Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Fluid intelligence and working memory capacity: Is the time for working on intelligence problems relevant for explaining their large relationship?

Roberto Colom [a,*], Jesús Privado [b,c], Luis F. García [a], Eduardo Estrada [a], Lara Cuevas [b,c], Pei-Chun Shih [a]

[a] Universidad Autónoma de Madrid, Spain
[b] Universidad Complutense de Madrid, Spain
[c] Colegio Universitario Cardenal Cisneros, Madrid, Spain

ARTICLE INFO

ABSTRACT

A recent report has shown that the relationship, at the latent variable level, between fluid ability and working memory capacity is affected by the time allowed for completing problems requiring the former (Chuderski, 2013): the greater the time, the lower the relationship. The underlying argument is that untimed administration of fluid ability problems compensates working memory capacity limitations. The present report analyzes a group of three hundred and two participants that completed a set of three fluid tests and six working memory tasks. Latent variable analyses revealed consistent correlations (weighted average r = .86) between fluid ability and working memory capacity irrespective of administration times. Furthermore, the lowest difference in fluid ability between individuals with high and low working memory capacity was observed for the highly speeded condition. Their difference was greater when increased time was allowed for completing the fluid problems. Therefore, the relationship between fluid ability and working memory capacity appeals to underlying general common mechanisms unrelated with time constraints. Here we suggest that the reliability by which the relevant information can be preserved in the short-term for successful on-line processing seems a likely candidate.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Fluid ability (Gf) and working memory capacity (WMC) are strongly related at the construct level. It is important to underscore that constructs and measures are not the same thing. Constructs are estimated using measures, but the latter are not the former (Jensen, 1998). From this perspective, only latent-variable analyses can provide valuable results for uncovering the most likely relationship between Gf and WMC. There are studies supporting their almost isomorphic nature, but there is not unanimity (Ackerman, Beier, & Boyle, 2005, but see the re-analysis by Oberauer, Schulze, Wilhelm, & Süb, 2005). We have underscored elsewhere that constructs must be sampled appropriately, meaning that several varied measures are required for tapping the same latent factor (Martínez et al., 2011). When this is done, results do support the quasi-isomorphic nature of Gf and WMC (Colom, Abad, Rebollo, & Shih, 2005; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Kyllonen & Christal, 1990; Oberauer, Süß,

Wilhelm, & Sander, 2007; Oberauer, Süß, Wilhelm, & Wittman, 2008).

Recently, Chuderski (2013) has published a thought-provoking report suggesting that the large correlation between Gf and WMC can be explained by time-constraints when completing fluid problems. This study supported their isomorphism when highly speeded Gf tests were administered. Increasing the time for solving the fluid problems substantially degrades the relationship between Gf and WMC (changing from 1.0 in highly speeded Gf tests to .62 for virtually untimed Gf tests). It is suggested that fluid reasoning is iterative on untimed intelligence testing. In this regard, low working memory individuals are thought to compensate their capacity limitations in unspeeded conditions. Hence the lower correlation observed between fluid ability and working memory capacity when the former was measured without severe time restrictions. However, if the argument is likely, then low working memory individuals must show fluid scores closer to high working memory individuals with increased administration times. This issue will be tested in the present study.

In short, here we firstly analyze the correlation, at the latent variable level, between Gf and WMC using a diverse set of measures. We will measure WMC by six verbal, numerical, and spatial tasks, whereas Gf will be measured by three standardized tests.

* Corresponding author at: Facultad de Psicología, Universidad Autónoma de Madrid, 28049 Madrid, Spain. Tel.: +34 91 497 41 14.
E-mail address: roberto.colom@uam.es (R. Colom).

The Raven Advanced Progressive Matrices Test (RAPM) will be administered under three timed conditions (20, 30, and 40 min) whereas the remaining fluid measures will be administered following recommendations of the tests' manuals (see below). Following Chuderski's report, we predict that the relationship between Gf and WMC must decline monotonically from the 20 min condition. Secondly, we tests if high and low working memory capacity individuals show reduced differences at increased administration times of fluid ability problems. Again, following Chuderski's rational we predict that the largest difference between these individuals varying in their working memory capacity must be observed in highly speeded conditions.

