Cognitive Development 76 (2025) 101641

Contents lists available at ScienceDirect

Cognitive Development

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

Children's use of semantic elaboration in immediate serial order memory

Luísa Superbia-Guimarães [a,*], Reese Lavers [b], Maya Steiger [c], Kellen Hendrix [d], Bret Glass [d], Nelson Cowan [d]

[a] School of Psychology, University of Leeds, University Rd, Woodhouse, Leeds LS2 9JT, United Kingdom
[b] Department of Psychology and Neuroscience, University of Colorado, Muenzinger Building, 1905 Colorado Ave, Boulder, CO 80309, United States
[c] Department of Psychology, Otterbein University, Psychology House, 55 North Grove St, Westerville, OH 43081, United States
[d] Department of Psychological Sciences, University of Missouri, 210 McAlester Hall, 320 S. 6th Street, Columbia, MO 65201, United States

ARTICLE INFO

Keywords:
Working memory
Semantic Elaboration
Serial Order
Child Development

ABSTRACT

Children's ability to use semantic elaboration to memorize novel associations of items greatly improves during the elementary school years. Yet, it is unclear whether this improvement contributes to the development of serial order memory. Here we investigated the role of semantic elaboration in the development of immediate memory for lists of object pictures. If the ability to use elaboration to encode new inter-item associations accounts for some of the developmental trend, then training children in an elaborative strategy should mitigate age differences in memory for serial order. We trained 7-year-olds (elaboration group) to create short stories to memorize lists of four items for serial order reconstruction and compared them to same-age and adult controls who received no strategy-specific instructions. All participants responded to questions on strategy use intermittently during the task and at its end. Although the story-making training markedly changed children's strategy reports, it did not improve their serial order memory. Moreover, although individual differences in the ability to create stories were strongly correlated with participants' serial order memory task performance, story quality and memory for serial order were only weakly correlated at the trial level. Results suggest that growing levels of use of semantic elaboration cannot account for the development of serial order memory in childhood. We discuss children's mnemonic strategies in serial order tasks in light of other studies.

A key aspect of development is the growing ability to memorize novel information over short delays, like being able to repeat novel verbal sequences. This ability is known as working memory (WM) – the capacity to maintain information at a heightened state of availability for ongoing processing (Cowan et al., 2021). WM greatly develops during elementary school years (Dempster, 1981; Huttenlocher & Burke, 1976), but the underlying driving mechanisms of such growth remain a topic of debate (e.g., Morra et al., 2024; Superbia-Guimarães & Cowan, 2023). In this study, we assess the notion that the ability to rapidly form new inter-item associations can partly account for the age-related increase in WM capacity for serial order (Cowan et al., 2018; Superbia-Guimarães and Cowan, 2023). This rapid encoding of associations can be achieved via semantic elaboration, i.e., the use of knowledge to enrich representations in WM. Yet, it is uncertain whether elaboration supports the development of serial order memory, which has mostly been investigated with verbal lists and under the lens of rehearsal mechanisms. For verbal lists, the development of rote rehearsal seems to play a

* Corresponding author.
E-mail address: L.SuperbiaGuimaraes@leeds.ac.uk (L. Superbia-Guimarães).

https://doi.org/10.1016/j.cogdev.2025.101641
Received 6 April 2025; Received in revised form 22 October 2025; Accepted 23 October 2025
Available online 8 November 2025

predominant role (e.g., Flavell et al., 1966; Ford & Silber, 1994; Hitch & Halliday, 1983) but, at least for visual materials, it is possible that a more semantically rich verbal encoding could play a larger role. Here we investigate whether children in early-elementary school can use semantic elaboration to memorize lists of object pictures for serial order reconstruction, and whether increasing proficiency in using this strategy could explain the developmental trend. We sought to answer these questions by training 7-year-olds in a story-making[1] strategy to encode lists of object pictures.

1. The development of serial order memory

Memory for verbal sequences steadily increases during middle childhood (Dempster, 1981), with recent large-scale studies showing curvilinear growth influenced by schooling and literacy (Ahmed et al., 2022; Kolinsky et al., 2020; Reynolds et al., 2022). Although this developmental increase has been attributed to various factors (see Superbia-Guimarães & Cowan, 2023 for a review), the dominant research approach has focused on verbal list tasks, emphasizing phonological coding and rote rehearsal as key mechanisms.

From this perspective, increasingly faster rehearsal rates with age would account for the development of the memory spans. Developmentally, children progress from spontaneously labelling items during kindergarten to using single-item rehearsal and then cumulative rehearsal during elementary school (Allik and Siegel, 1976; Bernbach, 1967; Elliott et al., 2021; Flavell et al., 1966; Hagen and Kingsley, 1968; Lehmann and Hasselhorn, 2007; Morey et al., 2024; Ornstein and Naus, 1978). Children start to use rehearsal and cumulative rehearsal around age 7, with these strategies becoming more automated until early adolescence (Allik and Siegel, 1976; Guttentag, 1984).

Although instructed rehearsal and training have been shown to improve children's serial order memory (Attout and Monnier, 2023; Miller et al., 2015), rehearsal-based accounts of development are insufficient to fully account for adult memory for verbal lists. While rehearsal requires less central attention compared to elaborative strategies (Naveh-Benjamin and Jonides, 1984; Thalmann et al., 2019), in adults, maintaining serial order in word lists strongly depends on attention and other strategies to preserve the relative relationships between item representations (Belletier et al., 2023; Guitard and Cowan, 2023; Guitard et al., 2022). Moreover, adults rely more on attention to maintain list lengths near the span level (Unsworth & Engle, 2007). Indeed, instructions to use single-item and cumulative rehearsal did not yield better serial recall in adults in two studies by Souza and Oberauer, 2018; Souza and Oberauer, 2020 – even though participants increased the frequency of rehearsal under the instructions. Therefore, rehearsal alone seems not to support memory for serial position in adults, let alone be a sufficient explanation for the steady capacity increase taking place during middle-childhood (see Jarrold & Hall, 2013 and Jarrold & Citroën, 2013, for discussions of limits in the evidence for a developmental role of rehearsal).

While much has been investigated about the use of rehearsal for serial recall with verbal lists, fewer studies on serial memory have used lists of object pictures. In a classic study by Flavell et al. (1966), 5- to 6-year-olds rarely verbalized to maintain a sequence of three object pictures, with a steady and constant increase in the frequency of verbalizations in the group of 7- to 8- and 10- to 11-year-olds. A more contemporary, large-scale replication of this study (Elliott et al., 2021) has shown that most children (75 %, as opposed to 10 % of children in the original Flavell study) in the younger group verbalized at least sometimes during the task, suggesting an earlier onset of verbal-based strategies to memorize pictorial lists. However, in another classic study by Hitch et al. (1989), the word length effect (a proxy of phonological rehearsal) emerged in children as young as 4 years of age for spoken materials, but considerably later – only in the 11-year-old group – for pictorially presented lists (Experiment 2).

For object pictures, semantic encoding might be a powerful way to maintain serial order, as semantics and imagery can reinforce each other and create stronger contextual bindings to serial positions. For example, in remembering that a picture of a teacup came before a picture of lamp, one could imagine a story in which a teacup was placed next to the lamp; whereas, in remembering that a picture of a lamp came before the teacup, one could imagine a story in which a lamp was turned on, revealing a previously hidden teacup. Below, we review the literature on the development of elaborative strategies in children and the role of semantics in maintaining serial order memory in adults. In anticipation, much of the developmental literature draws on studies using verbal stimuli but mainly word-pairs instead of lists of visual objects; adult evidence seems to suggest a limited role of semantics in memory for serial order but with studies typically manipulating list similarity rather than promoting the active use of elaboration by participants.

