Children's Working Memory Develops at Similar Rates for Sequences Differing in Compressibility

Fabien Mathy & Michael Fartoukh
Université Côte d'Azur

Ori Friedman
University of Waterloo

Nicolas Gauvrit
Université de Lille

Alessandro Guida
Université Rennes II

Determining the causes of working memory maturation is a global issue in developmental cognitive psychology. The question has most often been addressed by manipulating stimulus types within participants. In the present experiment, we instead addressed this question by examining development in children's working memory using a single stimulus type. Primary school children aged 6 to 10 years (total N = 256) participated in a task inspired by the game SIMON®. Children saw sequences of colored squares presented one-at-a-time, and then attempted to reproduce the sequences in the correct order. Some sequences were simple and contained regularities that made them compressible, while other sequences were more complex and less compressible. We found that children's memory for both kinds of sequences improved with age. However the ratio between children's memory for simple and complex sequences did not substantially change with age. These findings suggest that storage capacity of working memory increases in middle childhood, but that children's ability to optimize storage by compressing sequences remains stable across this age range.

Determining the causes of working memory maturation is a major issue in developmental psychology. How do improvements in working memory come about ? One explanation is that these improvements result because with age, stimuli take up less space in memory and are stored more compactly. Another explanation is that improvements result from developmental increases in the space available to store information. In the present study, we test these account using a task inspired by the game Simon N®.

Memory Optimization Ignored in Developmental Studies

Today, young children at all primary schools partake in math classes where they are taught how to divide numbers. Long ago, however, the algorithms for division were only taught in a much less accessible fashion in the best and first European universities of the 16th century, as these algorithms were perceived as extraordinary difficult (Swetz, 1987). This example illustrates that although new ideas and concepts are continuously being discovered, these same concepts can be modified and simplified so that they can be more readily conveyed to others. The rise of conceptual complexity with history is counterbalanced by simplifications (Berthoz, 2012). This 'simplexity' process is important because our ability to process and think about information is constrained by the limits on our ability to hold information in working memory (WM). As such, the ability to compress and simplify information for more efficient storage in WM may have been crucial for historical advances in human intellectual progress, in particular because simpler concepts can leave room for more complex reasoning[1].

This ability to compress and store information in WM may also be crucial for intellectual progress over the course of ontogeny. The expansion of working memory with age has been suspected to be a primary factor responsible for cognitive development for nearly half a century (Pascual-Leone, 1970). One reason is that task demands in any domain are determined by the extent to which the task requires maintenance of information (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002). For example, infants' difficulty with object permanence and the A-not-B task appears to depend on the challenge of maintaining in mind the current location of an object. This idea that working memory and reasoning share related limits concurs with the observation that WM capacity is strongly correlated with cognitive aptitude in adults (Colom, Flores-Mendoza, & Rebollo, 2003; Conway, Kane, & Engle, 2003; Halford, Cowan, & Andrews, 2007).

However, the importance of WM for children's cognitive development does not, in itself, imply that WM develops across much of childhood. Just as human intellectual progress over the last hundred years has not necessarily

[1] There are also conceptions in artificial intelligence which favor a similar general approach (Baum, 2004; Hutter, 2005).

corresponded with improvements in WM (e.g., see Gignac, 2015), it is likewise possible that children could also make progress with constant WM capacity. One motto could be that constant capacity is not incompatible with progress as long as information can be compressed. This observation highlights the importance of better determining the contributions of storage and compression processes to WM capacity development.

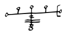
In the present paper, we therefore attempt to separate the ability to store and manipulate information from the ability to process information and to compress it. A better understanding of how storage capacity and compression ability contribute to the growth of WM should help shed light on their relation with the development of intelligence and knowledge in the future. We next review how the development of both factors (storage capacity, compression ability) might each lead to gains in WM performance, and then discuss one reason it is difficult to assess whether these factors are linked with increases in WM performance.

Storage Capacity and Compression Ability in Working Memory

According to one popular view of WM (Luck & Vogel, 1997; Miller, 1956; Rouder et al., 2008), storing information in separate slots in WM is much like packing possessions into boxes. Just as your ability to store possessions is limited by the number of boxes that you have available, people may be limited in the number of chunks of information they can store. The average WM span in adults is 4 ± 1 chunks of information (Cowan, 2001), and this limit applies across diverse kinds of information to be remembered. For example, depending on the task, this limit can apply to digits (running span task), letters (complex span task), multi-feature visual objects (change-detection paradigm), hence, the use of the term *chunk* to represent the variable content of a slot (for a review of these different tasks, see Brady, Konkle, & Alvarez, 2011; Conway et al., 2005). Developmental improvements in WM performance, then, could stem from changes in the number of chunks available to store information.

The number of chunks available may not be the only factor that affects WM performance. In packing boxes, the number of possessions that you pack depends not only on the number of boxes available, but also on your ability to arrange the possessions in the boxes. If you are efficient at packing, you may be able to optimize the available space and fit many possessions into just a few boxes[2].

A similar process has been shown to allow adults to compress information in WM (see Brady, Konkle, & Alvarez, 2009; Thalmann, Souza, & Oberauer, 2019; but ?, ?, for a rebuttal). This process could develop with age, and so developmental increases in WM performance could result (at least in part) from improvements in children's ability to compress information. In comparison with adults and older children, younger children might be inefficient in combining items to be remembered so that the items can be stored with fewer chunks (much as we might expect younger children to be inefficient in arranging possessions in boxes).

Importantly, developmental differences in the ability to compress should matter more for some types of items (or sets of items) than others. One reason is simply that information varies in its complexity and in the extent to which it can be compressed, and some very complex information might be entirely non-compressible. If the ability to compress information improves with age, this could lead to age-related improvements in WM for stimuli that can be easily compressed (e.g., 1 – 9 – 8 – 4 $\hookrightarrow$ 1984). However, development of this ability would not benefit memory for very complex stimuli (i.e., if stimuli such as  cannot be easily compressed, then improvements in compressing ability will be irrelevant).

A Challenge in Understanding Developmental Improvements in Working Memory

People's ability to compress information often depends on the extent to which they are familiar with the stimuli (or types of stimuli). For example, it will be easier to remember the letters J-S-Y-K using a single chunk (rather than four separate items) if you are familiar with the retail store JYSK®, or have used those letters as an acronym for "just so you know". Either way, numerous studies have shown evidence that information from long-term-memory contributes to adults' WM spans (Crannell & Parrish, 1957; Hulme, Maughan, & Brown, 1991; Roodenrys, Hulme, Alban, Ellis, & Brown, 1994; Walker & Hulme, 1999).

This raises the possibility that developmental improvements in WM could *primarily* result from the accumulation of knowledge across childhood (Chi, 1978; Halford et al., 2007; Jones, Gobet, & Pine, 2008). Increases in knowledge stored in long-term memory with age might allow children to optimize storage in WM because the familiar objects could be more rapidly encoded and manipulated (see Barrouillet, Gavens, Vergauwe, Gaillard, & Camos, 2009; Case, 1995; Jones & Macken, 2015, 2018; Reder, Liu, Keinath, & Popov, 2016, and Cowan, 2016 for a general presentation of this is-

[2]Inasmuch as analogies are often deceptive in science, like the analogy from water circuits to electric circuits, it is important to acknowledge that the boxes analogy is inexact. For instance, compressing information in WM might involve a redescription of the data, whereas the content of boxes is not usually transformed – unless using vacuum space saver bags beforehand to squish your possessions to save more space into the boxes)–. Also, our reference to chunks associated to slots is optional as WM capacity could instead depend on a shared continuous resource (Bays, Catalao, & Husain, 2009). One could assume a continuous buffer for holding information in WM, which would not necessarily break information up into discrete chunks. On that single buffer account (a single storage space to continue our analogy), we could still see the size of the buffer grow, and we could still see improvements in efficiency in storing information in the buffer. We only chose the chunking option because of its prevalence in current thinking about WM.

sue). A very nice study by Simmering, Miller, and Bohache (2015) has for instance showed that capacity for common colors increases faster with age than capacity for shapes, but that this difference can be attenuated using less-familiar colors.