2. Methods

2.1. Participants

Three hundred and two university students participated in this study to fulfill a course requirement. Seventy-seven percent were females and the mean age was 19 years (SD = 3.6). Participants were randomly assigned to three administration conditions regarding the RAPM. Ninety-three were submitted to the 20 min condition (high-speed), 99 to the 30 min condition (moderate speed), and one hundred and ten to the 40 min condition (low speed).

2.2. Measures

Fluid ability was measured by the RAPM (Raven, Raven, & Court, 2004), the reasoning tests from the Differential Aptitude Test Battery (DAT-AR) (Bennett, Seashore, & Wesman, 1990), and the dominoes test (D-48) (Pichot, 1961). Verbal working memory was measured by the ABCD and Alphabet tasks. Both tasks were modeled after Kyllonen and Christal (1990). The mental counters and computation span tasks measured numerical working memory. The counter task was modeled after Larson and Sacuzzo (1989). The computation span task was modeled after Ackerman, Beier, and Boyle (2002). Finally, spatial working memory was measured by the dot matrix and letter rotation tasks. Both tasks were modeled after Miyake, Friedman, Rettinger, Shah, and Hegarty (2001). A detailed description can be found in Appendix 1.

2.3. Procedure

Participants completed the intelligence measures in two separate testing sessions in groups of no more than 25 individuals.

The first session was devoted to the RAPM, whereas the second session included the DAT-AR and the D-48. The cognitive tasks were also completed in two sessions. In the first session the ABCD, the computation span, and the letter rotation task were administered. In the second session the remaining WM tasks were completed.

3. Results

Table 1 shows the descriptive statistics for the nine measures and the three RAPM groups. Results were largely similar regardless of the group, except for the RAPM. In this latter instance, increased administration time produced higher mean scores and these were the computed effect sizes: d (20 min group vs. 30 min group) = .39, d (30 min group vs. 40 min group) = .58, d (20 min group vs. 40 min group) = 1.01. Therefore, the 30 min group obtained an advantage equivalent to 6 IQ points over the 20 min group, the 40 min group obtained an advantage equivalent to 9 IQ points over the 30 min group, and the 40 min group obtained an advantage equivalent to 15 IQ points over the 20 min group. Skewness and kurtosis values were within the normal range. Reliabilities were also appropriate. Note that the reliability for the RAPM was almost identical for the three administration times (.77 for 20 min, .76 for 30 min, and .75 for 40 min).

Table 2 depicts the correlation matrix among the administered measures, again by RAPM group. Importantly, the correlation between the RAPM and the remaining two fluid measures were largely similar. This suggests that administration times do not change the nature of what is measured by the RAPM.

Afterwards, we defined the general latent-variable model including all the fluid ability and working memory capacity measures. The AMOS program (Arbuckle, 2006) was used for the computations (using maximum-likelihood estimation) testing the similarity among results for the three RAPM groups. Multivariate normality was confirmed using the Bollen-Stine Bootstrap method (p = .164). Model fitting was assessed using the Mean Square Error of Approximation (RMSEA) (Steiger, 1990), the χ^2/df ratio, the Tucker Lewis index (TLI) (Hu & Bentler, 1999), and the Comparative Fit Index (CFI) (Bentler, 1990). A χ^2/df ratio < 3.00, RMSEA values < .05, as well as TLI and CFI values > .95, are indicative of proper fit.

Figure 1 depicts the latent results for the three RAPM groups. The remarkable general finding is that there were large and consistent correlations between fluid ability (Gf) and working memory capacity (WMC) across groups. For the 20 min group the correlation was .89 (confidence interval = .84, .92), for the 30 min group

Table 1
Descriptive statistics for the three groups (20, 30, and 40 min).