1.1. Semantic elaboration in working memory development

Semantic elaboration involves creating new associations between items based on their meaning, thereby enriching an item's representation in WM through existing long-term representations. These new associations can be achieved through mental imagery or sentence generation. The former involves creating a mental image of elements interacting (e.g., for the pair *grape-desk*, imagine a grape sitting on a desk), while the latter involves forming a sentence linking the elements (e.g., *"I put the grapes on the desk"*). Both strategies appear to follow the same developmental progression, yielding benefits on memory recall.

In terms of the developmental trajectory, children as young as 4–5 years old can use interactive imagery to memorize pairs when given explicit instructions describing physical interactions between objects. However, the benefits to their performance are limited and depend on the presence of the objects and the opportunity to manipulate them (Bender and Levin, 1976; Levin et al., 1975; Wolff et al.,

[1] The term "story" is typically used to refer to complex verbal productions composed of multiple connected sentences forming a coherent narrative. In this manuscript, however, we use it more loosely to describe children's simpler and shorter verbal productions about object pictures, as "story" is a familiar and age-appropriate concept for using in task instructions.

1972). By ages 5–6, children can be effectively trained to form their own interactive images but still rely on pictorial materials (e.g., picture cards or toys images) (Varley et al., 1974; Yuille and Catchpole, 1973). By ages 7–8, children can be effectively trained to use elaboration to memorize purely verbal materials, such as word pairs or prose passages (Begg and Anderson, 1976; Pressley and Levin, 1977b; Pressley, 1976). Despite this ability to be trained, the spontaneous use of elaboration to support memory has a late onset: it typically starts in late elementary school years, around age 11, and becomes more frequent until the end of adolescence (Beuhring and Kee, 1987; Kemler and Jusczyk, 1975; Rohwer and Bean, 1973). Studies by Levin et al. (1973) and McCabe et al. (1974) investigated whether the joint use of the two forms of elaboration, sentence generation and visual imagery, would promote even stronger benefits in recall in children with ages ranging from 4 to 11 years. Children were instructed and trained to use either one or both strategies together in order memorize pairs (picture cards, written words, or toys). Both studies found that the positive effects upon recall were redundant, that is, either of these strategies alone was as good as both together. This suggests that the recall advantage lies in the deeper semantic processing (Craik & Lockhart, 1972) promoted by these elaborative strategies, rather than in imagery or verbal production itself, and that the two strategies are two ways of creating the same mnemonic code.

The notion that semantic processing yields better memory recall dates back to Craik and Lockhart's (1972) work on levels of processing. Developmental researchers incorporated this idea into the hypothesis that children gradually replace rote learning with elaborative strategies that promote deeper semantic processing (Rohwer et al., 1977; Rohwer, 1974). Indeed, the hypothesis of a growing use of elaboration is supported by numerous studies, mostly using imagery and/or sentence generation instructions in paired-associate tasks applied to classroom contexts (e.g., vocabulary learning). For example, between ages 7 and 11 years, there is a developmental shift from attending to superficial, phonological aspects to deeper, semantic aspects of stimuli (Hasher & Clifton, 1974); older adolescents spontaneously use more elaborative strategies than younger ones (Pressley and Levin, 1977a; Pressley, 1982); finally, elaboration becomes more effective and less effortful in adolescence (Beuhring and Kee, 1987; Kee and Davies, 1988; Kee and Davies, 1991). Moreover, evidence from adult research suggests that the use of elaboration, either spontaneous or under instructions, enhances both immediate and delayed serial recall (Bailey et al., 2009; Bailey et al., 2011; Bartsch et al., 2024; Loaiza and Lavilla, 2021), suggesting that high-capacity individuals are more proficient in using elaborative strategies. Despite evidence linking elaboration to better recall in adults, the developmental literature primarily focuses on word-pair learning, making it difficult to conclude that this strategy directly contributes to the age-related improvement of serial order memory.

Attempts to determine if elaboration enhances serial order memory have yielded mixed results, even in adults. For example, Thomas et al. (2023) instructed participants to use interactive imagery to improve memory for word pairs. Although memory for the associated pairs improved, the imagery instructions did not affect memory for the order of associations (e.g., whether the pair was originally presented as "A-B" or "B-A"). On the other hand, the self-reported use of elaborative strategies (imagery and sentence generation) enhanced immediate serial recall compared to superficial strategies (rehearsal, refreshing, or reading the items) in a complex span task reported by Leproult et al. (2024). Besides analyzing results based on participants' trial-by-trial strategy reports, the authors also manipulated the distribution of free intervals to enable semantic elaboration and included a between-group manipulation of strategy instructions. Participants had superior immediate serial recall when the task schedule was conducive of elaboration (i.e., fewer longer intervals instead of more short intervals) and when participants were instructed to use this strategy. However, these benefits were inconsistent in delayed recall, indicating a need for further research on whether semantic elaboration is a helpful strategy to maintain serial order.

1.2. The influence of semantics upon serial order memory

As explained, semantic elaboration operates by strengthening associations between items via deeper semantic processing. There is a large body of research on the influence of semantic knowledge upon serial recall, though not directly tapping the controlled use of elaborative strategies. Rather, these studies traditionally manipulate the levels of semantic similarity between items in word lists, for example by comparing serial recall accuracy of semantically related lists (e.g., big cats: *leopard – tiger – jaguar – puma – lion*) to that of unrelated lists (e.g., *glasses – meadow – galley – ruler – shoelace*).

In verbal serial recall, participants must retrieve both the identity of words and their contextual positions within a sequence, that is, both item and order information are required. Item information refers to the phonological and semantic features that define a word's identity and distinguish it from others. In contrast, serial order information relates to the temporal sequence of words within a list (Lee and Estes, 1981; Nairne and Kelley, 2004). Item retrieval is thought to rely on the temporary activation of consolidated concepts in long-term memory (Cowan, 1999), whereas remembering serial positions is thought to rely on temporary item-context associations taking place during the trial (Burgess and Hitch, 1999; Burgess and Hitch, 2006; we refer the reader to Majerus (2019) for a comprehensive review of the distinction between item and order memory). The latter would occur in WM because the positional arrangement of items in a sequence in memory studies is typically novel and arbitrary. For example, in the sequence *spoon – grass – ship – phone – floor*, each item represents a familiar concept already present in the knowledge base but there is no intrinsic connection between the items and their specific positions within the sequence. Memory for serial order is assessed through the analysis of transposition errors (e.g., swapping items across serial positions, like reporting "ACBD" instead of "ABCD"), which reveal confusions between contextual bindings of list items and their serial positions and incomplete item-to-item associations.

Unlike other types of similarity (e.g., phonological), semantic similarity does not consistently increase transposition errors in serial recall tasks (Kowialiewski et al., 2023; Neale and Tehan, 2007; Neath et al., 2022; Poirier and Saint-Aubin, 1995; Saint-Aubin and Poirier, 1999; but see also Saint-Aubin et al., 2005; Tse, 2010; Tse et al., 2011). In contrast, research consistently shows that item recall (retrieving the correct item, regardless of serial position) improves for semantically similar lists (Kowialiewski and Majerus, 2020; Neale and Tehan, 2007; Poirier and Saint-Aubin, 1995; Saint-Aubin and Poirier, 1999; Tse et al., 2011; Tse, 2010).

