



Age-related differentiation in verbal and visuospatial working memory processing in childhood

Frances Buttelmann^{1,2} · Tanja Könen^{2,3} · Lauren V. Hadley⁴ · Julie-Anne Meaney⁵ · Bonnie Auyeung⁵ · Candice C. Morey⁶ · Nicolas Chevalier⁵ · Julia Karbach^{2,3}

Received: 26 September 2018 / Accepted: 20 June 2019 / Published online: 12 July 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Working memory (WM), a key feature of the cognitive system, allows for maintaining and processing information simultaneously and in a controlled manner. WM processing continuously develops across childhood, with significant increases both in verbal and visuospatial WM. Verbal and visuospatial WM may show different developmental trajectories, as verbal (but not visuospatial) WM relies on internal verbal rehearsal, which is less developed in younger children. We examined complex VWM and VSWM performance in 125 younger (age 4–6 years) and 101 older (age 8–10 years) children. Latent multi-group modeling showed that (1) older children performed better on both verbal and visuospatial WM span tasks than younger children, (2) both age groups performed better on verbal than visuospatial WM, and (3) a model with two factors representing verbal and visuospatial WM fit the data better than a one-factor model. Importantly, the correlation between the two factors was significantly higher in younger than in older children, suggesting an age-related differentiation of verbal and spatial WM processing in middle childhood. Age-related differentiation is an important characteristic of cognitive functioning and thus the findings contribute to our general understanding of WM processing.

Introduction

Working memory (WM) is a key feature of our cognitive system, allowing for the simultaneous maintenance and processing of information in a controlled manner (Baddeley and Hitch 1994). WM is an excellent predictor for academic

development in childhood, as it helps children to process complex tasks, update and maintain task goals, and remember and follow instructions (Alloway et al. 2009). Consistently, WM, which improves most during the preschool period, but continues to develop well into adolescence (Gathercole et al. 2004), supports learning across academic domains, including reading and mathematics (Alloway et al. 2004, 2005; Cowan 2014; Titz and Karbach 2014, for reviews). Several WM models are currently existing and discussed (e.g., Baddeley 2000; Barrouillet and Camos 2012; Oberauer 2009). They all contributed substantially to our understanding of WM and agree on the basic assumptions that WM capacity is limited and that reliable individual differences in this capacity exist (e.g., Oberauer 2009). They do, however, differ in the assumptions about the components of WM. For example, there is debate of whether verbal and visuospatial information are processed and stored in different components (e.g., Alloway et al. 2006). One way to investigate this is by examining the developmental trajectory of verbal and visuospatial WM. Among the most commonly used tasks to assess WM abilities in childhood are complex WM span tasks. In WM span tasks, participants are presented a sequence of stimuli (encoding phase) and are asked to repeat them in the correct or reversed order after a

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00426-019-01219-w>) contains supplementary material, which is available to authorized users.

✉ Frances Buttelmann
frances.buttelmann@uni-jena.de

- ¹ Department of Developmental Psychology, Friedrich Schiller University Jena, Jena, Germany
- ² Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt, Germany
- ³ Department of Psychology, University of Koblenz-Landau, Landau, Germany
- ⁴ Hearing Sciences - Scottish Section, University of Nottingham, Glasgow, UK
- ⁵ Department of Psychology, University of Edinburgh, Edinburgh, UK
- ⁶ School of Psychology, Cardiff University, Cardiff, UK

short delay (recall phase). Complex WM span tasks increase demands on executive functioning by adding a processing task to the encoding phase, which requires attention and thus prevents it from contributing to storage during processing episodes. For instance, participants are asked to decide for every stimulus they see whether it is large or small while encoding the sequence. In large cross-sectional data sets, research repeatedly demonstrated substantial age-related WM-performance improvements across childhood on various WM span tasks (e.g., Michalczyk et al. 2013).

While different components of WM are clearly separable in adults (e.g., Jarrold and Towse 2006; Vallar and Papagno 2002), the extent to which these components differentiate during childhood remains to be clarified. Findings related to Baddeley's model demonstrated that the phonological loop and the visual-spatial sketchpad seem to be at least partially separable in both 4–8 years (Pickering et al. 1998; Roebbers and Zoelch 2005) and 11–14 years (Jarvis and Gathercole 2003). However, this differentiation is still not as well-established as the one between the two slave systems and the central executive (Alloway et al. 2006; Gathercole et al. 2004), indicating that the phonological loop and visuospatial sketchpad may continue differentiating across middle childhood. Consistently, performance on tasks requiring the phonological loop or the visuospatial sketchpad correlates with different outcomes, i.e., performance on verbal span tasks significantly correlates with measures of attainment levels in linguistic domains (e.g., English), whereas performance on spatial span tasks significantly correlates with mathematics and science measures (Jarvis and Gathercole 2003).