Some other studies suggest that WM capacity might develop independently of the accumulation of long-term memory (Burtis, 1982; Cowan, Ricker, Clark, Hinrichs, & Glass, 2015; Gilchrist, Cowan, & Naveh-Benjamin, 2009). For instance, data from Gathercole, Ambridge, Wearing, and Pickering (2004) re-analyzed by Cowan (2016) suggest a rate of development of memory capacity in childhood that is quite independent of the properties of different types of stimuli (e.g., words, non-words, and digits).

Regardless, the main point here is that the accumulation of knowledge in long-term memory can make it difficult to distinguish the extent to which growth of storage capacity and improvements in the ability to compress materials (i.e., the factors we aim to investigate) contribute to developmental increases in WM. To avoid this pitfall, our design used only one kind of stimulus type.

The Present Approach to Understanding Working Memory Capacity Growth

In conducting our study, our chief goal was to examine (A) whether developmental improvements in WM stem from an increase in storage capacity, and (B) whether it stems from improvements in compression ability. To examine these questions, we tested 6- to 10-year-old children using a simple span task inspired by the game (SIMON®). In the task (Gendle & Ransom, 2009; Karpicke & Pisoni, 2004; Mathy, Fartoukh, Gauvrit, & Guida, 2016), children saw sequences of colors (blue, green, yellow, red) which sometimes included repetitions that produce noticeable patterns. Children were asked to recall the sequences in their correct order, and we measured children's spans by determining the sequences that they could correctly reproduce (see the Methods for specific details).

Because the sequences consisted of familiar colors, any developmental increases in WM were unlikely to result from the accumulation of knowledge in long-term memory across childhood. As discussed above, increases in knowledge might allow children to optimize storage in WM for familiar objects. However, we assumed accumulation of knowledge would not matter in our study, because 10-year-olds are unlikely to have much more experience than 6-year-olds with colors as basic as blue, green, yellow, red, considered individually. As such, if the knowledge accumulation was the *only* factor underlying age-related improvements in WM, we would not expect WM to improve with age. Instead, any developmental improvements would be more likely to result from one or both of the factors we wanted to investigate (i.e., increases in storage capacity; improvements in compression ability). We should acknowledge that older children might have greater experience with basic patterns of colors, but this would likely contribute to improvements in compression ability

In our experiment, each child received two blocks of sequences (Simple vs. Complex). These blocks differed in the algorithmic compressibility of the sequences, which is a reliable inverse estimation of their compressibility (detailed in the Method section). This manipulation of complexity (or compressibility) allowed us to test whether improvements in WM resulted from increases in storage capacity and from improvements in compression-ability. Following Mathy et al. (2016), the complex sequences served as a base line for estimating participants' storage capacity. The main difference between this previous study and the present study is that we increased the simplicity of the Simple condition to better induce an optimization process, because this previous study did not find development in compression-ability. Given that complex sequences can scarcely be compressed or chunked, developmental improvements in WM for these stimuli can provides an estimate of increases in storage capacity[3]. In contrast, developmental improvements in WM for simple sequences could result both from gains in storage capacity and from improvements in compression-ability. Further, comparing improvements in WM between the Simple and Complex conditions provides an indication of the extent (if any at all) to which improvements in compression-ability contributes to increases in WM.

To understand which comparisons are most informative, it is helpful to first consider the developmental pattern we would expect if compression-ability does not improve with age. If compression-ability *does not* improve with age, we would expect spans for the simple and complex sequences to increase at the same rate (for related discussion see Burtis, 1982; Case, Kurland, & Goldberg, 1982). For example, if spans for the simple sequences doubled between younger and older children, we would likewise expect spans for the complex sequences to double over the same age range. To illustrate this example, suppose that younger children have a span of just 2 chunks, which can either be used to encode sequences of simple sequences of 4 colors, or complex sequences of just 3 colors (i.e., because the simple sequences are more compressible than the complex ones). If children at some older age have twice as many chunks available (i.e., 4), then they will also be able to remember sequences that are twice as long, that is simple sequences of 8 colors or complex ones of 6 colors. A key point to recognize, here, is that this predicted outcome involves an interaction between sequence complexity and age (i.e., $4 \longrightarrow 8$ involves a greater *absolute* increase than $3 \longrightarrow 6$), but a very specific kind of interaction. That is, in the predicted interaction, the ratio between the spans for simple and complex sequences is constant at

[3]Note that this approximation method can overestimate storage because a few regularities are still present in complex sequences.

each age (i.e., younger children's Simple:Complex ratio of 4:3 is identical to older children's ratio of 8:6). This type of progression corresponds to proportional growths as there is a 100% increase in the span for both the simple (4⟶8) and complex (3⟶6) conditions. Similarly, younger children can remember 33% more items (3⟶4) in the Simple condition than in the Complex condition, as well as older children (6⟶8 corresponding to a 33% increase).

Given this background, we can now describe the findings that should be expected if compression-ability does improve with age. As before, younger children would be expected to have longer spans for simple sequences than for complex ones. As before, this difference is expected given that simple sequences are more compressible than complex ones. Assuming that storage capacity (e.g., number of chunks) increases with age, we would also expect a developmental increase in children's memory for complex sequences. However, memory for simple sequences should increase at a greater rate because any improvements in compression-ability would have a *greater* impact on simple sequences than complex ones. So if we again suppose that the complex sequences that older children remember are twice as long as those that younger children can remember (or any other ratio > 1), we would expect an even greater improvement with age for simple sequences. As before, this predicted outcome involves an interaction between sequence complexity and age. But crucially, this interaction would *not* involve constant ratios between the spans for simple and complex sequences at each age. On the contrary, it predicts that Simple:Complex ratios will increase with age. Said otherwise, such a pattern of results would correspond to non-proportional growths (i.e., different percentages of increases across conditions).

Method

We examined whether the rate of improvement of children's WM capacity for sequences of items depends on their complexity. Children aged 6-11 completed a task inspired by the electronic children's toy, SIMON®, in which they attempted to reproduce sequences of colors varying in their complexity.

Participants. Our sample included 256 children. These were: 47 6-year-olds (*M* = 78 months, *SD* = 3.0), 48 7-year-olds (*M* = 89 months, *SD* = 3.1), 63 8-year-olds (*M* = 101 months, *SD* = 3.5), 50 9-year-olds (*M* = 113 months, *SD* = 3.3), and 48 10-year-olds (*M* = 126 months, *SD* = 3.4).

As local public schools were chosen randomly in the area, the children were on average from middle-class families. We tested the children after consent was obtained. None of the selected participants were colorblind (i.e., we asked children to name the four colors used in the task before the task began, and results were discarded for children who made any error in naming the colors). We also excluded results from eight children (all aged 6) who had a span of zero in the subsequent tested sequences.