Interestingly, two recent studies by Kowialiewski et al., 2022; Kowialiewski et al., 2024 showed evidence that, at least in healthy adults and for certain list arrangements, semantic information can scaffold the maintenance of serial order and even improve it, contrary to the prevailing findings. Grouping semantically related items enhanced both item and order memory, especially when the semantic categories were clustered (e.g., *leaf, tree, branch – sky, cloud, rain*) rather than interleaved (e.g., ***leaf*** *– sky –* ***tree*** *– cloud –* ***branch*** *– rain*). Moreover, transposition errors tended to occur within the semantic groups, with items migrating to positions occupied by items of the same semantic category (Kowialiewski et al., 2024). These results challenge the widespread view that order information in WM is impervious to the influence of semantic knowledge.

In children, the facilitative effect of semantic similarity on item retrieval has long been established (e.g., McFarland & Kellas, 1975). However, comparatively fewer studies have examined how semantic factors influence memory for serial order. One of the few exceptions is a study by Monnier and Bonthoux (2011), which investigated the serial recall performance of 5- and 9-year-old children using semantically similar and dissimilar word lists. The researchers also manipulated the taxonomic organization of the semantic categories. The semantically similar lists were either context-dependent, comprising items associated with a specific event schema (e.g., winter clothing: *coat – sweater – hat – trousers*) – or context-independent, comprising items from a broader, supra-ordinate category not tied to a specific context (e.g., clothing: *t-shirt – dress – sock – pajamas*). While children as young as two years can understand contextually grounded categories, the ability to grasp abstract supra-categories typically emerges around the age of nine (Nelson, 1985). Results showed superior serial recall of context-dependent lists from age 5 years, whereas improved serial recall of context-independent lists emerged only at age 9 years. These findings indicate that semantic knowledge can facilitates serial recall even in early childhood. Moreover, the nature of this facilitation appears to evolve with age, reflecting the developmental trajectory of conceptual and linguistic abilities. However, semantic organization did not specifically reduce the number of order errors, suggesting that facilitation of order might at least be dependent on a more extensive or detailed use of semantics.

A different result was obtained, however, by Rosselet-Jordan et al. (2022), who manipulated both the semantic similarity between list items for serial recall and the attentional constraints during the list encoding, in a complex-span task design. The study showed a strong beneficial effect of semantic similarity both in children (11-year-olds in the first experiment; 9-year-olds in the second) and young adults. However, the age groups differed in terms of transposition errors, which were less affected by relatedness for adults than for children. Related lists yielded more transposition errors. These results show that semantic information can influence order recall and suggest that the influence of long-term knowledge upon WM might even be greater in children than in adults, presumably because adults are better able to rely on phonological rehearsal rather than semantics. Attention demand did not matter overall, but high attentional demand did increase the number of transpositions as well as omissions, suggesting the use of an attention-related strategy at least some of the time. Overall, whereas semantic relatedness caused order confusion in this study, new semantic encoding of unrelated series of objects might be beneficial.

In sum, semantic information has a powerful influence upon item memory, but it affects memory for serial order in only specific conditions, with few studies showing a benefit in children's serial recall. Less is known about its effectiveness to scaffold serial order memory, whether children spontaneously use it for this aim, and whether it could partly account for the developmental trend in the verbal span. Traditionally, the efficiency of semantic elaboration and the development of serial order memory have been investigated separately in the literature, using paired associates versus serial recall or serial order reconstruction, respectively. These have been done mainly with verbal materials. In this study, we investigate the relationship between the spontaneous and trained use of semantic elaboration and the development of serial order memory for object pictures. We also collect subjective reports of strategies to be mindful that multiple strategies can co-exist.

2. Study rationale & hypotheses

Here we investigated whether the ability to use semantic elaboration contributes to age differences in WM capacity for serial order between 7-year-olds and young adults. A group of 7-year-olds s was trained to use an elaborative strategy – story-making – to associate the items at encoding, while same-age and adult controls were free to implement their preferred strategy. We purposefully chose serial order reconstruction because this type of task does not require item retrieval, thereby indexing the effects of processing specifically on order information. Moreover, we purposefully chose to use object pictures because of inconsistent findings that children aged 7 years prefer phonological-based to more semantic- or imagery-based strategies to remember lists the pictorial format (Elliott et al., 2021; Flavell et al., 1966; Hitch et al., 1989).

Our story-making instructions require children to recode items from phonological to semantic representations – if not already done by them – and then integrate the items via associations. If children's natural inclination is to retain items phonologically, then our story instructions require them to create semantic representations to be combined into a short fictional narrative, thereby enriching encoding. If, on the other hand, they are already using semantic- or imagery-based representations along with phonological coding, the effort to follow the story-making instructions and use elaboration should be lower. We addressed three classes of hypotheses:

H1. Adults and children spontaneously use elaboration to support serial order memory for series of pictures.

We were interested in whether child and adult controls (i.e., those not instructed to use a particular strategy) would spontaneously implement semantic elaboration in the context of a serial order reconstruction task. We did not have specific predictions for this hypothesis. On one hand, adults should be able to construct sophisticated stories that would assist in the encoding of order information for a series of pictured objects but, on the other hand, they might be more likely to rely on a verbal rehearsal strategy (e.g., Belletier et al., 2023). It is unclear how readily children would come up with satisfactory semantic encoding for order retention. Previous studies have primarily focused on paired-associate and vocabulary learning tasks, which are forms of associative learning where the order of items is not critical. For example, in a pair like "AB," either item should cue the recall of the other, making the sequence irrelevant. In such contexts, access to item meaning serves as a powerful cue for learning (e.g., *tea – lamp*, where one might imagine a cup of tea beside a table lamp). In contrast, our study employs a serial order reconstruction task, where item memory and semantic access are less central, as all items are presented again during the response phase. Instead, participants must reconstruct the relative positions of pictures within a sequence. This task demands the simultaneous maintenance of chained serial positions (i.e., identifying which item belongs in the first, second, third position, etc.), but it relies less on semantic memory. It is, therefore, unclear whether children and adults will spontaneously apply elaboration to memorize the relative serial positions of items in a sequence.

H2. There is developmental change in the use of elaboration.

If semantic elaboration is indeed spontaneously used to support serial order memory, does it change with age? Based on the literature above, we predicted that adults should use elaboration more often than children. If *H2* is true, we should observe a higher number of reports of using elaboration in adults compared to control children.

H3a-H3b. The story-making training will (3a) increase the amount of elaboration and (3b) improve serial order reconstruction.

If the training is successful, we expected to see an increase in the frequency of elaboration in comparison to control children. Moreover, if the efficiency of elaboration accounts for some of the developmental trend in the verbal span, then the story-making strategy should reduce the age gap to adults and improve performance as compared to same-age controls, who received no specific strategy instruction. The basis for our hypothesis (H3b) is twofold. First, elaboration creates more meaningful and interconnected memory traces of the items, which can facilitate the formation of contextual cues that aid retrieval. Second, story-making inherently possesses a temporal structure, as narratives unfold sequentially. If this temporal sequencing matches the order of items to be remembered, the story can provide a scaffold to the serial positions of items. By mapping each item onto a specific point within the story's timeline, individuals can encode the temporal context more effectively. Alternatively, a successful increase in the use of elaboration (confirmation of H3a) but with no benefits for order reconstruction performance (rejection of H3b) would lend support to the literature showing a very limited role of semantic knowledge in serial order within WM.

We also explored the relationship between the quality of the stories created by children and memory for serial order (H3b). According to this key hypothesis, the quality of the stories should correlate with responses given in the respective trials in which they were created. Moreover, individual differences in the ability to create stories should correlate with overall performance in the task. To our knowledge, our study is the first to specifically test the effectiveness of a story-making strategy to investigate the development of serial order memory in children.