	RAPM	DAT-AR	D48	ABCD	Alphabet	CompSpan	MentCount	LetterR	DotMatrix
20 min (N = 93)									
Mean	20.2	26.2	32.3	9.2	62.3	42.3	46.3	42.6	81.6
SD	4.4	6.3	6.5	3.9	15.6	9.3	8.7	9.7	9.9
Skew	−.05	−.51	−1.4	−.66	−.60	−.97	−1.3	.09	−.73
Kurt	−.26	.16	2.8	−.61	.50	.34	2.1	−.18	2.3
30 min (N = 99)									
Mean	22.0	25.7	31.4	9.4	59.2	42.5	46.1	40.9	80.0
SD	4.8	6.6	6.1	3.8	14.4	9.9	8.7	10.1	8.6
Skew	−.59	−.40	−.57	−.61	−.31	−1.4	−1.6	.32	.01
Kurt	.62	.34	.28	−.60	.35	2.0	3.6	−.44	−.49
40 min (N = 110)									
Mean	24.7	25.7	33.0	9.5	62.7	43.1	46.4	43.3	80.7
SD	4.5	6.2	6.2	3.7	13.7	8.4	7.1	9.9	9.1
Skew	−.29	−.07	−1.2	−.71	−.46	−1.1	−.68	.24	−.50
Kurt	.03	−.41	2.8	−.48	.73	1.5	.15	−.40	.86
Reliability	.77/.76/.75	.83	.85	.85	.78	.84	.79	.79	.83

Table 2
Correlation matrices for the three groups (20 min/30 min/40 min). N (20 min) = 93, N (30 min) = 99, N (40 min) = 110.

|	RAPM	DAT-AR	D48	ABCD	Alphabet	CompSpan	MentCoun	LetterR	DotMatrix
RAPM	–	.57/.66/.60	.48/.52/.60	.16/.31/.30	.21/.21/.47	.22/.17/.29	.50/.30/.45	.35/.22/.47	.43/.55/.45
DAT-AR		–	.59/.58/.50	.30/.38/.30	.30/.20/.46	.35/.10/.38	.54/.37/.57	.48/.33/.56	.57/.48/.55
D48			–	.29/.34/.35	.32/.39/.54	.25/.22/.30	.55/.28/.51	.41/.38/.41	.42/.31/.41
ABCD				–	.24/.31/.32	.33/.38/.25	.28/.30/.22	.29/.34/.16	.23/.35/.22
Alphabet					–	.29/.22/.26	.35/.21/.32	.34/.41/.46	.32/.40/.47
CompSpan						–	.32/.30/.33	.40/.35/.31	.40/.23/.25
MentCoun							–	.42/.19/.46	.49/.31/.53
LetterR								–	.48/.35/.57
DotMatrix									–
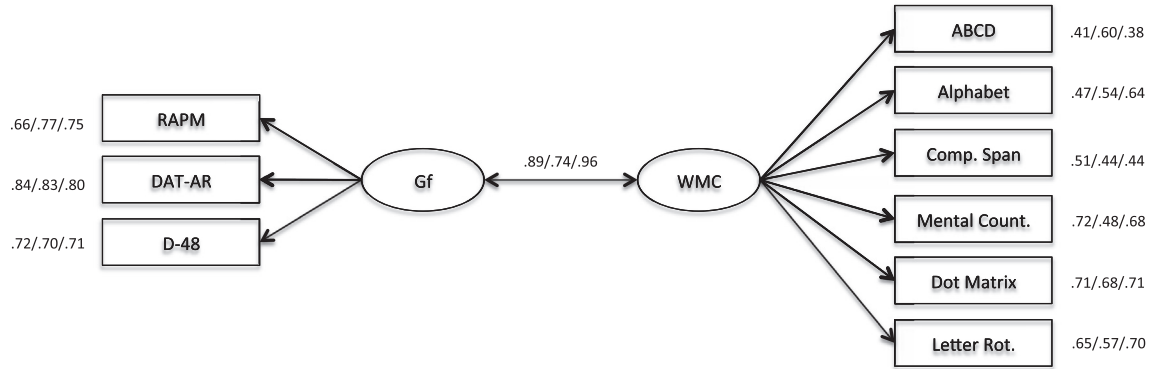
Fig. 1. Latent model for the three RAPM groups. Regression weights for the fluid ability (Gf) measures are depicted on the left (20 min/30 min/40 min), whereas regression weights for the working memory measures (WMC) are depicted on the right (20 min/30 min/40 min). Correlation values between Gf and WMC are also presented in the same order (20 min/30 min/40 min).

