex span task as an index of internal consistency. Validity evidence was evaluated by comparing participants' scores on the three shortened complex span tasks (operation span, symmetry span, and rotation span) with two measures of reasoning ability (Raven's Advanced Progressive Matrices and Number Series) and using confirmatory factor analysis. Results indicated that the short version of the Spanish complex span has satisfying qualities for assessing WMC in a sample of university students, which is an initial step toward providing a valid and standardized method for assessing WMC in the Peninsular Spanish-speaking population.

Public Significance Statement
This study advances the idea of providing a standardized method for assessing working memory capacity in Peninsular Spanish-speaking populations. The tool is based on a shortened format of the complex span task paradigm.

Keywords: working memory, complex span tasks, measurement, short version, validation

Working memory refers to a "limited-capacity system responsible for active maintenance, manipulation, and retrieval of task-relevant information that is needed for on-going cognition" (Unsworth, Redick, Heitz, Broadway, & Engle, 2009). Measures of WMC have been used in clinical, educational, personality, and developmental research disciplines because of the fundamental role WMC plays in predicting performance on a broad range of higher order capabilities (for reviews, see Conway, Jarrold, Kane, Miyake, & Towse, 2007; Engle & Kane, 2004; Unsworth et al., 2009) and managing rational and emotional situations in daily life (e.g., Cohen & Conway, 2008; Kane et al., 2007).

Mireia Felez-Nobrega, Sport and Physical Activity Research Group, Centre for Health and Social Care Research, University of Vic–Central University of Catalonia; Jeffrey L. Foster, School of Social Sciences and Psychology, Western Sydney University; Anna Puig-Ribera, Sport and Physical Activity Research Group, Centre for Health and Social Care Research, University of Vic–Central University of Catalonia; Christopher Draheim, Department of Psychology, Georgia Institute of Technology; Charles H. Hillman, Department of Psychology and Department of Health Sciences, Northeastern University.

This work has been supported by the FI-AGAUR Predoctoral Research Fellowship Program, Generalitat de Catalunya, 2014FI_B 00339.

Correspondence concerning this article should be addressed to Mireia Felez-Nobrega, Sport and Physical Activity Research Group, Centre for Health and Social Care Research, University of Vic–Central University of Catalonia, Pol, 1, 08500, Vic, Barcelona, Spain. E-mail: mireia.felez@uvic.cat

A wide variety of cognitive tasks are used to measure WMC, yet some may be more associated with short-term memory than WMC (Engle, Tuholski, Laughlin, & Conway, 1999). Simple span tasks (Turner & Engle, 1989) such as digit span, word span, or visuospatial span tasks, require holding in mind a short list of words, letters, positions, or digits with no requirement of transforming the information (Lezak, 1983; Rosen & Engle, 1997). Likewise, recalling items in the reverse order (e.g., backward-digit) arguably does not require more complex processing than forward recall (Rosen & Engle, 1997; St Clair-Thompson, 2010; St Clair-Thompson & Allen, 2013). In this context, validity of a WMC measurement using simple span tasks is controversial.

Contrarily, complex span tasks (CSTs; Daneman & Carpenter, 1980; Turner & Engle, 1989) are a different paradigm to measure WMC. CSTs interleave a processing task (e.g., operation span [OSpan]—solving simple equations; symmetry span [SymSpan]—judging image symmetries; rotation span [RotSpan]—judging rotated letters) with a short list of to-be-remembered items (e.g., OSpan, letters; SymSpan, square positions; RotSpan, arrows) that allow for the assessment of a dynamic working memory (WM) process, which involves both processing and storage capacity (Turner & Engle, 1989). Furthermore, although the nature of the storage component (domain-specific factors such as spatial, verbal, or numeric ability) can influence the WM variance score, CSTs are considered a measure of the underlying cognitive ability (Hambrick, Kane, & Engle, 2005; Kane et al., 2004). This general-capacity hypothesis is also supported in Spanish population studies (Pardo-Vazquez & Fernandez-Rey, 2012).