3. Method

3.1. Participants

The sample size rationale was based on a power analysis indicating that at least 82 participants were needed to achieve a power of 0.8 to detect a medium-to-large effect size (Cohen's $f = 0.35$) for group differences in a three-group design (Zhang & Yuan, 2018; https://webpower.psychstat.org/wiki/). The target sample size was then rounded up to 90 to allow for 30 participants per group. We collected data from more participants in anticipation of dropouts and technical problems associated with live online testing.

Eighty-two children (aged 6–7 years), and 37 adults (aged 18–30 years) took part in our study. After exclusions, our sample included 32 children in the elaboration group (M_{age} = 7 years and 2 months, SD = 6 months, 19 girls), 31 children in the control group

Table 1
Sample characterization by analysis.

Serial Order Reconstruction			
	Group		
	Elaboration children	Control children	Adults
	$N = 32$	$N = 31$	$N = 29$
Mean Age	7 y 2mo (6mo)	7 y 1mo (6mo)	24 y 2mo (3 y 5mo)
Mean Digit Span	5.38 (1.04)	5.13 (1.02)	7.24 (1.27)
Verbal Responses in Report Trials, Strategy Questionnaire			
	Group		
	Elaboration children	Control children	Adults
	$N = 31$	$N = 30$	$N = 27$
Mean Age	7 y 2mo (6mo)	7 y 1mo (6mo)	24 y 2 mo (3 y 3 mo)
Mean Digit Span	5.39 (1.04)	5.17 (1.04)	7.30 (1.21)
Within-Subject Comparison of Story Quality in Correct vs. Incorrect Serial Reconstruction Trials			
	Group		
	Elaboration children	Control children	Adults
	$N = 28$	$N = 0$	$N = 0$
Mean Age	7 y 2mo (7mo)		
Mean Digit Span	5.22 (0.92)		

Note. Values between parentheses represent standard deviations; y: years; mo: months.

($M_{age} = 7$ years and 1 month, $SD = 6$ months, 18 girls), and 29 adults ($M_{age} = 24$ years and 2 months, $SD = 3$ years and 5 months, 6 women). The two child groups were equated in age ($BF_{10} = 0.36$, $p = 0.38$). Reasons for excluding participants were the following: participant performance below 3 standard deviations of the group mean (3 adults, 5 children); participant not performing the task according to instructions (e.g., relying on parents, not understanding instructions, 4 children); participant dropout (1 adult, 2 children); experimental error (e.g., failure in internet connection, video call settings; 4 adults, 8 children). Some participants in the final sample were excluded from some analyses for reasons to be explained when each analysis is introduced. Accordingly, sample sizes and mean ages differ slightly between analyses. Table 1 details the number of participants, the mean age, and the mean digit span for each group in every analysis.

Written consent was obtained from all adult participants and children's parents. Participants received 15 dollars as compensation, and children who took part in the laboratory also received stickers and a book. All procedures complied with ethical standards and were approved by the Institutional Review Board at the University of Missouri (protocol #99–04-095).

3.2. *Materials and stimuli*

Materials are available on OSF (https://osf.io/nja4c/). The stimuli pool consisted of a set of 48 unique lists of four object pictures totaling 192 objects (original stimuli from Brady et al., 2008; Brady et al., 2013). Some of the original stimuli were replaced by images that are more familiar to American children, e.g., a honey pot was replaced by a commercial bottle of honey commonly found in U.S. supermarkets. Each object picture measured about 4 cm^2. Lists were created by randomly sorting four objects from the pool, with no repetitions. Another 22 object pictures were used as foils during the familiarization phase of the procedure (described below); these were never used in the lists for memorization. All objects corresponded to words with an age of acquisition below 5 years (average age of acquisition = 4.9 years) (Kuperman et al., 2012).

We created audio files with the corresponding spoken word (in a male voice) for each object picture in the stimuli pool, and another 22 files with non-matching words for the foil objects used in familiarization phase, totaling 236 audio files. See Appendix A for the full set of objects lists, the respective words, foil object pictures, and non-matching words used in the task. For the digit-span task (described below), we created 24 spoken lists of 2–9 digits (in a female voice), with three lists per list length. Lists were pseudorandom, with no repetitions of digits within a list and avoiding obvious digit sequences (e.g., 1–2-3). The experiment was programmed and implemented using Psytoolkit (v. 3.4.4., Stoet, 2010; Stoet, 2017).

3.3. *General procedure*

Participants were tested individually and the experimental sessions were audio recorded. Sessions began with a sound volume adjustment procedure, followed by a digit span task, and the main experiment – a serial order reconstruction task. We manipulated the use of semantic elaboration with a between-group manipulation of instructions and training. Participants answered a questionnaire on the use of WM maintenance strategies at the end of the experiment.

Experimental sessions were conducted either in the lab – with 15 children in the final sample (7 in the control condition and 8 in the elaboration condition) – or online via Zoom, remotely connected with an experimenter. The online sample included all adult participants and 48 children (24 in each condition). There were no significant differences in age ($BF_{10} = 0.78$, $p = 0.12$) or WM performance ($BF_{10} < 1$ and $p > 0.1$ in all measures) between children who participated online and those tested in the lab, as shown by independent samples *t*-tests. In online sessions, the program ran on the participant's own computer, in either Chrome or Firefox web browsers. Screen and sound sharing on Zoom were enabled during the online sessions, and the experimenter also had remote control over the participant's mouse and keyboard. Experimental sessions took about 60 min.

3.3.1. *Sound adjustment procedure*

At the beginning of the experimental session, participants and/or their parents were asked to adjust the volume to a comfort level and microphone and sound sharing settings were tested in online sessions. In-lab participants wore over-ear headphones, but the use of headphones was not guaranteed for all online participants. A sound message (*"Hey there!"*) with a female voice was played upon a mouse click to enable the sound adjustment. The experimenter prompted the program to the next procedure phase once the participant was satisfied with the volume level.

3.3.2. *Digit span procedure*

Digit lists were aurally presented. Lists started with two digits and increased to up to nine digits, with three lists per length. Each digit was presented for about 400 ms (with slight differences between digits), and the onset-to-onset time (i.e., the interval between the beginning of the presentation of each digit and the beginning of the next one) was fixed to 800 ms, holding the processing time per digit constant and yielding an interstimulus interval (ISI) of about 400 ms. Participants were asked to verbally repeat the lists, and the procedure was interrupted after three consecutive recall errors of the same list length. The release of the digit lists was controlled by the experimenter via keyboard presses.

3.3.3. *Object familiarization procedure*

Each block of trials of the main task (described next) was preceded by a familiarization phase to ensure that participants knew the vocabulary to describe the object pictures. Each object picture was centrally presented for 2500 ms, simultaneous to the aural presentation of a descriptive message (e.g., the picture of an onion with the audio *"This is an onion"*), with an ISI of 50 ms. To maintain participant's engagement with the task, they were instructed to press the space bar in case of mismatch between an object and its description, and 10 foil object pictures were presented with a mismatching audio description (e.g., a tiger accompanied by the audio "*This is a dog*"). These foils were never used in the memory task. Upon a keyboard press to a mismatched presentation of object picture and audio description, a correction message was played by the computer (e.g., *"Sorry, this is a tiger"* in the above example). The device of the computer offering feedback in which it apologetically acknowledges its own former mistake has been used before (Hentschel et al., 2023) and was shown to reinforce learning of the correct information. This result adds confidence that our procedure to assess whether children knew the names for the objects would reinforce the correct names and not instill incorrect names when the catch trials at first offered the wrong information. The cover story for children was that the computer was learning human language and practicing new words, and needed to be corrected sometimes. Object pictures were presented in a random order during the familiarization phase.