Therefore, our first aim was to clarify whether verbal and visuospatial working memory processing differentiates during childhood by comparing performance on complex verbal WM (VWM) and visuospatial WM (VSWM) span tasks in 4–6 years and 8–10 years.

Although research has repeatedly demonstrated substantial age-related WM performance improvements across childhood on various WM span tasks (e.g., Michalczyk et al. 2013), it is unclear whether children's performance on complex span tasks is influenced by the verbal or visuospatial nature to be recalled information. Young children may perform better on VSWM compared to VWM complex span tasks, because they have had less practice on verbal rehearsal strategies. For instance, Flavell, Beach, and Chin-sky (1966) tested 5, 7, and 10 years on a serial recall task and results show that the use of rehearsal strategies increased with age and improved performance of item recall. Whereas younger children only passively used rehearsal and mostly just repeated a single item, older children and adults actively used rehearsal strategies, e.g., repeating a series of items (e.g., Cuvo 1975; Hitch et al. 1993; Ornstein et al. 1977).

Further evidence for less practice in rehearsal with age comes from studies with adults, showing that articulatory suppression during item presentation significantly reduced recall performance, leading to recall patterns similar to those found in 5 years (Cowan et al. 1987). Direct evidence from studies with children showed that long and similar item names resulted in decreased rehearsal accuracy in 11 years (only based on VWM span tasks; Hitch et al. 1991). Once children get older and learn to effectively use rehearsal strategies, performance differences between VWM and VSWM may disappear. However, recent evidence based on directly comparable tasks tapping both VWM and VSWM within the same sample is missing. Thus, the present study addressed this question by directly comparing children's performance on a complex VWM span task and a structurally comparable complex VSWM span task.

In the current study, we investigated two questions. (1) Do verbal and visuospatial WM progressively differentiate with age? (2) Do younger children show asymmetrical performance between the verbal and visuospatial domains that no longer exists in older children? To investigate these questions, we examined children in two age groups (younger: 4–6 years and older: 8–10 years) that performed two complex WM span tasks, one verbal and one visuospatial task. If VWM and VSWM progressively differentiate with age, we expect (1) a stronger correlation between two latent factors representing VWM and VSWM in the younger compared to the older children. Moreover, if younger children are less experienced using verbal rehearsal than visuospatial WM processing strategies, we expect (2) that they should perform better on the VSWM span task than on the conceptually comparable VWM span task and that this difference should be reduced or non-significant in older children.

Method

Participants

The sample comprised a total of 226 children, 125 children aged 4–6 years ($M_{Age} = 59.7$ months, SE 0.59; 70 boys) and 101 children aged 8–10 years ($M_{Age} = 106.4$ months, SE 0.64; 59 boys). Eleven additional children were excluded from the analyses, because they did not finish both tasks (4), were diagnosed with ADHD, ASD, or epilepsy (6) or because of technical errors (1). Half of the data in both age groups were collected in Germany and the other half in Scotland. The children were either tested in the laboratory or in a quiet room in their kindergarten or daycare. Participants were randomly assigned to start with one of the two

adaptive complex working memory span tasks (VWM and VSWM). Informed parental consent was obtained and children received a certificate for participation.¹

Materials, design and procedures

The tasks were presented on laptop computers with the software package Eprime 2.0 (<http://www.pstnet.com/products/Eprime/>; Schneider et al. 2002). The test session started with a short game to familiarize the children with the use of the computer mouse. Afterwards, they were introduced to the WM tasks.

Adaptive complex verbal working memory (VWM) span task: In the encoding phase, children heard a series of words corresponding to food items while watching a picture of that item on the screen (e.g., bread, milk...). After each word, children indicated for each picture if it was large or small by clicking on the corresponding response buttons in the lower left and right corners of the screen. At the end of the series, they saw an array containing eight food items, including those that had just been presented, and were asked to reproduce the sequence of items in the correct order by clicking on the corresponding pictures. After each click, a grey dot was presented in the selected location for 500 ms to indicate that the click had been registered. Children performed four practice trials followed by 12 test trials. To ensure comparability between countries, we used one-syllable words in English as well as in German as stimuli. Furthermore, we considered the age of acquisition of the words (Kuperman et al. 2012) to ensure that all children were familiar with the stimuli (e.g., bread, milk, salt, egg, tea, juice, nuts, and rice).