Recent studies have argued that simple and CSTs may partially tap the same construct and underlying mechanisms (Bailey, Dunlosky, & Kane, 2011; Unsworth & Engle, 2007), but although clinical psychologists often use psychometric indices such as subscales of the Wechsler Adult Intelligence Scale and the Wechsler Memory Scale (Camara, Nathan, & Puente, 2000; Rabin, Barr, & Burton, 2005; Shelton, Elliott, Hill, Calamia, & Gouvier, 2009), in experimental psychology settings, CSTs are the most widely accepted instruments to measure WMC, and their use is supported by standardized administration and scoring procedures (Conway et al., 2005; Redick et al., 2012). Importantly, obtaining an accurate measure of WMC requires the use of more than one WMC task because otherwise, scores on any single indicator would be driven by the ability of interest (WMC) and other systematic and random influences such as the nature of the storage component itself (Conway et al., 2005; Foster et al., 2015; Unsworth et al., 2009). Nonetheless, this recommendation entails an extended length of time for both researchers and participants. Long and tedious experimental sessions can lead to decreases in participant motivation and a reduction of attention resources that negatively influence CST performance (Heitz, Schrock, Payne, & Engle, 2008). In order to manage this issue, recent studies have developed shorter versions of the CSTs, allowing the reduction of test administration time without affecting the accuracy measurement of the construct (Foster et al., 2015; Gonthier, Thomassin, & Roulin, 2016; Oswald, McAbee, Redick, & Hambrick, 2015).

Most available CSTs—both long and short versions—have been developed and validated for English- and French-speaking populations, but not Spanish (Conway et al., 2005; Foster et al., 2015; Gonthier et al., 2016; Oswald et al., 2015; Redick et al., 2012; Unsworth, Heitz, Schrock, & Engle, 2005). Sanchez et al. (2010) found that the English version of the OSpan accurately measured WMC for both native-English and nonnative populations (native-Spanish speakers who reported having spoken fluent English for at least 10 years) while the English version of the reading span was not a valid task in nonnative English speakers. Thus, linguistic status can have an influence on CST scores.

To our knowledge, only an automated group-administrable OSpan version that is based on a Dutch OSpan task by De Neys, D'Ydewalle, Schaeken, and Vos (2002) has been developed and validated in Spanish-speaking samples (Pardo-Vázquez & Fernández-Rey, 2008). Because population linguistic status needs to be considered, the Spanish-speaking population is limited in using CSTs to measure WMC. Thus, in the need for using multiple indicators to obtain an accurate measure of WMC and the scarcity of available and validated measures of WMC for Spanish-speaking populations, the present study aims to provide a time-effective measurement of WMC validating the shortened complex span version in Spanish university students.

Method

The aim of linguistic validation is to obtain translations that are conceptually equivalent to the original, comparable across languages, and easily understood by the people to whom the translated instrument is administered (Hambleton, 2005). To adapt these tasks for Spanish populations, the instructions of the CSTs were translated following the guidelines and recommendations of Hambleton (2005).

We used independent forward and backward translators who were experienced in translating health instruments and native speakers of the target languages. The English version was first translated into Spanish by one Spanish native-speaking translator. The first draft of the Spanish version was then back-translated into English by an English native-speaking translator. Finally, both translators reviewed and discussed the back-translation to identify any discrepancies between the meaning of the translation and the original instructions.

In total, 361 undergraduate students were recruited from the University of Vic (Spain) aged 18–36 years (M = 21.04, SD = 2.80). Participants were progressively recruited from April 2015 to March 2016 through personal invitation via email, placards, or study participation included in course curricula. Eligibility criteria included being native Spanish or native bilingual Catalan-Spanish speakers (i.e., individuals are considered bilingual when a Spanish citizen speaks a local official language in addition to Spanish) and not suffering from neurological disorders. The study was approved by the Research Commission of University of Vic, and all undergraduate students gave written informed consent to participate in the study.