3.3.4. Serial order reconstruction task

The main experiment consisted of a serial order reconstruction task with a between-group manipulation of the instructions and training procedure. We trained children in the experimental group to create short stories during the encoding of the object sequences. Child and adult controls were instructed to think about ways to memorize the sequences, but no specific strategy training or instruction was given. All instructions described in the next sections were written on the screen and read aloud by the experimenter.

3.3.4.1. Elaboration group. Fig. 1 (panel B) represents the order of events in a trial. On each trial, a list of four object pictures was sequentially presented on the screen and, after a brief delay, participants were asked to reconstruct the list. Each trial started with the central presentation of the word "*Prepare*" for 2000 ms, followed by a central fixation cross for 1000 ms. Then, the four object pictures were centrally presented, each for 1000 ms and with an ISI of 1000 ms. Participants were trained and instructed to create a short story linking the items during the encoding of lists. A retention interval of 3000 ms followed the offset of the last item in the list. Then, a response screen displayed the four list items horizontally aligned in random order, along with four response slots representing each serial position, in a left-to-right orientation. Participants had to reconstruct the list sequence by using the mouse to click on the response items in the correct order, thereby placing them in the slots corresponding to the serial positions. For example, the first clicked item would be placed in the leftmost square, so on. Once a response was made, the clicked item became unavailable for the upcoming responses by disappearing from the response options. Response corrections were not possible. Participants received feedback on their answers on every trial. After feedback, children were asked to report their stories in 25 % of the trials, henceforth "report trials". The instruction "*Tell me the story you created*" was displayed on the screen and read aloud by the experimenter. Stories were audio recorded and later transcribed. The release of trials was self-paced except in report trials, when the experimenter controlled it via a keypress unbeknownst to participants. This was to prevent children from skipping to the next trial before finishing their reports.

The experiment had a total of 40 trials, divided into two blocks. Each of the two blocks had 20 trials, presented in the same fixed order to every participant. Report trials were pseudo-randomly spaced during the task, the only constraint being that there were 5 of them in each block. Each block of trials, including the practice block, was preceded by a stimuli familiarization phase (described above). Fig. 1, panel A represents the stages in the experimental procedure and panel B represents a trial schema.

3.3.4.2. Story-making instructions and training. Seven practice trials preceded the experimental task. The timing of practice trials was identical to that of trials, except when stated otherwise. Participants in the elaboration group were instructed to *"memorize the order of pictures that will appear on the screen and make up short stories with them"*. In the first practice trial, a list (*toaster – onion – zebra – frog*) was sequentially presented and participants were given a head start to reconstruct the order (*"Can you put these in the order in which you saw them? The toaster was first."*). Participants received feedback on their answer and then a message was presented on the screen: *"Now let's make up a story with the pictures. Let's try?"*. In the second practice trial, the list items (*helmet – sneaker – cracker – door*) were simultaneously displayed side-by-side in a left-to-right orientation; an example story was written on the screen and read aloud by the experimenter (*"The helmet fell and hit the sneaker, then a furious cracker broke through the door"*).[2] Note that the example story preserves the serial order of items in the list. Then, the same list was sequentially presented, and children were asked to reconstruct the sequence. In the third practice trial, another list was simultaneously presented, and participants were asked to create a story by themselves (*"Now try to make up a story by yourself. It can be any story you can come up with."*); unlimited time was given, and the list remained on the screen. The response screen was released by the experimenter after the story was reported, and participants had to reconstruct the sequence. In the fourth practice trial, a list was sequentially presented, and participants created the story by themselves; unlimited time was given but the screen remained blank. The response screen was released by the experimenter after the story was reported. In the fifth practice trial, a list was sequentially presented, and children were instructed to create the story in silence; unlimited time was given, and the screen remained blank. The response screen was released after the story was reported. In the sixth practice trial, a list was sequentially presented, and children were instructed to create a story in silence and wait until the end of the trial, after response, to report it. In the seventh and final practice trial, a list was sequentially presented; children were instructed to create a story in silence but not report it immediately. They were instructed to report the story only when asked by the experimenter, who explained that this would happen occasionally during the task. Note that the training progressed in difficulty, from allowing children to create and report

[2] The example story contains a comic, implausible description of events to encourage children to report *any* story, regardless of adequacy or veracity. This decision was made after pilot testing because children were focusing too much on creating reasonable, veridical stories.

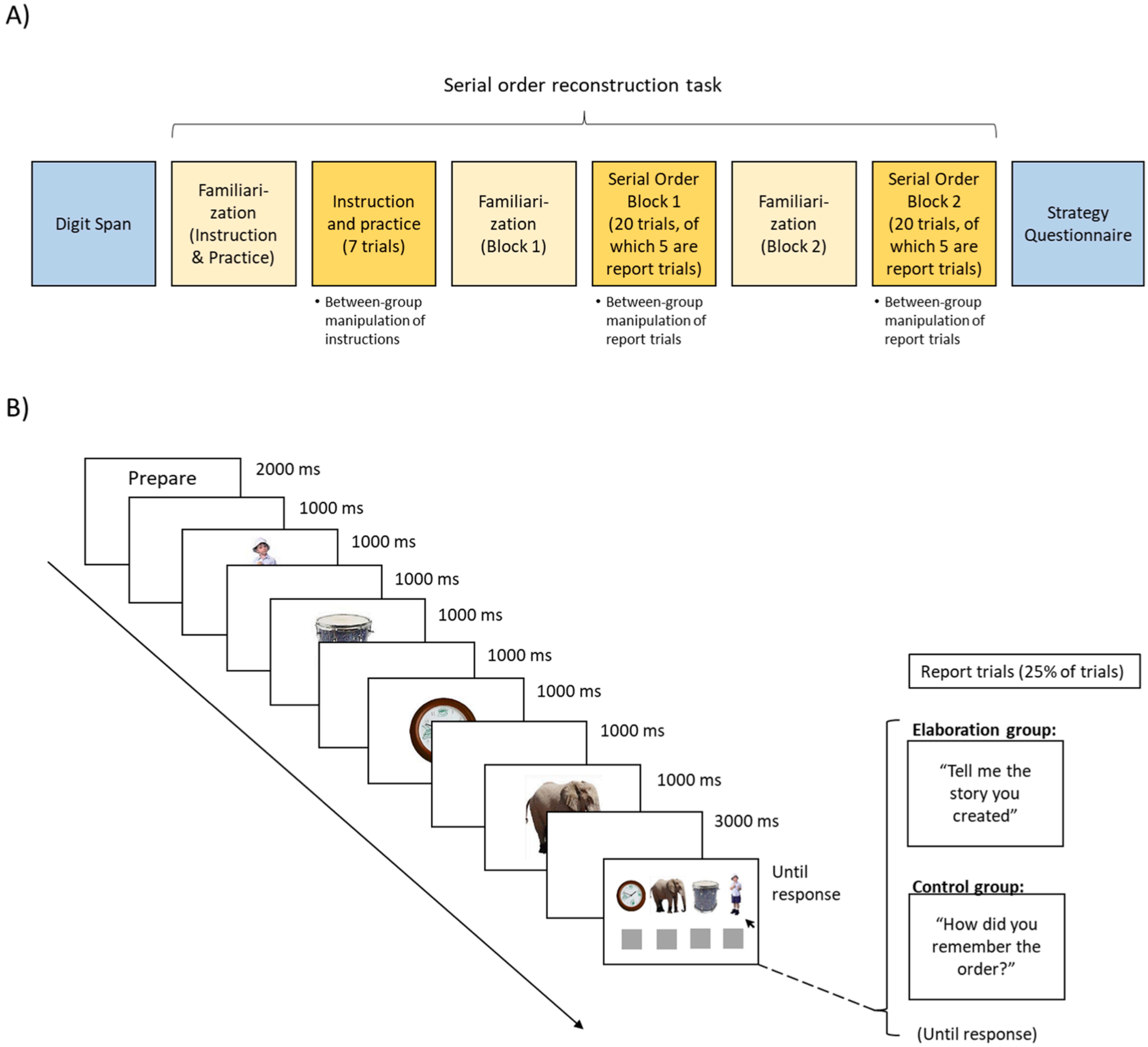

Fig. 1. Stages of the Experimental Procedure (A) and Trial Schema of the Serial Order Reconstruction Task (B). *Note.* Panel (A) represents the stages of our experimental procedure. Participants were given a 5-minute break between blocks 1 and 2. Panel (B) shows a trial schema in the serial order reconstruction task, with the sequence *boy – drum – clock – elephant*. In the response screen, the grey squares represent response slots for each serial position, in a left-to-right orientation. The response items were placed in the slots upon a mouse click, in the order they were clicked.