In recruiting participants, we aimed to have [approximately] 50 participants per age group. The sample exceeds this size at some ages for which we received more consents than expected. Our sample also greatly exceeds the recommended sample size of 10 participants per age group yielded by a power analysis[4] requiring a difference of at least one item between the two conditions at a given age, a standard deviation equal to 1, and a power of .80.

Procedure and Materials. Participants were seen individually. They were first instructed on how to use the computer and on how to best recall the sequences (e.g., in the correct order and by paying attention to eventual repetitions of colors). Participants were also told that the game would become more and more difficult as they would have to progressively recall longer sequences.

In the task, participants completed a series of trials in which they saw sequences of colored rectangles (6 cm long × 4 cm wide) on a computer screen, and attempted to recall each sequence in the correct order. Each sequence started with a fixation cross centered on the screen for 1000 ms. The colored squares then serially appeared at the centre of a grey screen (i.e., one at-a-time) at the speed of 1000 ms per square with a 400 ms inter-item blank. After each sequence, the grey screen displayed four colored buttons (four rectangles of 6 cm long × 4 cm wide, and separated spaces of 1 cm) arranged in a quadrant, and participants clicked these using a computer mouse to recall the sequence. Participants indicated they had completed each sequence by pressing the space bar. The screen then provided feedback with the message "perfect" or "not exactly", which the experimenter read aloud to make sure the children would do their best the next trial.

To estimate WM capacity[5], participants received test sequences in two successive within-subject conditions: Simple or Complex (order counterbalanced between participants). In each condition, the sequences were three items long initially, but progressively increased in length every four trials up to a maximum length of 8 successive items. As such, participants could complete up to 24 test sequences (per condition),

[4]We used the function pwr.t.test of the pwr package in R (R Core Team, 2017), with the parameters $d = 1, sig.level = 0.05, type = paired, alternative = two.sided$).

[5]Memory tasks presented over short retention intervals can be divided into two categories: short-term memory (STM) tasks (or simple span tasks) where individuals are asked to just maintain items (e.g., "keep in mind the following sequence: 7 4 3 9") and WM tasks (or complex span tasks) where a processing task is added (e.g., "keep in mind the following sequence: 7 4 3 9, while saying continuously *super*"). Although STM and WM are estimated differently, it has been shown that memory span tasks of both types tap the same underlying construct (Colom, Shih, Flores-Mendoza, & Quiroga, 2006; Unsworth & Engle, 2006), and so in the present paper we refer to WM only (and do not distinguish it from STM).

though testing in each condition was stopped after four consecutive errors at a given list length.

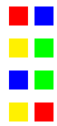
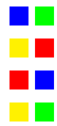
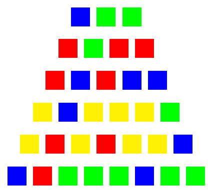
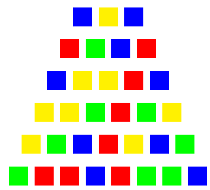
Using the R function ACSS, we assigned sequences to each condition (Simple vs. Complex) by first computing (for thousands of sequences of different lengths) their algorithmic complexity (Kolmogorov, 1965; Li & Vitányi, 1997), which inversely corresponds to their compressibility. The algorithmic complexity of a sequence can be thought as the shortest algorithm able to reproduce it. Simple sequences can be encoded with short algorithms, and are therefore highly compressible. In contrast, complex sequences can only be encoded with longer algorithms, and are less compressible. The same R function also provided an estimation of the complexity of all participants' responses.[6]. From this set, sequences in the Simple condition were at the 25^{th} percentile of complexity (a lower percentile would have produced sequences of several similar colors), while those in the Complex condition were close to the 100^{th} percentile. Table 1 shows a sample sequences given to participants. Note that the actual colors were randomly chosen, after the pattern was selected for the current trial (for instance, the pattern 112 could be displayed as blue-blue-red, green-green-red, or any other combinations). It must be emphasized that the four colors are monosyllabic in French.

Each condition began with four warm-up sequences, included to allow participants to get accustomed with the procedure. The warm-up sequences each featured two different colors, as shown in Table 1. Warm-up sequences were programmed such that each of the four colors appeared at least once in the first two sequences, and at least twice across all four warm-ups. Participants were invited to respond slowly during the first warm-up trial to make sure they responded correctly. All participants succeeded in every warm-up trial.

In each new sequence, the positions of the colors in the quadrant of response options were randomized to discourage spatial encoding (for instance red on north-west, blue on north-east, green on south-west, and yellow on south-east, but this configuration randomly switched every next trial). To pace the task and to mimic the ergonomy of the original game, whenever participants clicked one button of the quadrant, the color lit up to show a lighter color (e.g., the red turned light red) for a brief duration (300 ms) during which all buttons were disabled.

Scoring and Analysis. Participants' spans were calculated using a refined partial-credit scoring system (Conway et al., 2005), in which we computed the proportion of items recalled at their correct position for each sequence. This proportion was then multiplied by 0.25 points for each sequence (i.e., since there was 4 trials per list length). These points were summed across sequences to obtain an estimate of the span. Two points were automatically granted to each participants to take into account the absence of one-color trials and the successful warm-up trials. The maximum possible score

Table 1
Sample of sequences in the Simple and Complex conditions.
Note. *After the warm-up trials, we only show one example of sequence for a given length. The complete list of sequences and their corresponding complexity is provided in the data file on http://XXX-OSF-XXX.*

Length	(Simple)	(Complex)
2 2 2 2		
	— End of Warm-up—	— End of Warm-up—
3 4 5 6 7 8		

was 8 (corresponding to the maximal sequence length).

Results

Table 2 shows the mean spans and transformed spans averaged across participants as a function of age group, separated by the complexity of the sequences (Simple vs. Complex).

Our main analysis tests whether WM develops at different rates or at the same rate for simple and complex sequences. Findings that WM develops at different rates for both kinds of sequences would mean that compression-ability and storage capacity both increase with age. In contrast, findings that WM develops at the same rate for both kinds of sequences would suggest that storage capacity increases with age, but compression-ability does not. We also report some further analyses using other techniques (detailed in the Appendix) that follow up the main analyses.

All inferential statistics were run in JASP (retrieved from http://jasp-stats.org/) with defaults parameters. We first an-

[6]We refer the reader to Gauvrit, Zenil, Delahaye, and Soler-Toscano (2014); Gauvrit, Singmann, Soler-Toscano, and Zenil (2015), who developed the technique consisting in first generating billions of deterministic algorithms to obtain the Solomonoff-Levin algorithmic probability of a string s. This method provides a reliable approximation of the algorithmic complexity of short strings (see also Delahaye & Zenil, 2012; Soler-Toscano, Zenil, Delahaye, & Gauvrit, 2013; Gauvrit et al., 2014) and the metric has already proven useful in cognitive psychology (Chekaf, Gauvrit, Guida, & Mathy, 2018; Dieguez, Wagner-Egger, & Gauvrit, 2015; Gauvrit et al., 2015, 2014; Kempe, Gauvrit, & Forsyth, 2015).