The shortened original version of the CSTs was created by Foster and colleagues (2015). This measure includes three complex span subtests: OSpan, SymSpan, and RotSpan. These tasks are divided into three blocks of trials that maintained similar measurement characteristics (Foster et al., 2015). Analyzing different combinations of tasks and blocks, the authors show 39 models indicating the percentage of the fluid intelligence (Gf) variance that each model can predict. This based-on-blocks methodology allows researchers to select the number of blocks for each task considering the percentage of the Gf predicted and the length of administration time needed (for detailed information, see Foster et al., 2015).

Given the importance of using more than one WMC indicator to obtain an accurate measure of WMC (Conway et al., 2005; Foster et al., 2015), all participants completed the Spanish version of the three shortened CSTs (OSpan, SymSpan, RotSpan) and two measures of Gf (Raven's Advanced Progressive Matrices [RAPM] and Number Series). Considering Foster's previous work, participants completed two blocks of the OSpan, two blocks of the SymSpan, and three blocks of the RotSpan, which is represented as Model 29 (accounts for 98.6% of the predicted Gf variance and can be administered in less than 55 min). This model was chosen because it affords the ability to obtain the minimum best recommendation of 90% factor variance while also allowing for the completion of the WMC subtests in less than 1 hr. Data collection was conducted in one single 60-min session. All tasks were group administered with the presence of a researcher supervisor and computer paced. The five tasks were presented following the order listed above.

Tasks

OSpan (Redick et al., 2012; Turner & Engle, 1989). Participants mentally solved a simple math operation and decided whether a suggested answer was correct or incorrect. Subsequently, a letter (that was the item to-be-remembered) appeared on the screen (1,000-ms duration for the storage components in all WMC measures). After three to seven randomly presented math-letter sequences, participants were asked to recall—in correct serial order—the letters that they had seen by clicking on a letter grid displayed on the screen. Each block contained each of one trial with 3, 4, 5, 6, and 7 math-letter sequences totaling 25 letters and math problems per block (50 across two blocks). The study trial was preceded by an instruction phase where first the letters were shown alone, a second trial where the mathematic equations were shown alone, and a third trial where both items were paired. During this second practice trial, participants were timed and were required to respond within 2.5 standard deviations (*SD*s) of their average response time to each distraction item (Conway et al., 2005; Redick et al., 2012). Participants received feedback on their level of accuracy for the distractor task (mathematic equation).

SymSpan (Kane et al., 2004; Redick et al., 2012). This task followed the same methodology as the OSpan with a practice phase preceding study trials where participants received feedback on their level of accuracy for the distractor task. The items to be remembered were locations of red squares shown in a 4 × 4 grid while the distractor task was deciding whether a black image pattern was symmetrical along its vertical axis. After two to five randomly presented image-square sequences, participants were asked to recall in correct serial order the square positions by clicking on a grid template. Each block contained a total of 14 image-symmetry judgments (28 across two blocks).

RotSpan (Kane et al., 2004). Similar to the other two tasks, a practice phase preceded the study trials. The items to be remembered were arrows of varying lengths (short or long) and direction in which they point. The distractor was deciding whether a rotated letter (i.e., "G" "F" or "L" "J" "R") was presented normally or mirror-reversed. After two to five randomly presented letter-arrow sequences, participants were asked to recall—in correct serial order—the length and direction of the arrows they had seen by clicking on a grid template. Each block contained a total of 14 arrow-letter judgments (42 across three blocks).

The scores for the three tasks were calculated by summing the total number of items recalled in the correct order for every trial (i.e., letters for the OSpan, square positions for the SymSpan, and arrows for the RotSpan), which is referred to as the partial span score (Conway et al., 2005; Turner & Engle, 1989). Finally, the common variance from the three complex span tasks was extracted by calculating the average of the current sample *z* scores for each of the three tasks (further referred to as the WMC factor score).

Fluid Intelligence Measures

RAPM (Raven, Raven, & Court, 1998). A computer-administered Spanish version of the RAPM was used as a measure of Gf. The task consisted of the 18 odd-numbered items to be resolved in a maximum of 10 min (this methodology is consistent with previous research, e.g., Foster et al., 2015; Kane et al., 2004). For each item, participants see a 3 × 3 matrix of shapes with the bottom-right pattern missing. The participant's task is to figure out the logical rule of the matrix and then click the correct item among the eight possible options. Items are presented in ascending difficulty, and three practice trials preceded the study trials. Scores were calculated by summing the number of correct answers.