Table 3
Fitted models and results for model comparisons.

Models	χ^2	gl	χ^2/gl	TLI	CFI	RMSEA	AIC
Model A. Unconstrained	101.29	78	1.299	.960	.971	.032	215.29
Model B. Structural weights	112.18	92	1.219	.971	.975	.027	198.18
Model C. Structural covariances	121.01	98	1.235	.969	.972	.028	195.01
Model D. Measurement residuals	148.71	116	1.282	.963	.960	.031	186.71
Model comparison	$\Delta\chi^2$	Δgl	p		ΔCFI		
Modelo A & B (metric invariance)	10.89	14	>.500		.004		
Models B & C (strong metric invariance)	8.83	6	>.150		.003		
Models C & D (strict metric invariance)	27.70	18	>.050		.012		

it was .74 (confidence interval = .63, .82), and for the 40 min group the correlation was .96 (confidence interval = .94, .97). Therefore, the relationship between these constructs was not monotonically reduced at increased RAPM administration times. The weighted average correlation was .86 (confidence interval = .83, .89).

For testing the invariance of the factor structure across RAPM groups, we applied multi-group confirmatory factor analysis (MG-CFA) (Vanderberg & Lance, 2000). This was the analytic sequence: first, in the unconstrained model (Model A), factor loadings, the variance–covariance matrix and error variances were allowed to differ across groups; second, for the metric invariance model (Model B), factor loadings were constrained to be equal; third, in the strong invariance model (Model C), the factor variance–covariance matrix was also constrained to be equal; finally, for the strict invariance model (Model D), error variances were also constrained to be equal.

These models were compared using these indices: the difference between χ^2 associated with two alternative models, RMSEA,

CFI, and Akaike (1987) information criterion (AIC). The smaller the AIC value, the better the model fit. Furthermore, a difference in CFI values < .01 is usually taken as supporting an invariant solution (Cheung & Rensvold, 2002). Table 3 shows the results.

As revealed by $\Delta\chi^2$ results supported metric invariance (Model A vs. Model B), strong metric invariance (Model B vs. Model C), and strict metric invariance (Model C vs. Model D). Also, ΔCFI values were consistent with the presence of structural invariance (Cheung & Rensvold, 2002), and, therefore, factor loadings and variance–covariance matrices were identical for the three RAPM groups.

We checked if correlations between Gf and WMC were significantly different from 1. This was done computing if the χ^2 difference is statistically significant (Yung, Thissen, & McLeod, 1999) and these were the results: $\Delta\chi^2 = \chi^2 (87) - \chi^2 (78) = 230.00 - 101.29 = 128.71$, $p < .001$. Therefore, correlations cannot be fixed to 1. Given the sensitivity of the χ^2 statistic to sample size, computing ΔTLI is also recommended. A difference greater than .01 is

considered statistically significant (Gignac, 2007). ΔTLI = .960 − .782 = .178. Thus, the difference was significant, and, again, we conclude the correlation cannot be fixed to 1.

Finally, we checked how low and high working memory individuals scored on the RAPM across administration times. A general index (Mean = 100, SD = 15) was calculated after the six working memory tasks. High working memory individuals obtained scores greater than 115, whereas low working memory individuals obtained scores lower than 85. Results showed that high working memory individuals obtained better RAPM scores than low working memory individuals regardless of administration time (Fig. 2). Indeed, high working memory individuals completing the RAPM in 20 min obtained a closely similar score (20.9) to the low working memory individuals completing the RAPM in the 40 min condition (19.6). But the key finding was that the lowest difference (4 items on average) between these working memory groups was observed for the highly speeded condition (20 min). The difference doubled in the 30 min condition (8 items on average) and remained at the same level in the 40 min condition. This result is hardly supportive of Chuderski's conceptual perspective.