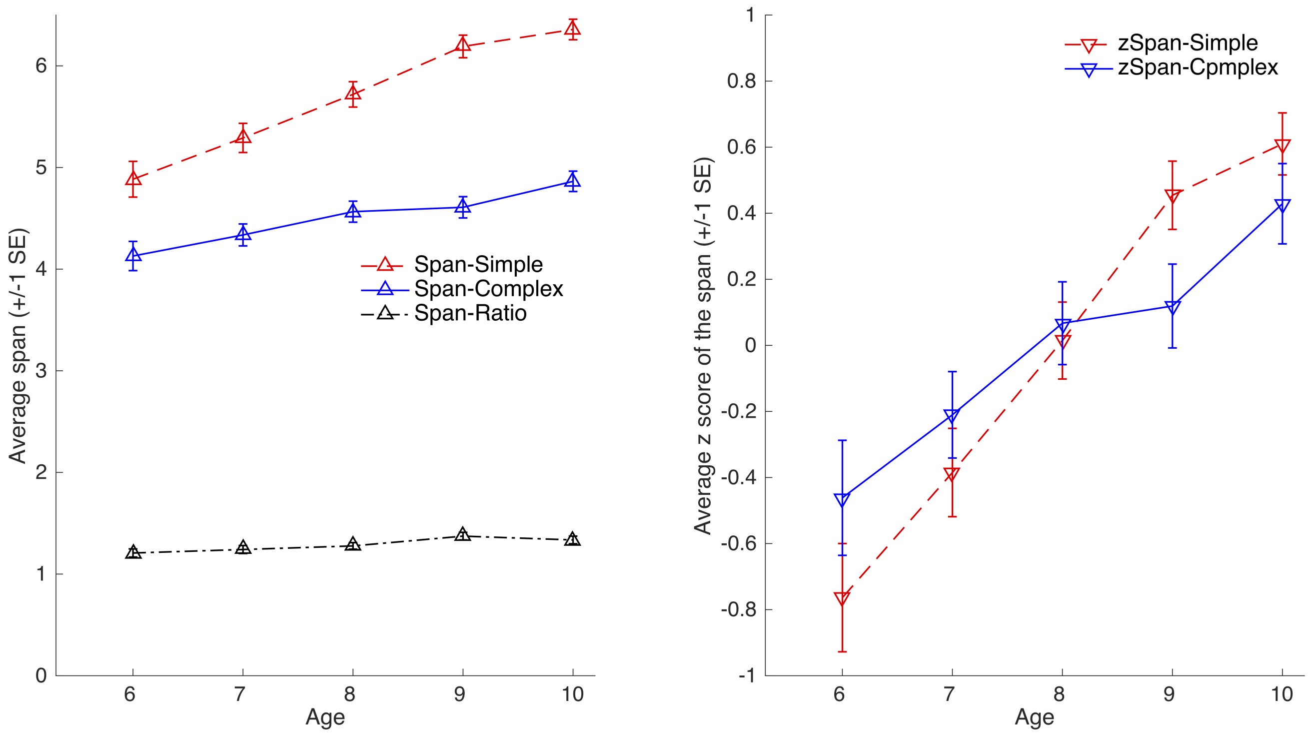


Figure 1. Mean spans averaged across participants and average ratios across participants between the two average spans (LEFT side), and average z-scores based on the mean and standard deviations of the spans across age groups (RIGHT side), as a function of age and complexity of the sequences. *Note.* Error bars are +/-1 standard errors.

alyzed the individuals' spans using a two-way repeated-measures analysis of variance with the within-subject factor Complexity (Simple vs. Complex) and the between-subjects factor Age group (6, 7, 8, 9, and 10); see Figure 1. The principal goal of such analysis is to detect an interaction, without which it could not be shown that WM develops at different rates for simple and complex sequences. The results revealed an effect of Age ($F(4, 251) = 17.1, p < .001, \eta_p^2 = .21$); the posthoc tests with a Bonferroni correction showed significant differences between ages whenever there was at least a two-year gap between two age conditions. The results also showed an effect of Complexity ($F(1, 251) = 428, p < .001, \eta_p^2 = .63$) and an interaction between the two factors ($F(4, 251) = 7.2, p < .001, \eta_p^2 = .10$). The interaction was a bit weaker in analyses using the log of the data ($\eta_p^2 = .06$), but still remained significant ($p = .004$). Given the interaction, the subsequent analysis based on the ratios will provide more information on whether Simple and Complex show different slopes with age.

A Bayesian repeated-measures ANOVA (using the within-subject factor Complexity and the between-subjects Age group) confirmed that the full model (i.e., both factors including the interaction term) showed evidence against the null model ($BF_{10} = 2.4e + 63$) and the interaction term increased the model probability ($BF_M = 3650$). All other more restrictive models had BF_M inferior to 1.

We next ran a one-way ANOVA to test whether the Simple:Complex ratio increased with age. The ANOVA showed a significant effect of Age group, but with a very small effect size ($F(4, 251) = 3.26, p = .012, \eta_p^2 = .05$). The posthoc analysis with a Bonferroni correction showed that this significant difference was only due to a significant effect of the difference between the 6-year-old group and the 9-year-old group. The respective means and SE were 1.20 (0.04), 1.24 (0.04), 1.28 (0.03), 1.37 (0.04), 1.34 (0.04). The ratios ranged from 1.2 to less than 1.4. This does not correspond with a dramatic increase or with meaningful gain. Thus, even though we found a significant effect of the Ratio factor, which corresponded to a linear increase confirmed by a linear regression using a continuous variable Age ($F(1, 254) = 10.2, p = .002, R^2 = .04$), the Bayesian ANOVA found no evidence of the effect of Age group ($BF_{10} = 1.8$) on the ratios.

We next examined z scores, computed separately for each Complexity condition. This analysis was included to allow us to detect coincidental growth curves, which would indicate that WM develops at similar rates for simple and complex sequences. Analysis of the z scores showed only a sig-

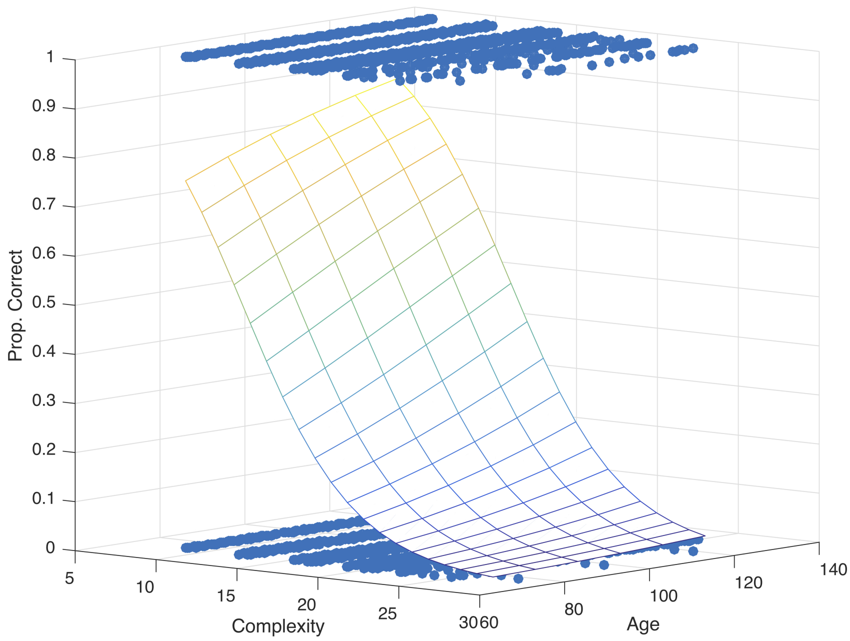


Figure 2. Logistic regression of correct response as a function of sequence complexity and age (in months).