Adaptive complex visuospatial working memory (VSWM) span task: the task procedure was identical to the verbal WM span task, except for the following differences. Children saw an apple presented at different locations in a 3 × 3 grid. For each apple, they were asked to indicate if the apple was complete or if bites were taken from it by clicking on the corresponding response button in lower left and right corners of the screen. At the end of the series, children were asked to reproduce the sequence of locations by clicking on the corresponding grid.

To control for between-group differences, we also assessed children's verbal abilities and processing speed. Analysis of variance (ANOVA) showed that there neither were differences in terms of language abilities and processing speed between the children tested in Germany and Scotland nor between those performing the VWM or the VSWM task first. However, older children outperformed younger ones on both tasks (see Supplementary Material for details

on the tasks, descriptive statistics, reliabilities, and ANOVA results).

Task difficulty adjustment and scoring

Task difficulty (i.e., number of items to be remembered) was individually adapted as a function of each child's performance, starting with one item in the first trial and continuing for a total of 12 trials (see Studer-Luethi et al. 2015, for a similar procedure). If children responded correctly to both the WM span task and the processing task, task difficulty was increased by one item in the following trial. If they responded incorrectly to both tasks or to the span task only, task difficulty was decreased by one item in the next trial. If they responded correctly to the span task, but incorrectly to the processing task, task difficulty did not change. We calculated two scores for each child and each WM span task (VWM and VSWM) representing average (average trial length) and maximum (maximum correct trial length) performance.

Statistical modeling procedures

All models are confirmatory factor models (CFA) based on maximum likelihood estimation with robust standard error correction (MLR; *Mplus* 7.4). For model identification, the first factor loading of each latent variable was fixed to one. Model fit was evaluated with the χ^2 test, the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR) in reference to Beauducel and Wittmann (2005). We tested measurement invariance by comparing different levels of invariance (configural, metric, scalar, and strict, e.g., Cheung and Rensvold 2002). None of the variables had missing cases.

Results

Descriptive statistics and reliabilities are shown in Table S1 and correlations of all measures are shown in Table S2. We tested the measurement invariance of VWM and VSWM performance separately for age group (age 4–6 vs. 8–10), country (Germany vs. Scotland), and sex (female vs. male). In each case, invariance was evaluated in a two-group CFA with two factors (VWM and VSWM) having two indicators each (average and maximum performance; cf. Fig. 1 for an example). For all cases, the fit indices of all invariance levels are included in Table 1. Overall, all models fitted the data perfectly (all χ^2 with $ps > .11$) and eight out of nine χ^2 -difference tests were non-significant (indicating that the assumption of a higher level of measurement invariance did not result in a significant drop of model fit). The exception

¹ Ethical approval was given by the research ethics committee of the University of Edinburgh, protocol number 111-1516/2.

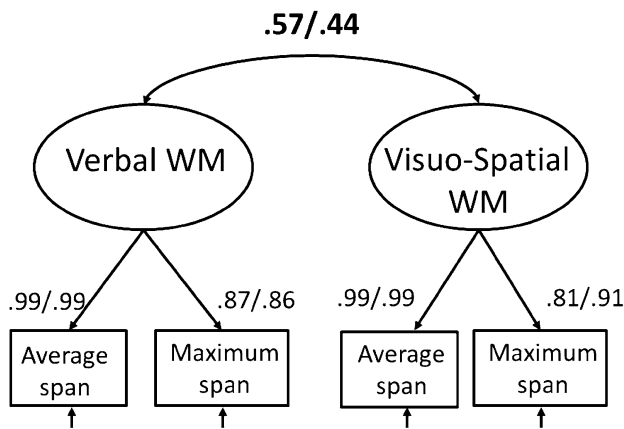


Fig. 1 Multi-group model with two age groups (younger/older children) and strict measurement invariance across age groups. Read from top to bottom, the digits represent latent factor correlations and standardized factor loadings, respectively, for both age groups 4–6/8–10 years

was metric measurement invariance for age group, which resulted in a significantly higher χ^2 , but not in a severe drop in the CFI ($\Delta\text{CFI} = .003$), indicating that there is a small but not severe violation of measurement invariance (e.g., Cheung and Rensvold 2002). Thus, overall, the data supported the assumption of measurement invariance over age group, country, and sex (Table 1), allowing for comparisons of means and variances between these groups (e.g., Steenkamp and Baumgartner 1998).