a story while looking at list items on the screen (trials 2 and 3), then creating a story in a timely manner during a sequential presentation (trials 4 and 5), and finally creating a story completely in silence (trials 6 and 7). Practice trials 2–7 were repeated up to three times upon error, in which case the experimenter would correct participants and, if necessary, help them to reformulate their story in a way that preserved serial order.

3.3.4.3. Control groups. Control children and adults underwent the same procedure as described above, except for report trials, instructions, and practice trials. In report trials, the prompt to report a story was replaced by the question "*How did you remember the order?*", and participants were asked to describe their strategy in that trial.

3.3.4.4. Control instructions and training. Control participants were instructed to "*Remember the order of the pictures*" but were not trained in the story-making strategy. To control for the interaction with the experimenter, control instructions mimicked the schedule of the elaboration training, but without directing participants to any specific strategy. Instead, participants were instructed to "*Memorize the order of pictures that will appear on the screen*" and to "*Think about how you can memorize the order of the pictures*". The prompt to report the story in practice trials was replaced by "*Could you tell me how you remembered the order of the pictures?*". All other events during practice trials were identical to the elaboration condition. Appendix B provides a detailed comparison of the instructions given to the elaboration and control groups.

3.3.5. Strategy questionnaire

Participants answered to a questionnaire about the use of maintenance strategies after the experiment (adapted from Belletier et al., 2023). The questionnaire had an open-ended question ("*How did you remember the sequences of pictures?*") followed by seven closed (Yes/No) questions on the use of elaboration, verbal rehearsal, visualization, visual imagery, acoustic imagery, temporal clustering, and memory reduction (Table C1, Appendix C). Questions were read aloud by the experimenter to participants. We originally intended to ask all questions to all participants; however, when transcribing the session recordings, we realized that the closed question on the use of elaboration was not consistently asked to all children in the elaboration group. For this reason, we will refrain from reporting and interpreting results from the closed items in our questionnaire and present only results from the open-ended question.

3.4. Data generation

3.4.1. Serial order reconstruction measures

Our main variables of interest are the percentage of lists perfectly reconstructed, henceforth "perfect lists", and the number of items correctly placed within a list, henceforth "number of correct items".

3.4.1.1. Perfect lists. To calculate the percentage of perfect lists, we scored one point for each correctly reconstructed list and divided this sum by the total number of trials in the experiment.

3.4.1.2. Number of correct items. One point was scored to each item placed in the correct serial position in a list (e.g., if the list *floor – cat – mug – clock* was reconstructed as *floor – mug – cat – clock*, 2 points were given). Those measures were calculated per participant and averaged across groups.

3.4.2. Story characterization

We characterized stories created by children in the elaboration group according to (1) their completeness and (2) by applying the edit distance scoring method (Demerau-Levenshtein distance, scoring algorithm by Gonthier, 2023).

3.4.2.1. Story completeness. This refers to the number of elements from the current trial list that were included in the story, irrespective of the order that they were mentioned. According to this criterion, stories were classed as "complete" (all four list items included), "partial" (at least two, but not all items included), "single-item" (only one list item included), "unrelated" (a story was reported, but none of the current list items were included), and "no-story" (no story was reported in the trial).

3.4.2.2. Edit distance. The edit distance measure corresponds to the number of changes required to make a string match another string, and thus it encompasses the order that list items were mentioned in the stories. This scoring method bears the advantage of not overly penalizing transposition errors.[3] For example, for a given list "ABCD" that was recalled as "BCD" (one omission error), the edit distance is 1 (one change required: add item A). If "ABCD" is recalled as "BACD" (one transposition error), the edit distance is also 1 (one change required: move A by one position); if it is recalled as "ADBC" (three transposition errors), the edit distance is 2 (two changes required: remove D, insert D at the end); if it is recalled "AD" (two omissions), the edit distance is 2 (add item B, add item C). Shorter edit distances, thus, correspond to better stories both in terms of preserving the items and their relative order within the story. This scoring method has an advantage over traditional serial position scoring in giving credit for partial sequences in correct order. For example, if the list is represented as "ABCD," then the response "BCDA" receives partial credit for order despite all responses being in the incorrect serial positions.

To apply the edit distance scoring to stories, we transformed each story into a short "ABCD" format that included only words corresponding to list items, ignoring the extra words used to connect them. For example, for the list *umbrella-purse-jug-guitar* (i.e., ABCD), the story "*When I went to the store, I had a[n] umbrella, a purse, and I was singing music on my guitar*" was converted into *umbrella – purse – guitar* (i.e., ABD), which yielded the edit distance of 1 (i.e., include item C, *jug*). Order reconstruction responses were also scored with the edit distance system to allow comparison with the stories created in the corresponding trials.

[3] It has been argued that the edit distance scoring produces a bias against children with weaker WM who produce shorter stories, e.g., including only two items from the original sequence. An example from our data is the following: Given the trial *truck - teddy bear - teapot – baseball* (ABCD), the child reported the story "*The bear (teddy bear) went with the teapot to go to the store*", which was then converted into *teddy bear – teapot* (BC). The edit distance here is 2, corresponding to the inclusion of items A (*truck*) and D (*baseball*) in order to match the encoded sequence (ABCD). We sustain that the edit distance method cannot bias the results against children with worse performance. On the contrary, it even benefits them, given that participants are not penalized for intrusions, nor penalized for omissions in initial serial positions. For example, with a simpler and widely used scoring method for serial order, the story exemplified above ("*The bear went with the teapot to go to the store*" → BC) would have been scored as 0: because A was omitted, B and C automatically are scored as zero because they were recalled in the wrong serial positions – even though memory for item was perfectly preserved in the story. Moreover, other scoring methods implement penalties for intrusions, for example, the incorrect item *store* (not originally encoded) would have been given a −1. We decided to use the edit distance in order to cope with these limitations of other scoring methods that we judged too harsh.

3.4.2.3. Number of connective words within stories. We computed the frequency with which each participant in the elaboration group used the words "and" and "then" in their stories, on a trial-by-trial basis.

3.4.3. Coding of answers in the strategy questionnaire

Answers to the open-ended question were categorized by two independent raters to determine the use of elaboration or story-making, as well as other reported strategies. Table C2 in Appendix C reports the detailed rules, categories, and outcomes. The inter-rater agreement was 96 % for the elaboration criteria, with only three instances of disagreement; these were resolved through discussions between authors L.S.G. and N.C.[4]

3.5. Data analysis

The equality in age and digit spans between children in the control and elaboration groups was verified with Bayesian *t*-tests for independent samples. Group means for the perfect lists and the number of correct items were compared with a one-way Bayesian ANOVA – this was our main analysis of interest. To assess participants' strategy use, we compared the frequency of strategy mentions reported in the questionnaire across groups by using Bayesian contingency tables and Chi-squared tests. Finally, we assessed whether individual differences in creating stories are associated to individual differences in WM performance byrunning Bayesian correlations 1) between the average story edit distances, the number of correct items, the percentage of perfect list reconstruction, and digit spans; and 2) between the number of connective words used in the stories and the number of correct items in the corresponding trials.