Table 2
Average spans and their ratio as a function of age group, with standard errors in parentheses. Note. The last two columns are the respective z-scores of the spans. The z transformation was obtained by taking the overall mean and SD of each condition of the complexity factor across age groups. The ratios and z scores were computed for each participant before being averaged.

Age	Simple	Complex	Ratio	z_{Simple}	z_{Complex}
		Partial-credit-scoring system			
6	4.86	4.11	1.20	-.77	-.48
7	5.30	4.36	1.24	-.36	-.18
8	5.72	4.57	1.28	.03	.07
9	6.18	4.62	1.37	.45	.13
10	6.36	4.86	1.34	.61	.43
6	(.18)	(.14)	(.04)	(.16)	(.17)
7	(.14)	(.11)	(.04)	(.13)	(.13)
8	(.12)	(.10)	(.03)	(.12)	(.13)
9	(.11)	(.10)	(.04)	(.10)	(.13)
10	(.10)	(.10)	(.04)	(.09)	(.12)

nificant effect of Age group ($BF_{10} = 6e + 8$) and the interaction term increased the model probability ($BF_M = 12$). No other model showed a $BF_M > 1$. Therefore, the analysis based on the z score provides a conclusion similar to the one using the ratios. However, as can be seen in Figure 1 (graph on right), there is no intuitive way to interpret the differential developmental effect represented by the two non-coinciding curves (the null hypothesis of an absence of effect of the complexity as a function of age corresponds to two coinciding curves).

To obtain more precise results, we also ran analyses using continuous values of age-in-months and complexity level of each sequence. Specifically, we conducted a logistic regression on all trials ($N = 7454$, scored as correct vs. incorrect recalled sequence) as a function of complexity and age levels. The logistic regression plotted in Figure 2 showed a significant effect of age, a significant effect of complexity and a significant interaction between the two variables. The respective estimates were .04 ($p < .001$), -.19($p < .001$), and -.001 ($p = .006$) with an intercept of .44. The interaction is visible in Figure 2 in which the detrimental effect of complexity is more pronounced in younger participants than in older participants. For instance, age did not affect the proportion correct when the complexity was the highest (purple lines), whereas older children most benefited from the sequences with the lowest complexity (yellow line). Again, as demonstrated above, it does not mean that this gain was sufficiently important to correspond to different growths between complexity conditions.

When we measured the actual algorithmic complexity of the responses themselves and compared it to the original complexity of the stimulus sequences, we observed a lower average complexity across trials for the responses ($M = 14.1$) than for the stimuli ($M = 14.8$), $M_{diff} = .66, t(7453) = -20, p < .001, d_{Cohen} = .23$. The same analysis aggregated by participant produced a larger size effect: $M_{diff} = .63, t(255) = -12, p < .001, d_{cohen} = .77$. Figure 3 shows that this trend was stable with age. We observed no correlation between age and the average difference of complexity between responses and sequences, in both experimental conditions (Figure 3). In the Discussion, we suggest an explanation for this reduction in complexity.

Discussion

Compressibility effect. We found that participants' spans for simple and complex sequences of colors improved with age. However, memory for simple and complex sequences improved at about the same rate. This finding was revealed by our analysis of the ratios between children's memory for the simple and complex sequences. We also found a main effect of complexity, as children showed better memory for simple sequences than complex ones. This difference suggests that children were able to compress information in the simple sequences. Further evidence that children's memory was aided by compression comes from the finding that mis-remembered sequences were less complex than the original sequences that children attempted to reproduce.

Further evidence that children compressed the stimuli comes from their patterns of error. As noted above, participants' generally erred by *reducing* the complexity of the to-be-remembered sequences (i.e., their wrongly recalled sequences were simpler than the originals). This would not be expected if participants' had made random errors. In fact, random errors would have led to the opposite result, and would have increased the complexity of participants' output (recall that algorithmic complexity is the highest for random sequences). Hence, one explanation for the observed reduction in complexity is that a compression process occurred.

Age-related increase in storage capacity . Our findings support three main conclusions. First, they suggest that with age, there is an increase in the storage capacity of WM. Second, they suggest that children aged 6 are already able to compress information in working memory. Third, they suggest that, at least within the age range examined, children's ability to compress information does not improve with age. We next expand on these three conclusions.

Our findings suggest an age-related increase in the storage capacity of children's working memory. This could stem from an age-related increase in the number of chunks children can encode[7]. Previous studies have also shown that improvements in storage capacity remain when controlling for chunking-ability or other strategies that could potentially develop with age (Burtis, 1982; Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2010; Cowan, AuBuchon, Gilchrist, Ricker, & Saults, 2011; Cowan, Morey, et al., 2010; Cowan et al., 2015). Our results are generally consistent with such findings, because we observed a similar rate of increase across our conditions. The general growth we observed corresponds approximately to a 5% increase between age groups (after computing the geometric mean of the growths observed in the spans). For instance, had a material of average complexity produced a base line span of 4.5 at 6 years old, we could have expected a span equal to $4.5 \times 1.05^4 = 5.5$ at 10 years old with a constant 5% increase (the exponent 4 represents the 4 steps between the 5 age groups).

Our findings also suggest that children were able to compress information (at least for simple sequences) in WM. We draw this conclusion from a somewhat novel feature of our study. Specifically, the simple and complex sequences both consisted of the same kinds of items, colored squares. As such, the difference in WM spans we observed between the sequences likely resulted because the simpler sequences were more compressible than the complex ones. This is especially likely given that categorization of sequences as simple or complex was itself based on a measure of complexity. Children's superior memory for the simple sequences therefore provides evidence of their ability to rapidly compress stimulus materials (see Kibbe & Feigenson, 2014a, 2016; Feigenson & Halberda, 2008 for previous evidence of this ability in infants). This rapid process that seems to take place exclusively in WM (since none of the color patterns corresponded to pre-learned chunks) casts doubt that chunking could be achieved by redintegration (Norris, Kalm, & Hall, 2019). Redintegration would occur if participants used information previously stored in long-term memory to reconstruct degraded traces in working-memory. However, this account cannot explain our findings because the patterns were likely unfamiliar to participants, and hence not already in their long term memory. Our findings instead suggest that the sequences discovered on the spot are actively recoded into chunks directly in WM.

The final major point suggested by our findings is that children's ability to compress sequences does not improve with age. We chiefly base this conclusion on the ratio between memory for simple and complex sequences. As noted above, the simple sequences were more compressible than complex ones. This means that improvements in compression-ability memory should lead to greater improve-

[7]However, as noted above, increases in storage capacity could also occur even if WM does not encode information in discrete chunks (WM capacity could depend on a continuous resource according to Bays et al., 2009).