In a CFA model comparison, we found that two factors for WM performance (VWM and VSWM) explained the data better than one factor (WM) for both the younger and older children. Two WM factors fitted the data well in both the younger [$\chi^2(df=2) = 0.49, p = .79; \text{CFI} = 1; \text{RMSEA} = 0$ (90% CI = 0–.12); $\text{SRMR} < .01$] and older age group [$\chi^2(df=2) = 0.15, p = .93; \text{CFI} = 1; \text{RMSEA} = 0$ (90% CI = 0–.06); $\text{SRMR} < .01$]. Testing their correlation against 1, however, resulted in a severe drop of model fit in both age groups [younger: $\Delta\chi^2(df=1) = 83.06, p < .01; \Delta\text{CFI} = .28$; older: $\Delta\chi^2(df=1) = 157.99, p < .01; \Delta\text{CFI} = .41$], indicating that two factors represent the data better than one. This is not surprising when considering the strength of the observed correlations between the factors of .57 in the younger and .44 in the older age group. Using a two-group CFA (Fig. 1) with strict measurement invariance across age groups and perfect model fit (Table 1, line 4), we found that both correlations were significantly different from each other [$\Delta\chi^2(df=1) = 9.60, p < .01; \Delta\text{CFI} = .01$], indicating that the relation between VWM and VSWM was higher in younger children than in older children.

Using this two-group CFA (Fig. 1) with strict measurement invariance across age groups, we compared latent factor means (see Table S3). Performance on both WM factors was significantly higher for older children with very large effect sizes for both factors. Furthermore, performance on VWM was significantly better than on VSWM for both age groups. This effect was large in the younger and medium-sized in the older age group (see Table S3).

Table 1 Tests of measurement invariance for verbal and visuospatial working memory performance in children

Comparison	Invariance Level	$\chi^2 (df)$	CFI	RMSEA (90% CI)	SRMR	SB-scaled $\Delta\chi^2 (df)$	ΔCFI
Age group (age 4–6 vs. age 8–10)	1. Configural	0.61 (6), $p = .99$	1.00	0 (0–0)	.004		
	2. Metric	10.17 (8), $p = .25$.997	.05 (0–.13)	.112	^{1,2} 8.50(2), $p = .01$.003
	3. Scalar	10.11 (10), $p = .43$	1.00	.01 (0–.10)	.109	^{2,3} 0.12(2), $p = .94$	–.003
	4. Strict	12.08 (12), $p = .44$	1.00	.01 (0–.10)	.102	^{3,4} 1.96(2), $p = .38$	0
Country (Germany vs. Scotland)	1. Configural	5.60 (6), $p = .47$	1.00	0 (0–.12)	.010		
	2. Metric	9.76 (8), $p = .28$.999	.04 (0–.12)	.033	^{1,2} 4.28(2), $p = .12$.001
	3. Scalar	12.13 (10), $p = .28$.998	.04 (0–.12)	.039	^{2,3} 2.37(2), $p = .31$.001
	4. Strict	15.25 (12), $p = .23$.997	.05 (0–.11)	.046	^{3,4} 3.30(2), $p = .19$.001
Sex (female vs. male)	1. Configural	10.21 (6), $p = .12$.996	.08 (0–.16)	.011		
	2. Metric	10.57 (8), $p = .23$.998	.05 (0–.13)	.013	^{1,2} 0.19(2), $p = .91$	–.002
	3. Scalar	10.75 (10), $p = .38$.999	.03 (0–.11)	.014	^{2,3} 0.23(2), $p = .89$	–.001
	4. Strict	13.95 (12), $p = .30$.998	.04 (0–.11)	.020	^{3,4} 3.46(2), $p = .18$.001

Models were compared with χ^2 difference tests and the superscripts indicate which models were compared in which test (e.g., a model with scalar and a model with strict measurement invariance)

We tested measurement invariance separately for age group, country, and sex. For each case, invariance was evaluated in a two-group CFA with two factors (VWM and VSWM) having two indicators each (average and maximum performance; cf. Fig. 1 for an example)