4. Discussion

Here we have shown that the strong correlation between fluid ability and working memory capacity at the latent variable level is virtually unaffected by the time allowed for completing the Raven Progressive Matrices Test (RAPM). The observed correlation values were not distinguishable statistically across the three considered conditions and the average value was .86. Nevertheless, this correlation cannot be fixed to unity, and, therefore, some degree of uniqueness in fluid ability and working memory capacity not captured by their shared variance must be acknowledged, at least in this dataset.

The main message that might be derived from the present study is that the time allowed for working on fluid problems is not behind the high correlation observed between fluid ability and working memory capacity. This result is in tension with the main conclusion supported by the Chuderski's (2013) study. Importantly, our results showed that the correlation between the RAPM administered under quite different timing conditions and the remaining fluid measures remains unaltered. This allowed the

delimitation of a fluid latent factor that was correlated with the working memory factor defined by six diverse verbal, numerical, and spatial tasks.

The observed strong correlation between the considered latent constructs (the shared variance was 74%) is consistent with previous findings (Colom, Abad, Rebollo, et al., 2005; Colom, Rebollo, Palacios, et al., 2004; Martínez et al., 2011; Oberauer et al., 2007) but not with all the published results. Thus, for instance, Redick, Unsworth, Kelly, and Engle (2012) reported a correlation of .53 between Gf and WMC. Engle, Tuholski, Laughlin, and Conway (1999) found a correlation of .59 between these constructs. Kane et al. (2004) reported a correlation of .64. As noted by Chuderski (2013) comparisons across studies with respect to administration time of the completed fluid tests are not straightforward.

We have shown that fluid ability measured under non-standardized (high speeded) time constraints and this same ability measured under standardized time constraints tap the same construct. The construct validity is not threatened. The argument endorsed by Chuderski (2013) along with his interesting thoughts regarding the role of relational learning seems unwarranted. He reasoned that fluid reasoning could be iterative on untimed fluid intelligence testing and this implies that low working memory individuals would compensate their capacity limitations in unspeeded conditions. Closer scores are expected in these latter conditions, but we did find just the opposite pattern: low speeded conditions revealed the largest difference between high and low working memory individuals in our dataset.

In closing, there should be some general mechanism underlying the strong correlation between fluid ability and working memory capacity. Here we have shown that this correlation is virtually unaffected by the time allowed for working on the fluid problems. In addition, sharp differences in working memory capacity are not more visible in highly speeded conditions than in low speeded conditions for completing fluid ability problems. Elsewhere we have suggested that the common general mechanism might be related with the reliability by which the information required for on-line processing can be temporarily preserved (Colom, Rebollo, Abad, & Shih, 2006; Martínez et al., 2011). This kind of processing is intimately related with short-term memory, and this cognitive function is hardly distinguished from working memory capacity (Colom, Shih, Flores-Mendoza, & Quiroga, 2006; Unsworth & Engle, 2007). Speed factors are much less relevant for
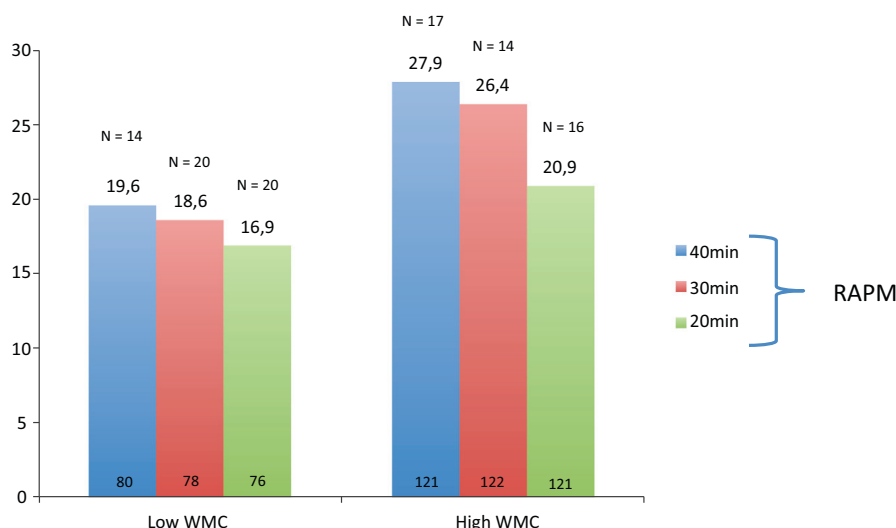