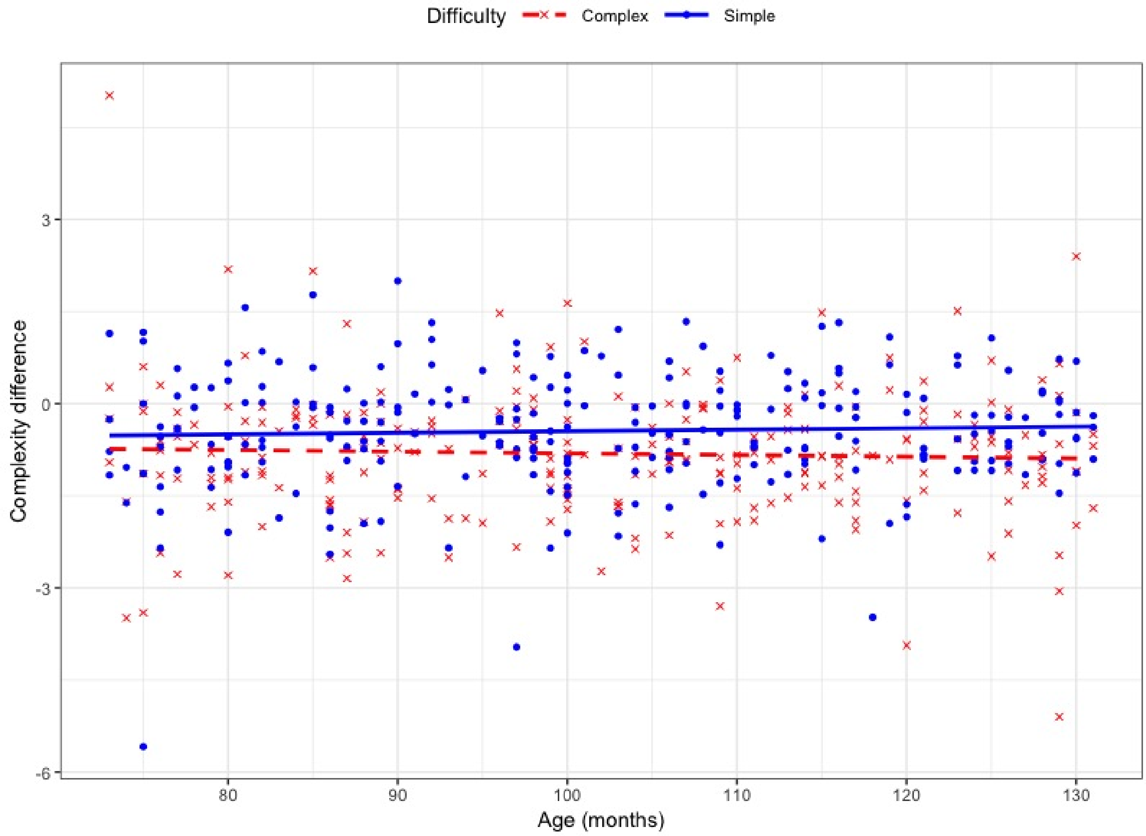


Figure 3. Difference in algorithmic complexity between the stimulus sequences and the participants' responses as a function of age, collapsed by participants and experimental condition (one dot = one participant in one condition).

ments in memory for simple sequences than for complex ones. Against this prediction, we found that memory improved at similar rates for both kinds of sequences. That is, the Simple:Complex ratio did not substantially improve with age and as such we cannot conclude that a concurrent optimization process develops to alleviate the storage process.

As speculative as it might sound, the absence of increase in storage optimization with age (i.e., better compression of simpler sequences) in our results could mean that a greater compactness of information could not be achieved by older children, and as such, this suggests that chunk size did not increase with age. To return to the boxes analogy, the number of boxes available for packing possessions increased, but not ability to arrange the possessions in the boxes. If this interpretation of our results is correct, these findings are broadly consistent with previous developmental studies which have found that with age, children increase the number of chunks they can recall, without a corresponding increase in the size of the chunks (Cowan, Hismjatullina, et al., 2010; Gilchrist et al., 2009; one exception is Ottem, Lian, & Karlsen, 2007). However, our results (and those previous ones), might seem to contradict other findings showing that adults benefit from increases in chunk size (i.e., rather than increase in the number of chunks; (Cowan, Chen, & Rouder, 2004; Gobet & Clarkson, 2004; Ericsson, Chase, & Faloon, 1980; Tulving & Patkau, 1962)). But overall, we can reasonably expect the cognitive system to favor an increase in chunk size once working memory capacity in adults has reached a fixed capacity and all of these results finally appear consistent.

Limitations. One potential limitation of our finding relates to children's speech rate. Previous studies have shown a clear relation between speech rate and memory span (Hulme, Thomson, Muir, & Lawrence, 1984). This introduces a potential confound for effects of age, because speech rate increases with age. In fact, children with a high speech rate have a span similar to adults with a low speech rate. This is a concern for our study because participants were free to encode items verbally. Higher speech rate could especially have impacted results for simple sequences, as they included more repetitions than complex sequences (e.g., it may be faster to repeat 'blue-blue' than 'blue-red'). However, we do not believe that this potential confound undermines our conclusions because higher speech rate with age would have especially favored the Simple condition and this should have caused increasing ratios between the Simple and Complex condition.

A more definite limitation is that because our sample of children ranged between ages 6 and 10, our findings do not speak to the working memory abilities of younger children. It is possible, for example, that we would have found evidence for improvements in compression-ability if we had included younger children in our sample. The same reasoning also applies for children older than 10. Findings also might have differed with other kinds of stimulus materials. We have already reviewed some benefits of having both simple and complex sequences only consisting of colored squares. However, this decision could limit the generalizability of our findings, as it remains possible that improvements in compression-ability might turn up for other kinds of stimuli. It is even possible that improvements in compression-ability would turn up using a task removing the requirement that children recall items in the correct order. One reason for this is that less constrained tasks may be more amenable to compression (e.g., items could be re-ordered more freely for optimal storage in memory). We hope to address some of these possibilities in future research.

The comparison of our findings with previous studies (which for instance manipulated the number of chunks and chunk size, see Norris et al., 2019) is only speculative because our method does not provide a direct measure of the chunks formed by participants. In the present study, we still do not know whether (or how) chunks were formed in either of the experimental conditions. Algorithmic complexity is not yet useful for determining how a sequence can be optimally recoded into separate chunks. A similar issue interests computer scientists who seek to split a file into smaller chunks for optimizing cloud storage (Widodo, Lim, & Atiquzzaman, 2017). Here, we simply assume that greater compressibility most probably correlates with more occasions to recode a sequence into chunks to optimize the storage process. Transition errors probabilities could have been computed (e.g., Johnson, 1969; Mathy et al., 2016) to detect whether repetitions of colors helped participants better recall the simple sequences, but this method can be deceptive (mainly because chunks do not necessarily correspond to color repetitions only). Also, comparing the responses and the stimulus sequences on a trial-by-trial basis requires one to compute an alignment between the chunked responses and chunked stimulus sequences (see Mathy & Varré, 2013). However, alignment algorithms need to be parameterized to indicate the plausibility of errors (deletions, substitution, transpositions, insertions, etc). Such decisions would thus require estimating the probabilities of these errors for each age groups to obtain the best alignments. We believe that this method would bring more questions than actual responses. Future work could attempt to identify the exact chunks participants formed to better understand how they mentally reorganized the material. Related to this, compression could be related to prediction. Simple sequences contain regularities that make them more predictable. Optimizing memory through the prediction of ongoing to-be-encoded events is not the same as compressing information after their encoding. If participants expect more regular patterns as the presentation of a sequence progresses, this would facilitate the retention of the sequences. One potential solution to this issue would be that prediction and compression work together as the sequence progresses to encode information optimally.