Number Series (Thurstone, 1938). In this computer-based Spanish version, participants view a series of numbers. The participant's task is to decipher the logical rule in order to choose which of the five numbers presented would continue the numerical series. Three practice items preceded the study trials. There were 15 items presented in ascending difficulty and a maximum of 10 min to complete the task. Scores were calculated by summing the number of items solved correctly.

The range of scores for the WMC and Gf measures were OSpan = 0–50, SymSpan = 0–28, RotSpan = 0–42, RAPM = 0–18, and Number Series = 0–15.

Cronbach's coefficient alpha was computed for each complex span task as an index of internal consistency. For validity, it was assessed whether measures of fluid intelligence and the shortened CSTs constituted distinct but related constructs by comparing Pearson correlations between the three CSTs subtests (OSpan, SymSpan, and RotSpan) and the two measures of Gf (RAPM, Number Series) using SPSS Statistics 23.0. Further, confirmatory factor analyses and structural equation models were used to analyze the structure of the data and the relationship between constructs. Latent variables were WMC with three shortened span measures as indicators (OSpan, SymSpan, and RotSpan) and Gf with two measures as indicators (RAPM and Number Series). We report the chi-square, the comparative fit index (CFI), root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMS). Indicative values of an acceptable model fit are as follows: CFI > 0.95, SRMR < 0.05, and RMSEA < 0.06 (Hu & Bentler, 1999; Kline, 1998). To conduct these analyses, we used the statistical package STATA.

Results

A total of 36 participants were excluded from analysis. Sixteen participants were eliminated because they reported not being either native Spanish or native bilingual Catalan-Spanish speakers, 10

Table 1
Descriptive Statistics for Working Memory and Gf Scores

Measure	*M*	*SD*	Skew	Kurtosis
Operation span				
Partial score	33.33 (35.59)	8.16 (10.81)	−.86	1.21
Symmetry span				
Partial score	19.71 (17.73)	4.36 (6.23)	−.46	.06
Rotation span				
Partial score	28.76 (24.56)	6.66 (9.84)	−.59	.11
RAPM correct	7.99	3.24	.06	−.43
Number correct	8.65	3.12	−.15	−.72

Note. Operation span = Blocks 1 and 2; symmetry span = Blocks 1 and 2; rotation span = Blocks 1, 2, 3. Range scores are as follows: operation span = 0–50; symmetry span = 0–28; rotation span = 0–42; RAPM = 0–18; Number Series = 0–15. Foster et al.'s (2015) partial scores and *SD* for the same number of blocks in each task are shown in parentheses. RAPM = Raven's Advanced Progressive Matrices.

participants because they were not following task instructions as observed by the experimenter, 6 because they did not complete all five tasks, 2 participants for technical problems, and 2 participants due to incorrect selection of the number of blocks for each task. The final data set included 325 participants (40% female).

A descriptive analysis of score distribution (see Table 1) showed that storage scores are normally distributed with values of skewness and kurtosis under the generally accepted values (skewness $<$ 2 and kurtosis $<$ 4; Kline, 1998).

For internal consistency, the values of Cronbach's alpha were overall acceptable (OSpan, $\alpha = .72$; RotSpan, $\alpha = .72$). The only measure with values substantially lower than the recommended level of 0.70 (Nunnally, 1978) was the SymSpan ($\alpha = .52$).

In addition, although our interpretation of the shortened CSTs focuses on latent-variable analyses (see below), we first explored Pearson correlations between the three CST subtests and Gf tasks (see Table 2). The three CSTs showed moderate correlations with each other (0.32–0.46), and the WMC factor score showed high correlations with all three CSTs (0.74–0.78). Additionally, the three CSTs generally correlated more strongly with each other than with the Gf tasks (0.18–0.40), and the same trend was observed for Gf correlation coefficients, which suggests that these two sets of tasks measured related but distinct abilities, a finding consistent with previous studies employing WMC measures (Conway et al., 2005; Foster et al., 2015).