CFI, Comparative Fit Index; RMSEA, Root Mean Square Error of Approximation; CI, Confidence Interval; SRMR, Standardized Root Mean Square Residual; SB-scaled, Satorra–Bentler scaled; $\Delta\chi^2$, χ^2 -difference test

Discussion

In the current study, we investigated younger and older children's performance on adaptive complex VSWM and VWM span tasks. Specifically, we tested (1) whether VWM and VSWM progressively differentiate with age and (2) whether younger children show performance differences between VWM and VSWM span tasks that are no longer evident in older children. Results showed that older children performed significantly better than younger children in both span tasks, replicating former findings (e.g., Gathercole et al. 2004; Michalczyk et al. 2013). Using confirmatory factor modeling, we found that VSWM and VMW were significantly correlated. Critically, the correlation was significantly higher in younger than in older children.² Moreover, both age groups performed significantly better in the VWM than in the VSWM span task, which is inconsistent with previous data (e.g., Cowan et al. 1987).

The current findings show that working memory processes in the verbal and visuospatial domains are already partially separable in children, which is consistent with the distinction between the phonological loop and visuospatial sketchpad proposed by Baddeley (Baddeley and Hitch 1994). Our findings, however, are also consistent with other models of WM that do not include distinct visuospatial resources (e.g., Barrouillet and Camos 2012; Oberauer and Lewandowsky 2013). Such recent models conceive of a separable verbal component, but take visuospatial processing within the general WM resource. For instance, the latest time-based resource-sharing (TBRS) model of Barrouillet and Camos (2014) differs from Baddeley's model in that it does not include a visuospatial WM sub-system. This model includes a verbal rehearsal component and a general resource (coordinating attentional refreshing), but nothing comparable to the verbal rehearsal component in the visuospatial domain. Consistently, there is evidence, suggesting that visuospatial measures may be part of the general WM resource (i.e., sustained attention; Vergauwe et al. 2010, see also Gray et al. 2017), while verbal WM seems to be a distinct component (e.g., Kane et al. 2004; Park et al. 2002). The separable but

correlated VWM and VSWM processes found in the current study are both consistent with the WM structure in Baddeley's model (i.e., representing the phonological loop and the visuospatial sketchpad) and WM structure proposed by Barrouillet and Camos (i.e., representing the verbal rehearsal component and the general resource). However, although the VWM task primarily relied on verbal rehearsal, it included visual stimuli. Thus, one may argue that it was possible to rely on visual processes instead of verbal rehearsal to solve the task. If children would have mainly used such a visual strategy—even though visuospatial processing is more demanding than verbal processing—this would have contributed to a one-factor solution. Given that our data clearly showed a two-factor model, we consider it unlikely that visual strategies were the primary means to solve the task. Although the two tasks used in this study do not allow distinguishing between general and domain-specific variance in the model, higher correlations between the complex VWM span task and the complex VSWM span task for younger children than for older children suggest at least an age-related differentiation into different components.

Notably, we did not test different WM models against each other (see, for example, Wilhelm et al. 2013) and our data would not even allow for such an approach. We were interested in whether VWM and VSWM progressively differentiate with age, which we found to be consistent with prominent WM models (e.g., Baddeley 2000; Barrouillet and Camos 2012). Our main focus was the comparison of age groups in childhood. Age-related differentiation and age-related performance differences in childhood are important characteristics of WM processing in specific and cognitive functioning in general, because WM capacity places limiting constraints for performing a wide range of other cognitive activities (e.g., Oberauer 2009).

Older as well as younger children performed better in the VWM than in the VSWM span task. This is surprising given the hypothesis that especially, younger children should be better in the VSWM span task, since the use of rehearsal strategies might not be fully developed. It was evident during testing that both age groups tended to name aloud the stimulus they saw on the screen in the VWM span task (cf. Karbach and Kray 2007), indicating that younger children already used verbal rehearsal strategies. This is consistent with the fact that the younger sample in this study was already very verbally capable (see Table S1). Interestingly, children did not use overt verbal strategies in the VSWM span task. Moreover, the VWM and VSWM span tasks used in the present study were structurally and conceptually very similar, which may also have contributed to the fact that children did not perform worse on VWM span tasks.