Conclusion. Our results are relevant for previous studies showing that recoding of information can operate on the spot (Bor & Owen, 2007; Bower & Winzenz, 1969; Mathy & Feldman, 2012). The rapidity of the recoding process does not seem to depend much on whether chunking is explicit (e.g., Ericsson et al., 1980; Gobet et al., 2001; Hu & Ericsson, 2012; Klahr, Chase, & Lovelace, 1983), or implicit (Perruchet & Pacteau, 1990; Servan-Schreiber & Anderson, 1990). For instance, it has been suggested that children acquire language by relying on such rapid chunking mechanisms (Perruchet, Poulin-Charronnat, Tillmann, & Peereman, 2014; Saffran, Aslin, & Newport, 1996), that might help them compress information rapidly to deal with the deluge of linguistic input (Chater, Clark, Goldsmith, & Perfors, 2015; Christiansen & Chater, 2016). Rapid chunking appears to be a fundamental process found in pigeons (Terrace, 1987), infants (Kibbe & Feigenson, 2016; Feigenson & Halberda, 2004, 2008), and toddlers (Kibbe & Feigenson, 2014b). Future studies could thus further study whether the capacity to recode on the spot information into separate chunks can subserve a potential compression process that the present study aimed to highlight in children.

Acknowledgements

This research was supported in part by a grant from the Agence Nationale de la Recherche (ANR-17-CE28-0013-01) awarded to Fabien Mathy. Author contributions: F.M. designed the research; M.F. performed the research; F.M. analyzed the data; N. G. computed the algorithmic complexity; All authors wrote the paper. Correspondence concerning this article should be addressed to fabien.mathy@univ-cotedazur.fr or BCL Lab, CNRS, UMR 7320, Université Côte d'Azur, Nice, France.

Appendix

We posit that if optimization of storage capacity develops with age (i.e., due to improvements in compression-ability), then WM should improve at a greater rate for simple sequences than for complex ones. If not, WM should improve at approximately similar rates.

In the main text, the main method we used to test between these two possibilities was examining whether the

Simple:Complex ratios changed with age. Here We argue that the ratio analysis is the most convenient to detect different rates. Let's imagine that an old 1-liter engine used to produce 100 horse powers, but has been upgraded so that it now produces 120 horse powers. In comparison, an old 2-liter engine used to produce 150 horse powers, but has been upgraded so that it now produces 180 horse powers. The absolute difference is larger for the 2-liter engine. However, as both new engines increased power by 20%, it would be unfair to say that the larger engine has been better optimized than the first. To conclude that one engine was optimized more than the other, we cannot look to the magnitude of improvement. Instead, we should look to the relative gain that can be measured by ratios or odds-ratios. Here, the ratio $130 : 100 = 1.3$ is superior to $180 : 150 = 1.2$, with odds-ratio $(130 : 100) : (180 : 150) > 1$). One needs to take into account the size of the engine to get a sense of proportion and we believe the same reasoning applies to the measure of WM storage capacity.

Here we present three analytic approaches for examining...:

1. *Detecting increasing ratios of the spans*. Constant ratios are the signature for proportional growths and increasing ratios can represent different rates of improvements intuitively.

2. *Detecting identical z-scores of the spans*. Identical z scores correspond to proportional growths.

3. *Detecting an interaction after a logarithmic transformation of the spans*. The interaction detects different growth rates.

The method based on the spans' ratios is straightforward. For instance, doubling capacity for simple stimuli at a younger age, but tripling capacity at an older age, no matter the base line. The spans' ratios (span for the Simple condition divided by span for the Complex condition) can provide an intuitive estimate of chunk size. If children at a given age succeed in doubling their span in the simpler condition, it means that they can compact twice as much information than in the complex condition. If this ratio is greater for older children, we can assume that the material benefited from a greater compactness. Also, the odds ratio are convenient to indicate for instance that older children could compact $3/2 = 1.5$ times more material than younger children.

The method using z scores is less direct than computing the spans' ratios; z scores are easy to interpret when they perfectly coincide (when the curves overlap, they reflect equivalent rates of development). However, they are more difficult to interpret when the scores do not coincide (for instance, when crossing curves are obtained, like in our results).

The method based on logarithms might leave the impression that the radical non-linear transformation is a way to cheat with numbers, whereas they simply transform proportional growths into similar growths. The logarithm (no matter its base, since $LOG_a(x)LN(a) = LN(x)$) is an interesting transformation of the data to capture a multiplicative effect (such a transformation places numbers into a geometric domain for numbers that are multiplicative by nature; see Kerkhoff & Enquist, 2009). The log transformation is thus informative when for instance it reveals parallel curves. The method can also capture different growths if one interaction is present after the log transformation. Still, most readers do not appreciate juggling from multiplicative effects to additive effects. For that reason mainly, a non-linear transformation of the data using a logarithm might not be the most direct approach and we conclude with the idea that computing the ratios remains best.

References

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: a time-based resource-sharing model account. *Developmental psychology*, *45*, 477-490.

Baum, E. B. (2004). *What is thought ?* Cambridge, MA: MIT Press.

Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*, 1-11.

Berthoz, A. (2012). *Simplexity: Simplifying principles for a complex world (g. weiss, trans.).* Yale University Press.

Bor, D., & Owen, A. M. (2007). A common prefrontal-parietal network for mnemonic and mathematical recoding strategies within working memory. *Cerebral cortex*, *17*, 778-786.

Bower, G. H., & Winzenz, D. (1969). Group structure, coding, and memory for digit series. *Journal of Experimental Psychology*, *80*, 1 - 17.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*, 487-502.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*, 1-34.

Burtis, P. J. (1982). Capacity increase and chunking in the development of short-term memory. *Journal of Experimental Child Psychology*, *34*, 387-413.

Case, R. (1995). Memory performance and competencies: Issues in growth and development. In F. Weinert & W. Schneider (Eds.), (chap. Capacity based explanations of working memory growth: a brief history and reevaluation). Mahwah, NJ: Erlbaum.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*, 386-404.

Chater, N., Clark, A., Goldsmith, J. A., & Perfors, A. (2015). *Empiricism and language learnability*. OUP Oxford.

Chekaf, M., Gauvrit, N., Guida, A., & Mathy, F. (2018). Compression in working memory and its relationship with fluid intelligence. *Cognitive Science*, *2*, 904-922.

Chi, M. T. (1978). Knowledge structures and memory development. *Children's thinking: What develops*, *1*, 75-96.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*.

Colom, R., Flores-Mendoza, C., & Rebollo, I. (2003). Working memory and intelligence. *Personality and Individual Differences*, *34*(1), 33–39.

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

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163-183.

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*(5), 769-786.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*, 547-552.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-185.

Cowan, N. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science*, *11*, 239–264.

Cowan, N., AuBuchon, A. M., Gilchrist, A., Ricker, T. J., & Saults, J. S. (2011). Age differences in visual working memory capacity: not based on encoding limitations. *Developmental Science*, *14*, 1066-1074.

Cowan, N., Chen, Z., & Rouder, J. N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to miller (1956). *Psychological Science*, *15*, 634-640.

Cowan, N., Hismjatullina, A., AuBuchon, A. M., Saults, J. S., Horton, N., Leadbitter, K., & Towse, J. (2010). With development, list recall includes more chunks, not just larger ones. *Developmental Psychology*, *46*, 1119-1131.

Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science*, *13*, 120-133.

Cowan, N., Ricker, T. J., Clark, K. M., Hinrichs, G. A., & Glass, B. A. (2015). Knowledge cannot explain the developmental growth of working memory capacity. *Developmental science*, *18*, 132-145.