We next analyzed whether the data were better represented by a single-factor model or a two-factor model (see Figure 1) differentiating between two latent variables: WMC (three shortened span measures as indicators) and Gf (RAPM and Number Series tasks as indicators).

Following the indices of model fit (Hu & Bentler, 1999; Kline, 1998), the two-factor model ($\chi^2(4) = 9.71$, $p = .05$; CFI = .98; RMSEA = .07; SRMR = .02) fit the data better than the one-factor model ($\chi^2(5) = 38.90$, $p < .01$; CFI = .89; RMSEA = .14; SRMR = .05).

To further corroborate this hypothesis, the two models were compared, and the difference in chi-square for the two models was significant, $\Delta\chi^2(1) = 29.19$, $p < .01$, suggesting that the two-factor model was a better fit.

The endorsed two-factor model is shown in Figure 1. Three important features are evident; first, each complex span measure had a strong positive loading on the latent WMC factor (0.50–0.72). Second, factor loadings of the tasks measuring Gf also loaded well on the Gf factor (0.66–0.77). Finally, there was a strong relationship between WMC and Gf ($r = .69$), which fits squarely in the expected range of this relationship established in other research (roughly 0.6–0.8; see Ackerman, Beier, & Boyle, 2005; Kane, Hambrick, & Conway, 2005). In short, this model suggests that the translated versions of the CSTs were able to effectively measure WMC as an ability that is distinct but related to Gf.

Discussion

This study aimed to validate an English shortened version of the CSTs into Spanish (Castilian), providing a time-effective measure of WMC for this population. The results show that the Spanish shortened version of the CSTs is a valid tool to assess domain-general WMC in a sample of Spanish-speaking university students.

Internal consistencies of the Spanish version are lower (OSpan, $\alpha = -0.09$; SymSpan, $\alpha = -0.23$; RotSpan, $\alpha = -0.14$) compared to the reference study (Foster et al., 2015). Lower alpha coefficients could be attributed to a restricted range of general abilities (participants obtained overall higher mean scores and smaller *SD*s; see Table 1).

Correlations between the three CSTs are similar to other studies using shortened French CSTs versions with university students as participants (Gonthier et al., 2016) but lower compared to the reference study (Foster et al., 2015). Lower correlations could also be attributed to a restricted range of scores obtained in the present study. In addition, CST loadings on the latent WMC factor were similar to other shortened CSTs (0.43–0.86; Foster et al., 2015; Gonthier et al., 2016; Oswald et al., 2015).

Overall, the psychometric properties of the tool are within the range expected based on previous research using shortened CSTs (Foster et al., 2015; Gonthier et al., 2016; Oswald et al., 2015). However, it is important to note that the validation evidence provided here is based on the use of two blocks of trials in the OSpan and SymSpan and three blocks of trials in the RotSpan. Alterations in choosing the number of blocks may lead to variations in the validity of the measure.

This study has several limitations. The shortened Spanish version has demonstrated satisfying qualities for assessing WMC, but although we aimed to ensure variability of domain-specific cognitive abilities by recruiting participants from different schools that represent a wider selection of undergraduate studies, the study was

Table 2
Pearson Correlations Among Tasks

Characteristic	1	2	3	4	5	6
1. Factor score WMC	—					
2. OSpan	.74	—				
3. SymSpan	.78	.38	—			
4. RotSpan	.77	.32	.46	—		
5. RAPM	.35	.18	.27	.36	—	
6. Number Series	.43	.30	.29	.40	.50	—

Note. WMC = working memory capacity; OSpan = operation span; SymSpan = symmetry span; RotSpan = rotation span; RAPM = Raven's Advanced Progressive Matrices.