Fig. 2. RAPM scores of low (low WMC) and high (high WMC) working memory individuals in the three RAPM administration times. Numbers within bars denote average working memory scores. N = number of participants within each condition.

understanding the construct of fluid ability (Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008), and its assessment in applied settings, which is largely consistent with the three-stratum theory (Carroll, 1993).

Appendix A

Fluid ability measures

The RAPM comprises a matrix figure with three rows and three columns. There are eight possible answers but only one fits the 3×3 matrix. The 36 matrices increase progressively their complexity.

DAT-AR is a series test based on abstract figures. Successive figures follow a rule and the figure fitting the series must be chosen among five possible answers. The test includes 40 items.

D-48 is a dominoes test based on finding rules within series varying in their format. Each domino piece shows variations in the number of dots (from 0 to 6) in the top and bottom area. There is always an empty piece and the participant must find the answer properly completing the series. The test includes 44 items.

The number of hits was the dependent variable in these three tests.

Working memory measures

ABCD. Two categories were used, with five words comprised by each category. The categories were trees and food. Three study frames were displayed on the computer screen for 3 s each. The first frame indicated the order of two members from the same category (El cedro precede al roble); the second frame indicated the order of two members from the other category (El ajo no precede a la sal); the third frame indicated the order of the categories (Los árboles no preceden a la comida). After the third study screen, an eight choice answer screen was displayed from which participants selected the correct order of the words. Participants were allowed 10 s to select a response. The use and ordering of category members were balanced across items, as were the variations of order (precede, no precede). There were 14 trials and two practice trials. The number of hits was obtained as the participant's score.

Alphabet. This task required participants to apply successor and predecessor operations to a string with a given number of letters. A typical trial was

(Screen 1) M A D
(Screen 2) + 2
(Screen 3) ¿

for which the correct response would be: P D G.

The string of letters is presented for 3 s, the operation to apply is presented for 1500 ms, and the participant has unlimited time to enter a response. The number of letters increased from three to seven (four trials at each level and a total of 20 trials). For two trials within a given block the participant must add 1 or 2 positions, while for the other two trials the participant must subtract 1 or 2 positions. The number of additions and subtractions are randomized within a given block of trials. The number of hits is obtained as the participant's score.

Mental Counters. Three boxes representing counters appeared on the computer screen. At the beginning of each trial, the value of the three counters was set to 0, 0, 0. A flash appeared above or below one counter for 500 ms. If the flash appeared above the box, the participant must add one (+1) to that counter, but if the flash appeared below the box, the participant must subtract one (−1) to that counter. The participant's task is to keep a running track of the value of the three counters. At the end of each trial, participants reported the cumulative total of all three counters.

The participant had unlimited time to enter a response. The task comprised 10 trials with five counter changes and 10 trials with seven counter changes (and three practice trials with three counter changes). Maximum and minimum values used were +3 and −3. The number of hits was obtained as the participant score.

Computation span. This task included a verification task and a recall task. Participants were allowed 6 s to verify the accuracy of a math equation and were instructed to remember the displayed solution, irrespective of its accuracy. After the final equation of the trial was displayed, participants were prompted to remember in the correct serial order each of the presented solutions from the equations. Each math equation included two operations using digits from 1 to 10, and the provided and actual solutions were always single-digit numbers. Set size ranged from three to seven equation/solutions (15 trials total). The participant score was obtained after the number of hits in the verification and remembering tasks.

Dot Matrix. The requirement is to verify a matrix equation while simultaneously remembering a dot location in a 5×5 grid. In the matrix equation display, a simple addition or subtraction equation was presented on the computer screen. Participants were given a maximum of 4.5 s to verify each equation. Immediately after their response the computer displayed a dot grid for 1.5 s. After a sequence of between two and five equation-grid pairs, participants recalled, in any order, which grid spaces contained dots clicking with the mouse on an empty grid. There were three practice trials with two equations and two dots each, after which sets progressively increased in size from two to five equations and dots for a total of 12 sets, 3 of each size. The number of hits in the verification and remembering tasks were obtained as the participant's score.