Crannell, C. W., & Parrish, J. M. (1957). A comparison of immediate memory span for digits, letters, and words. *Journal of Psychology: Interdisciplinary and Applied*, *44*, 319-327.

Delahaye, J.-P., & Zenil, H. (2012). Numerical evaluation of algorithmic complexity for short strings: A glance into the innermost structure of randomness. *Applied Mathematics and Computation*, *219*(1), 63–77.

Dieguez, S., Wagner-Egger, P., & Gauvrit, N. (2015). Nothing happens by accident, or does it? a low prior for randomness does not explain belief in conspiracy theories. *Psychological science*, *26*, 1762-1770.

Ericsson, K. A., Chase, W., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, *208*, 1181–1182.

Feigenson, L., & Halberda, J. (2004). Infants chunk object arrays into sets of individuals. *Cognition*, *91*(2), 173-190.

Feigenson, L., & Halberda, J. (2008). Conceptual knowledge increases infants' memory capacity. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 9926-9930.

Gathercole, S., Ambridge, B., Wearing, H., & Pickering, S. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177-190.

Gauvrit, N., Singmann, H., Soler-Toscano, F., & Zenil, H. (2015). Algorithmic complexity for psychology: a user-friendly implementation of the coding theorem method. *Behavior research methods*, 1-16.

Gauvrit, N., Zenil, H., Delahaye, J.-P., & Soler-Toscano, F. (2014). Algorithmic complexity for short binary strings applied to psychology: a primer. *Behavior research methods*, *46*, 732-744.

Gendle, M. H., & Ransom, M. R. (2009). Use of the electronic game simon® as a measure of working memory span in college age adults. *Journal of Behavioral and Neuroscience Research*, *4*, 1-7.

Gignac, G. E. (2015). The magical numbers 7 and 4 are resistant to the Flynn effect: No evidence for increases in forward or backward recall across 85 years of data. *Intelligence*, *48*, 85-95.

Gilchrist, A. L., Cowan, N., & Naveh-Benjamin, M. (2009). Investigating the childhood development of working memory using sentences: new evidence for the growth of chunk capacity. *Journal of Experimental Child Psychology*, *104*, 252-265.

Gobet, F., & Clarkson, G. (2004). Chunks in expert memory: Evidence for the magical number four...or is it two? *Memory*, *12*, 732-747.

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C.-H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*, 236-243.

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: a new hypothesis. *Trends in Cognitive Sciences*, *11*, 236-242.

Hu, Y., & Ericsson, K. A. (2012). Memorization and recall of very long lists accounted for within the long-term working memory framework. *Cognitive Psychology*, *64*, 235-266.

Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of memory and language*, *30*, 685-701.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory. *Journal of Experimental Child Psychology*, *38*, 241-253.

Hutter, M. (2005). *Universal artificial intelligence: Sequential decisions based on algorithmic probability*. Springer.

Johnson, N. F. (1969). The effect of a difficult word on the transitional error probabilities within a sequence. *Journal of Verbal Learning & Verbal Behavior*, *8*, 518-523.

Jones, G., Gobet, F., & Pine, J. M. (2008). Computer simulations of developmental change: the contributions of working memory capacity and long-term knowledge. *Cognitive Science*, *32*, 1148-1176.

Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition*, *144*, 1-13.

Jones, G., & Macken, B. (2018). Long-term associative learning predicts verbal short-term memory performance. *Memory & cognition*, *46*(2), 216–229.

Karpicke, J. D., & Pisoni, D. B. (2004). Using immediate memory span. *Memory & cognition*, *32*(6), 956-964.

Kempe, V., Gauvrit, N., & Forsyth, D. (2015). Structure emerges faster during cultural transmission in children than in adults. *Cognition*, *136*, 247–254.

Kerkhoff, A. J., & Enquist, B. (2009). Multiplicative by nature: why logarithmic transformation is necessary in allometry. *Journal of Theoretical Biology*, *257*(3), 519–521.

Kibbe, M. M., & Feigenson, L. (2014a). Developmental origins of recoding and decoding in memory. *Cognitive psychology*, *75*, 55–79.

Kibbe, M. M., & Feigenson, L. (2014b). Developmental origins of recoding and decoding in memory. *Cognitive Psychology*, *75*, 55-79.

Kibbe, M. M., & Feigenson, L. (2016). Infants use temporal regularities to chunk objects in memory. *Cognition*, *146*, 251-263.

Klahr, D., Chase, W. G., & Lovelace, E. A. (1983). Structure and process in alphabetic retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 462.

Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission*, *1*, 1-7.

Li, M., & Vitányi, P. (1997). *An introduction to Kolmogorov complexity and its applications*. New York, NY: Springler Verlag.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279-281.

Mathy, F., Fartoukh, M., Gauvrit, N., & Guida, A. (2016). Developmental abilities to form chunks in immediate memory and its non-relationship to span development. *Frontiers in psychology*, *7*, 201.

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, *122*, 346-362.

Mathy, F., & Varré, J. S. (2013). Retention-error patterns in complex alphanumeric serial-recall tasks. *Memory*, *21*, 945-968.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Norris, D. G., Kalm, K., & Hall, J. (2019). Chunking and redintegration in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, Cognition*. doi: 10.1037/a0013355

Ottem, E. J., Lian, A., & Karlsen, P. J. (2007). Reasons for the growth of traditional memory span across age. *European Journal of Cognitive Psychology*(19), 233-270.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in piaget's developmental stages. *Acta psychologica*, *32*, 301-345.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge?. *Journal of Experimental Psychology: General*, *119*, 264-275.

Perruchet, P., Poulin-Charronnat, B., Tillmann, B., & Peereman, R. (2014). New evidence for chunk-based models in word segmentation. *Acta psychologica*, *149*, 1-8.

R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Reder, L., Liu, X., Keinath, A., & Popov, V. (2016). Building knowledge requires bricks, not sand: The critical role of familiar. *Psychonomic bulletin & review*, *23*, 271-277.

Roodenrys, S., Hulme, C., Alban, J., Ellis, A. W., & Brown, G. D. (1994). Effects of word frequency and age of acquisition on short-term memory span. *Memory & Cognition*, *22*(6), 695–701.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, *105*, 5975-5979.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 592-608.

Simmering, V. R., Miller, H. E., & Bohache, K. (2015). Different developmental trajectories across feature types support a dynamic field model of visual working memory development. *Attention, perception & psychophysics*, *77*, 1170-1188.

Soler-Toscano, F., Zenil, H., Delahaye, J.-P., & Gauvrit, N. (2013). Correspondence and independence of numerical evaluations of algorithmic-theoretic information measures. *Computability*, *2*, 125-140.

Swetz, F. J. (1987). *Capitalism and arithmetic: the new math of the 15th century, including the full text of the treviso arithmetic of 1478, translated by david eugene smith*. Open Court Publishing.

Terrace, H. S. (1987). Chunking by a pigeon in a serial learning task. *Nature*, *325*, 149-151.

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 37.

Tulving, E., & Patkau, J. E. (1962). Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Experimental Psychology*, *16*, 83-95.

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language*, *54*(1), 68-80.

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1256.

Widodo, R. N., Lim, H., & Atiquzzaman, M. (2017). A new content-defined chunking algorithm for data deduplication in cloud storage. *Future Generation Computer Systems*, *71*, 145–156.