We opted for Bayesian hypothesis testing because it allows quantifying evidence in favor of the null hypothesis, which is not possible with traditional frequentist *p*-values that can only reject or fail to reject the null, but cannot provide support for it (Wagenmakers et al., 2017). We adopted the criteria of a Bayes factor (BF_{10}) > 3 as evidence for the alternative hypothesis and $BF_{10} < 0.3$ as evidence for the null (Jeon & De Boeck, 2017). By editorial request, we also report *p*-values for the convenience of readers who prefer traditional metrics. Additionally, we employed generalized linear mixed models to address power limitations in the pairwise comparisons between groups.

4. Results

First, we show that there were no baseline differences between the child groups in terms of the digit span. Then we report on children and adults' use of elaboration, relying on both intermittent reports collected during the memory procedure and retrospective strategy questionnaire responses. After the use of elaboration in each group has been described, we examine performance in order reconstruction to assess the role elaboration may have played. Finally, we present an analysis of the quality of reported stories in the elaboration group and their relation to serial order reconstruction performance. The results of the supplementary generalized linear mixed models (available on OSF: https://osf.io/nja4c/) yielded same the conclusions as the Bayesian ANOVAs reported below.

4.1. Digit spans

There were no baseline differences between the two child groups in their baseline WM ability, as shown by equivalent digit spans ($M = 5.13$, $SD = 1.02$ for control children and $M = 5.38$, $SD = 1.04$ for elaboration children, $BF_{10} = 0.37$, $p = .35$, Cohen's $d = -0.24$). In adults, the mean digit span was 7.24 ($SD = 1.27$).

4.2. Verbal responses in report trials: Compliance to story-making in the elaboration group and spontaneous elaboration in the control group

We assessed children's compliance with the story-making strategy and the spontaneous use of elaboration in control groups by looking into their verbal responses in report trials. Due to audio recordings of bad quality or with recording errors, the following data excludes 1 child in the elaboration group, 1 control child, and 2 adults, resulting in samples of 31 ($M_{age} = 7$ years and 2 months, $SD =$ six months), 30 ($M_{age} = 7$ years and 1 month, $SD = 6$ months), and 27 ($M_{age} = 24$ years and 2 months, $SD = 3$ years and 3 months) participants, respectively. Since each participant completed 10 report trials, the corresponding number of trials entered per group was 310, 300, and 270 respectively in the analysis below.

4.2.1. Elaboration children

All children in the elaboration group reported a story in at least one report trial, with most children (28 children, 90.3 % of participants in this group) being able to produce a complete story (i.e., including all four items) at least once during the task. Most children in the elaboration group (25 out of 31, 80.6 % of participants in this group) reported stories including at least two list items in at least half of their report trials. Only 3 children did not report any complete story during the task; yet they were able to report partial stories, i.e., including at least two items. The mean number of list items included in the stories, regardless of their order of mention

[4] In these three cases, we decided for scoring elaboration to one child who spontaneously mentioned using stories when asked the directed question on acoustic imagery; one child who said *"First I had to repeat what the Tigress said and then I had to help the computer… And then I did the stories";* and one child who said *"I just kinda looked at the pictures and thought about what I would do with some"*, by considering *"what I would do with some"* as elaboration.

within the narrative, was 2.39 (*SD* = 0.83). There were no cases of children in the elaboration group who never reported a story, either partial or complete. Together, these results suggest that participants in the elaboration group mainly complied with the story-making strategy during the task.

Examples of complete stories created by children are *"The butterfly went into the nest and found a bottle with milk in it"*, for the list *butterfly – nest – bottle – milk*, and *"I was eating pizza for dinner while watching a crocodile swimming in a lake and then I heard a bell for the pasta"*, for the list *pizza – crocodile – bell – pasta*. Examples of partial stories are *"The scissors cut the penguin then the penguin had only one arm and then the one arm he had to eat so he couldn't eat"*, for the list *bench – penguin – scissors – doll*, and *"When I went to the store, I had a [n] umbrella, a purse, and I was singing music on my guitar"*, for the list *umbrella – purse – jug – guitar*.

4.2.2. Control children

In stark contrast to the elaboration group, only 3 control children (10 % of participants in this group) described using stories or some sort of elaborative strategy when asked how they remembered the list during report trials. Across these three children, elaboration was observed in 17 out of 30 trials, i.e., 56 % of their report trials.

4.2.3. Control adults

Of all adults, 9 (30 % of participants in this group) mentioned using stories or other elaborative strategies (interactive imagery, sentence generation, self-referencing actions involving the objects) during report trials. Across these 9 participants, such strategies were reported in 69 out of their 90 trials, representing 76 % of their trials.