The present study presents some limitations that need to be discussed. First, one might reason that the less pronounced differentiation of WM processes in younger

² One might criticize the fact that the two indicators for each WM domain were derived from the same task. Of course, two values derived from the same tasks should naturally correlate more strongly than different tasks. However, this has no influence on whether or not the four indicators (two VWM and two VSWM indicators) load on one common factor. Reliable tasks measuring the same underlying psychological construct have common variance over and above the task specific variance some indicators might share. Further, the question of one versus two factors is not central here. The distinction between two strongly correlating factors and one common factor is a cut off on a continuous scale. Rather the notable different correlations between VWM and VSWM for the younger and older children are theoretically important.

children may be due to greater error measurement at that age. However, the correlations were tested in a multi-group model with strict measurement invariance. Thus, the intercepts and residuals for both groups were estimated on the same values, allowing for interpretation of the correlation between the factors. Second, we acknowledge that there is a slight difference between the two complex span tasks because of the processing task used during the encoding phase. Whereas the VWM task taps verbal memory and combines it with visual processing of the stimuli for the processing task (a cross-domain combination), the VSWM task taps spatial memory and visual processing only (a within-domain combination). Thus, interference between storage and processing may have been stronger in the VSWM task, which may have contributed to the finding that children performed better on the VWM than the VSWM task. However, this is just an assumption, since we did not include simple-span tasks to measure storage only.

In sum, we found age-related improvements in both VWM and VSWM from 4–6 to 8–10 years of age. We demonstrated measurement invariance across the age groups, which is encouraging for future longitudinal research, because it shows that these age groups can be compared on the same WM measures. At the same time, a differentiation into more separable VWM and VSWM components occurred, pointing to significant changes in the structure of WM during middle childhood. Further research with more working memory tasks is needed to further differentiate between different theoretical models and their subcomponents and to understand how the system matures during normal development.

Acknowledgements This work was supported by joint grants from the UK Economic and Social Research Council (ESRC, ES/N018877/1) and the German Research Foundation (DFG, KA 3216/2-1). We thank A. Cloes, G. Schaarschmidt, M. Hansmann, R. Sperlich, A. Hall, J. Middleton, S. Ju, and C. Serrano for their help with testing, KoWo - Kommunale Wohnungsgesellschaft mbH Erfurt for providing a test room, and the children and their parents for their participation.