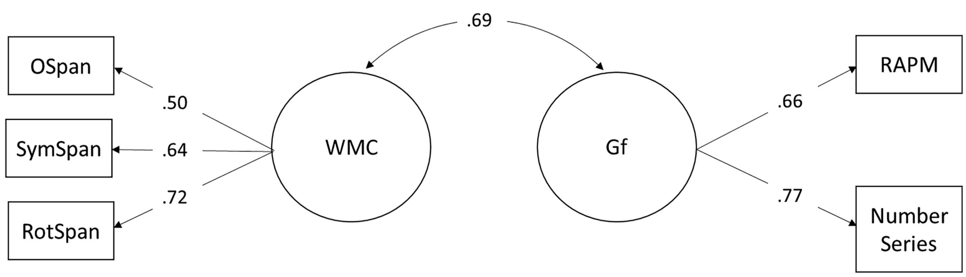

Figure 1. Structural equation model of working memory capacity (WMC) and fluid intelligence (Gf). Paths connecting latent variables (WMC and Gf) indicate correlations between constructs, and the numbers from the latent variables to the manifest variables (seen in squares) represent the loadings of each task onto the latent variable. Error terms are omitted to simplify. OSpan = operation span; SymSpan = symmetry span; RotSpan = rotation span.

conducted with a university student sample. Consequently, the range of variability in general abilities is restricted and people with low Gf and low WMC were underrepresented. Further studies with more cognitively diverse samples are required to test generalizability of the results to other populations, such as those including noncollege students, young children, the elderly, or clinical populations.

In addition, these data may not generalize to Latin American Spanish speakers, given the well-known differences in the Spanish language between Peninsular Spanish and Latin American Spanish.

The present study also used the CSTs as a single tool to measure WMC, and despite its validity, broader perspectives on WMC measurement and different paradigms to measure it exist. Indeed, it is recommended to use different methods from diverse paradigms to obtain a better measurement of WMC (Wilhelm, Hildebrandt, & Oberauer, 2013).

In conclusion, the present study is an initial step toward providing a standardized method for assessing WMC in the Peninsular Spanish-speaking population that is based on a shortened format of the CSTs that affords the ability to obtain an accurate measurement of the construct using multiple tasks and a reduced test administration time. Although further studies are needed to test the generalizability of the results to other populations, this study contributes in furthering the understanding of the nature of WMC. The Spanish version is accessible at the Georgia Tech Attention and Working Memory Lab website (http://englelab.gatech.edu) for other researchers with the aim of promoting valid and standardized methods for assessing WMC and facilitating generalization of results as well as intercultural study replications.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131,* 30–60. http://dx.doi.org/10.1037/0033-2909.131.1.30

Bailey, H., Dunlosky, J., & Kane, M. J. (2011). Contribution of strategy use to performance on complex and simple span tasks. *Memory & Cognition, 39,* 447–461. http://dx.doi.org/10.3758/s13421-010-0034-3

Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice, 31,* 141–154. http://dx.doi.org/10.1037/0735-7028.31.2.141

Cohen, G., & Conway, M. A. (2008). *Memory in the real world* (3rd ed.). Hove, UK: Psychology Press Taylor & Francis Group.

Conway, A. R., Jarrold, C. E., Kane, M. J., Miyake, A., & Towse, J. N. (2007). *Variation in working memory*. New York, NY: Oxford University Press.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12,* 769–786. http://dx.doi.org/10.3758/BF03196772

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19,* 450–466. http://dx.doi.org/10.1016/S0022-5371(80)90312-6

De Neys, W., D'Ydewalle, G., Schaeken, W., & Vos, G. (2002). A Dutch, computerized, and group administrable adaptation of the operation span test. *Psychologica Belgica, 42,* 177–190. Retrieved form http://psycnet.apa.org/psycinfo/2003-02354-004

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 145–199). New York, NY: Elsevier.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128,* 309–331. http://dx.doi.org/10.1037/0096-3445.128.3.309

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition, 43,* 226–236. http://dx.doi.org/10.3758/s13421-014-0461-7

Gonthier, C., Thomassin, N., & Roulin, J. L. (2016). The composite complex span: French validation of a short working memory task. *Behavior Research Methods, 48,* 233–242. http://dx.doi.org/10.3758/s13428-015-0566-3

Hambleton, R. K. (2005). Issues, designs and technical guidelines for adapting test into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3–38). Mahwah, NJ: Erlbaum.

Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2005). The role of working memory in higher-level cognition: Domain-specific versus domain-general perspectives. In R. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence: Identifying the mechanisms of the mind* (pp. 104–121). New York, NY: Cambridge University Press.

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology, 45,* 119–129. http://dx.doi.org/10.1111/j.1469-8986.2007.00605.x

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. http://dx.doi.org/10.1080/10705519909540118

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science, 18,* 614–621. http://dx.doi.org/10.1111/j.1467-9280.2007.01948.x

Kane, M. J., Hambrick, D. Z., & Conway, A. R. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131,* 66–71. http://dx.doi.org/10.1037/0033-2909.131.1.66

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133,* 189–217. http://dx.doi.org/10.1037/0096-3445.133.2.189

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Lezak, M. (1983). *Neuropsychological assessment*. New York, NY: Oxford University Press.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods, 47,* 1343–1355. http://dx.doi.org/10.3758/s13428-014-0543-2

Pardo-Vázquez, J. L., & Fernández-Rey, J. (2008). External validation of the computerized, group administrable adaptation of the "operation span task." *Behavior Research Methods, 40,* 46–54. http://dx.doi.org/10.3758/BRM.40.1.46

Pardo-Vazquez, J. L., & Fernandez-Rey, J. (2012). Working memory capacity and mental rotation: Evidence for a domain-general view. *The Spanish Journal of Psychology, 15,* 881–890. http://dx.doi.org/10.5209/rev_SJOP.2012.v15.n3.39381

Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology, 20,* 33–65. http://dx.doi.org/10.1016/j.acn.2004.02.005

Raven, J. C., Raven, J., & Court, J. H. (1998). *A manual for Raven's progressive matrices and vocabulary scales*. London, UK: H. K. Lewis.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28,* 164–171. http://dx.doi.org/10.1027/1015-5759/a000123

Rosen, V. M., & Engle, R. W. (1997). Forward and backward serial recall. *Intelligence, 25,* 37–47. http://dx.doi.org/10.1016/S0160-2896(97)90006-4

Sanchez, C. A., Wiley, J., Miura, T. K., Colflesh, G. J., Ricks, T. R., Jensen, M. S., & Conway, A. R. (2010). Assessing working memory capacity in a non-native language. *Learning and Individual Differences, 20,* 488–493. http://dx.doi.org/10.1016/j.lindif.2010.04.001

Shelton, J. T., Elliott, E. M., Hill, B. D., Calamia, M. R., & Gouvier, W. D. (2009). A comparison of laboratory and clinical working memory tests and their prediction of fluid intelligence. *Intelligence, 37,* 283–293. http://dx.doi.org/10.1016/j.intell.2008.11.005

St Clair-Thompson, H. L. (2010). Backwards digit recall: A measure of short-term memory or working memory? *European Journal of Cognitive Psychology, 22,* 286–296. http://dx.doi.org/10.1080/09541440902771299

St Clair-Thompson, H. L., & Allen, R. J. (2013). Are forward and backward recall the same? A dual-task study of digit recall. *Memory & Cognition, 41,* 519–532. http://dx.doi.org/10.3758/s13421-012-0277-2

Thurstone, L. L. (1938). *Primary mental abilities (kursiva)*. Chicago, Il: University of Chicago Press.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127–154. http://dx.doi.org/10.1016/0749-596X(89)90040-5

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133,* 1038–1066. http://dx.doi.org/10.1037/0033-2909.133.6.1038

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498–505. http://dx.doi.org/10.3758/BF03192720

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory, 17,* 635–654. http://dx.doi.org/10.1080/09658210902998047

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology, 4,* 433. http://dx.doi.org/10.3389/fpsyg.2013.00433

Received July 18, 2016
Revision received March 7, 2017
Accepted March 17, 2017 ■