Letter Rotation. Each trial consisted of the sequential presentation of a set of the same capital letter (F, J, P, or R) each of which appeared on a computer screen either normal or mirror imaged and rotated in one of seven possible orientations (multiples of 45°). The participant's task was to decide whether each letter was normal or mirror imaged as quickly and as accurately as possible and remember its spatial orientation (where the top of each letter was pointing). Participants were given a maximum of 3 s to respond normal or mirror, and then the subsequent letter is displayed. After each set, the participant saw a grid and marked the places corresponding to the positions of the tops of the presented letters in the correct serial order. There were three practice trials with two letters each, after which sets progressively increased in size from two to five letters for a total of 12 sets, 3 of each size. The number of hits in the verification and remembering tasks were obtained as the participant's score.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General, 131*, 567–589.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30–60.

Akaike, H. (1987). Factor Analysis and AIC. *Psychometrika, 52*, 317–332.

Arbuckle, J. L. (2006). *Amos 7.0 User's Guide*. Chicago: SPSS.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1990). *Differential Aptitude Test* (5th ed.). Madrid: TEA.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246.

Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge: UK, Cambridge University Pres.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Colom, R., Abad, F. J., Quiroga, M. A., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs but why? *Intelligence, 36*, 584–606.

Colom, R., Abad, F. J., Rebollo, I., & Shih, P. C. (2005). Memory span and general intelligence: A latent variable approach. *Intelligence, 33*, 623–642.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence, 32*, 277–296.

Colom, R., Rebollo, I., Abad, F., & Shih, P. C. (2006). Simple span tasks, complex span tasks. And cognitive abilities: A re-analysis of key studies. *Memory and Cognition, 34*, 158–171.

Colom, R., Shih, P. C., Flores-Mendoza, C., & Quiroga, M. A. (2006). The real relationship between short-term memory and working memory. *Memory, 14*, 804–813.

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence, 41*, 244–262.

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

Gignac, G. E. (2007). Multi-factor modelling in individual differences research: Some suggestions and recommendations. *Personality and Individual Differences, 42*, 37–48.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.

Jensen, A. R. (1998). *The g factor*. CT, Praeger: Westport.

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 visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189–217.

Kyllonen, P., & Christal, R. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence, 14*, 389–433.

Larson, G. E., & Sacuzzo, D. P. (1989). Cognitive correlated of general intelligence: Toward a process theory of g. *Intelligence, 13*, 5–31.

Martínez, K., Burgaleta, M., Román, F. J., Escorial, S., Shih, P. C., Quiroga, Ma., et al. (2011). Can fluid intelligence be reduced to 'simple' short-term storage? *Intelligence, 39*, 473–480.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General, 130*(4), 621–640.

Oberauer, K., Schulze, R., Wilhelm, O., & Süb, H. (2005). Working memory and intelligence— their correlation and their relation: comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 61–65.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 49–75). Oxford: Oxford University Press.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittman, W. W. (2008). Which working memory functions predict intelligence? *Intelligence, 36*, 641–652.

Pichot, P. (1961). *D48 (Dominoes Test)*. Madrid: TEA.

Raven, J., Raven, J. C., & Court, J. H. (2004). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Redick, T. S., Unsworth, N., Kelly, A. J., & Engle, R. W. (2012). Faster, smarter? Working memory capacity and perceptual speed in relation to fluid intelligence. *Journal of Cognitive Psychology*. http://dx.doi.org/10.1080/20445911.2012.704359.

Steiger, J. H. (1990). Structural model evaluation modification: An interval estimation approach. *Multivariate Behavioral Research, 25*, 173–180.

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

Vanderberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organization research. *Organizational Research Methods, 3*, 4–70.

Yung, Y. F., Thissen, D., & McLeod, L. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika, 64*, 113–128.