References

- Alloway, T. P., Gathercole, S. E., Adams, A.-M., Willis, C., Eaglen, R., & Lamont, E. (2005). Working memory and phonological awareness as predictors of progress toward early learning goals at school entry. *British Journal of Developmental Psychology*, *23*, 417–426. <https://doi.org/10.1348/026151005X26804>.
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, *80*, 606–621. <https://doi.org/10.1111/j.1467-8624.2009.01282.x>.
- Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, *77*, 1698–1716. <https://doi.org/10.1111/j.1467-8624.2006.00968.x>.
- Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, *87*, 85–106. <https://doi.org/10.1016/j.jecp.2003.10.002>.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, *4*, 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2).
- Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*, 485–493. <https://doi.org/10.1037/0894-4105.8.4.485>.
- Barrouillet, P., & Camos, V. (2012). As time goes by: Temporal constraints in working memory. *Current Directions in Psychological Science*, *21*, 413–419. <https://doi.org/10.1177/0963721412459513>.
- Barrouillet, P., & Camos, V. (2014). On the proper reading of the TBRS model: Reply to Oberauer and Lewandowsky (2014). *Frontiers in Psychology*, *5*, 1331. <https://doi.org/10.3389/fpsyg.2014.01331>.
- Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling*, *12*, 41–75. https://doi.org/10.1207/s15328007sem1201_3.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255. https://doi.org/10.1207/s15328007SEM0902_5.
- Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*, 197–223. <https://doi.org/10.1007/s10648-013-9246-y>.
- Cowan, N., Cartwright, C., Winterowd, C., & Sherk, M. (1987). An adult model of preschool children's speech memory. *Memory & Cognition*, *15*, 511–517. <https://doi.org/10.3758/BF03198385>.
- Cuvo, A. J. (1975). Developmental differences in rehearsal and free recall. *Journal of Experimental Child Psychology*, *19*, 265–278. [https://doi.org/10.1016/0022-0965\(75\)90090-9](https://doi.org/10.1016/0022-0965(75)90090-9).
- Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development*, *37*(2), 283–299. <https://doi.org/10.2307/1126804>.
- Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177–190. <https://doi.org/10.1037/0012-1649.40.2.177>.
- Gray, S., Green, S., Alt, M., Hogan, T. P., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young children and its relation to intelligence. *Journal of Memory and Language*, *92*, 183–201. <https://doi.org/10.1016/j.jml.2016.06.004>.
- Hitch, G. J., Halliday, M. S., & Littler, J. E. (1993). Development of memory span for spoken words: The role of rehearsal and item identification processes. *British Journal of Developmental Psychology*, *11*, 159–169. <https://doi.org/10.1111/j.2044-835X.1993.tb00595.x>.
- Hitch, G., Halliday, M., Schaafstal, A., & Heffernan, T. (1991). Speech, inner speech, and the development of short-term-memory—Effects. *Journal of Experimental Child Psychology*, *51*(2), 220–234. [https://doi.org/10.1016/0022-0965\(91\)90033-O](https://doi.org/10.1016/0022-0965(91)90033-O).
- Jarrold, C., & Towse, J. N. (2006). Individual differences in working memory. *Neuroscience*, *139*, 39–50. <https://doi.org/10.1016/j.neuroscience.2005.07.002>.
- Jarvis, H. L., & Gathercole, S. E. (2003). Verbal and nonverbal working memory and achievements on national curriculum tests at 11 and 14 years of age. *Educational and Child Psychology*, *20*, 123–140.
- 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. <https://doi.org/10.1037/0096-3445.133.2.189>.
- Karbach, J., & Kray, J. (2007). Developmental changes in switching between mental task sets: The influence of verbal labeling in childhood. *Journal of Cognition and Development*, 8, 205–236. <https://doi.org/10.1080/15248370701202430>.
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods*, 44, 978–990. <https://doi.org/10.3758/s13428-013-0348-8>.
- Michalczyk, K., Malstädt, N., Worgt, M., Könen, T., & Hasselhorn, M. (2013). Age differences and measurement invariance of working memory in 5- to 12-year-old children. *European Journal of Psychological Assessment*, 29(3), 220–229. <https://doi.org/10.1027/1015-5759/a000149>.
- Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, 51, 45–100. [https://doi.org/10.1016/S0079-7421\(09\)51002-X](https://doi.org/10.1016/S0079-7421(09)51002-X).
- Oberauer, K., & Lewandowsky, S. (2013). Evidence against decay in verbal working memory. *Journal of Experimental Psychology: General*, 142, 380–411. <https://doi.org/10.1037/a0029588>.
- Ornstein, P. A., Naus, M. J., & Stone, B. P. (1977). Rehearsal training and development differences in memory. *Developmental Psychology*, 13, 15–24. <https://doi.org/10.1037/0012-1649.13.1.15>.
- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, 17, 299–320. <https://doi.org/10.1037/0882-7974.17.2.299>.
- Pickering, S. J., Gathercole, S. E., & Peaker, M. (1998). Verbal and visuo-spatial short-term memory in children: Evidence for common and distinct mechanisms. *Memory and Cognition*, 26, 1117–1130. <https://doi.org/10.3758/bf03201189>.
- Roebbers, C. M., & Zoelch, C. (2005). Erfassung und Struktur des phonologischen und visuell-räumlichen Arbeitsgedächtnisses bei 4-jährigen Kindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37, 113–121. <https://doi.org/10.1026/0049-8637.37.3.113>.
- Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.
- Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *The Journal of Consumer Research*, 25, 78–107. <https://doi.org/10.1086/209528>.
- Studer-Luethi, B., Kodzhabashev, S., Hogrefe, A., & Perrig, W. J. (2015). *BrainTwister2—An extended collection of cognitive training tasks*. Fachstelle für Lernen und Gedächtnis: Universität Bern.
- Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research*, 78, 852–868. <https://doi.org/10.1007/s00426-013-0537-1>.
- Vallar, G., & Papagno, C. (2002). Neuropsychological impairments of verbal short-term memory. In M. Kopelman, A. D. Baddeley, & B. Wilson (Eds.), *Handbook of neuropsychology* (2nd ed., pp. 249–270). Chichester: Wiley.
- Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological Science*, 21, 384–390. <https://doi.org/10.1177/0956797610361340>.
- Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory, and how can we measure it? *Frontiers in Psychology*, 4, 433. <https://doi.org/10.3389/fpsyg.2013.00433>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.