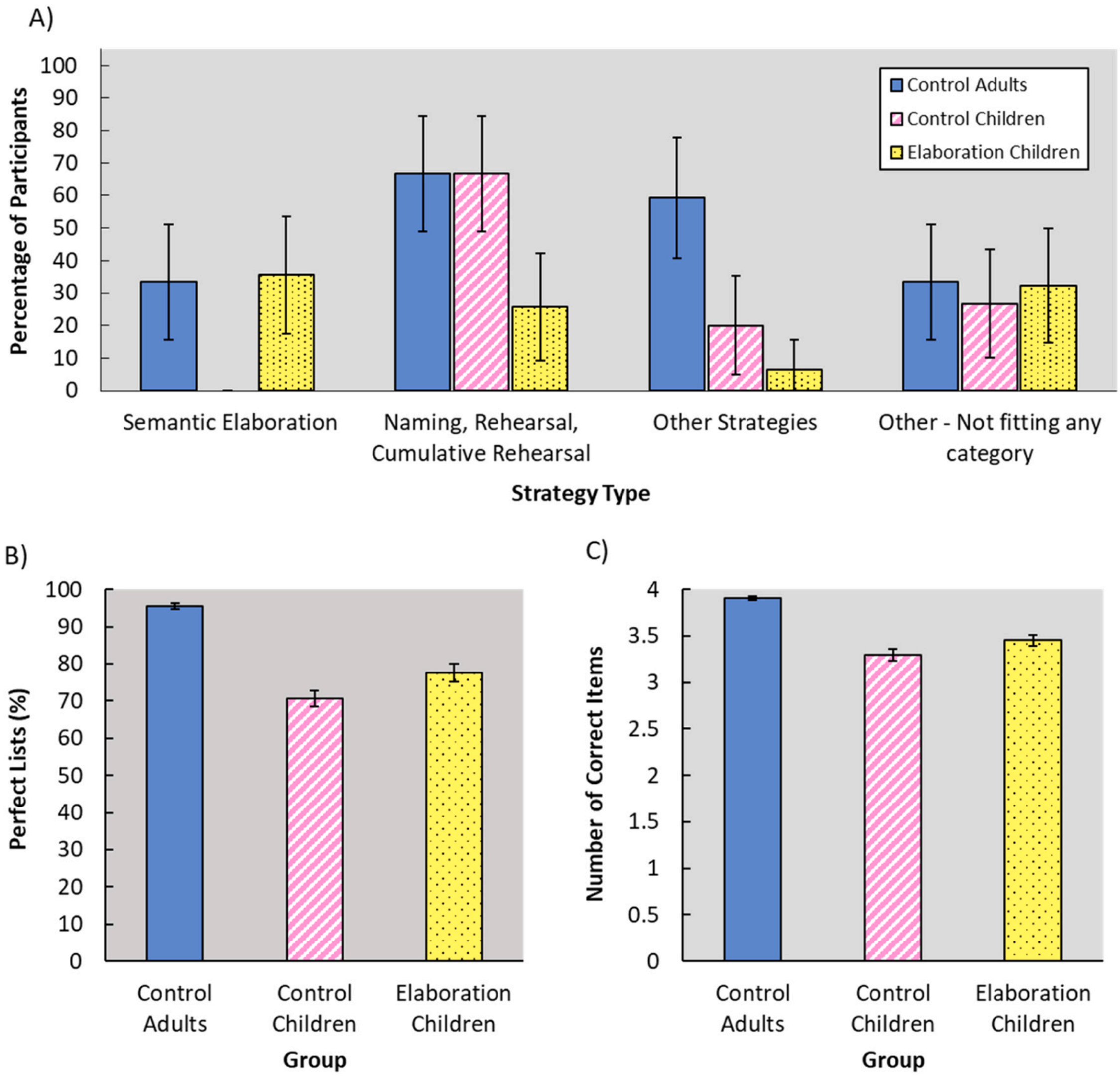

Fig. 3. Percentage of lists correctly reconstructed and number of correct items, per participant group. *Note*. Panel (A) represents the percentage of perfect lists in each group and panel (B) represents the average number of correct items. Error bars represent the standard errors of the mean.

In sum, results from report trials suggest that the use of elaboration was rare among control children and that trained children complied with the story-making strategy instruction, even though not always including all list items in the stories. The complete transcription of stories and strategy descriptions provided in report trials can be found on OSF (https://osf.io/nja4c/).

4.3. Mentions to elaboration in the questionnaire

Participants responded to a questionnaire about their strategy use at the end of the serial reconstruction task. Here we report only answers to the open-ended question, which reveals participants' spontaneous meta-cognitive descriptions of strategy use. First and foremost, no child in the control group spontaneously mentioned the use of elaboration in the open-ended question, compared to 11 children (35 %) in the elaboration group, yielding decisive evidence ($BF_{10} = 308.61$, $p < .001$) for a difference between the two child groups. Of the adults, 9 of them (33 %) mentioned elaboration, the same ones who also mentioned it in report trials, with no evidence for a difference compared to children in the elaboration group ($BF_{10} = 0.31$, $p = .724$).

Among controls, verbal strategies (naming, rehearsal, cumulative rehearsal) were the most cited ones, mentioned by 66.7 % of adults and 66.7 % of control children, compared to only 25.8 % of children in the elaboration group. The proportion of control children mentioning verbal strategies was significantly higher than elaboration children, as supported by Bayesian contingency tables and Chi-square tests ($BF_{10} = 282$, $p < .01$, conclusive evidence). The proportion was also significantly higher in control than elaboration children when strategies were broken down into naming ($BF_{10} = 400$, $p < .001$, conclusive evidence) and rehearsal (including cumulative rehearsal, which is grouped here with simple rehearsal) ($BF_{10} = 72.35$, $p < .001$, conclusive evidence).

About 30 % of participants across all groups provided descriptions that did not match any classic strategy type, such as using fingers, letters, renaming the stimuli, and vague references to mental effort or using the brain/mind – and the latter were especially common among children from both groups.

In sum, responses in report trials and answers to the questionnaire suggest that our story-making training effectively changed children's approach to the task, making them use elaboration to the same extent as adults. Next, we assess whether elaboration improved children's performance in serial order reconstruction, which was our main variable of interest.

4.4. Serial order reconstruction

The reconstruction results are shown in Fig. 3. On average, the percentage of perfect list reconstructions was 70.6 % (*SD* = 12 %) for control children, 77.6 % (*SD* = 13 %) for elaboration children, and 95.6 % (*SD* =4 %) for adults. There was a trend for a numerical advantage for elaboration children over control children (a 7.6 % increase), but this was not confirmed by the statistical analysis. We found decisive evidence ($BF_{10} = 3.035 \times 10^{11}$, $p < .001$, $\eta_p^2 = 0.48$) for a group effect, but it was driven exclusively by the adult group. Post-hoc pairwise comparisons showed that the percentage of perfect lists in the adult group was superior to control children ($BF_{10} = 8.66 \times 10^{11}$, $p < .001$) and elaboration children ($BF_{10} = 1.48 \times 10^{6}$, $p < .001$), but the difference between the two child groups was in the indeterminate range ($BF_{10} = 1.70$, $p = .04$).

We also measured the number of items placed in correct serial positions within the lists. The average number of correct items was 3.30 (*SD* = 0.34) for control children, 3.45 items (*SD* = 0.35) for elaboration children, and 3.90 items (*SD* = 0.10) for adults. Again, we found decisive evidence of a group effect ($BF_{10} = 1.14 \times 10^{9}$, $p < .001$, $\eta_p^2 = 0.44$) that was driven by the adult group, and no evidence

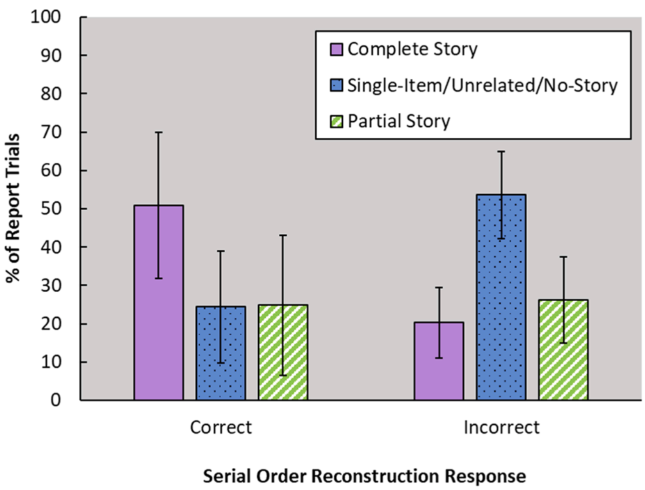

Fig. 4. Percentage of stories in correct and incorrect order reconstruction trials, according to story completeness criterion. *Note*. Complete story: all four list items were included; Partial story = at least two, but not all four list items were included. Single-item story: only one list item was included; Unrelated story: a story was reported but it did not include any of the items in the current list; No-story = no story was reported in the trial. Error bars represent the margin of error at 95 % confidence level.

for a difference between the two child groups ($BF_{10} = 0.95$, $p = .12$). Note, however, that we did not find evidence in favor or the equality of the two child groups either, given the BF being greater than 0.3. The number of correct items in the adult group was greater than in control children ($BF_{10} = 8.34 \times 10^9$, $p < .001$) and in elaboration children ($BF_{10} = 1.25 \times 10^6$, $p < .001$).

In sum, we found a large age effect, but we did not find the anticipated effect of elaboration training on children's serial order reconstruction scores.

4.5. Reported stories in the elaboration group and WM performance

Although the strategy reports in the elaboration group suggest compliance with the story-making instruction, the ability to create stories varied markedly between children. The reported stories varied in length, complexity, in how many items were included, and in whether the order of items was preserved in the narrative. This variability could account for the absence of a beneficial effect of elaboration at the group level. Here we characterize the variety of stories and then examine the relation between story quality and order reconstruction (Hypothesis 3b) in three ways. First, we pool all report trials to examine how those with correct order reconstruction compare to those with incorrect reconstruction. Second, we examine whether the quality of the stories correlates on a trial-by-trial basis with WM accuracy within each participant. Third, we examine whether individuals' average story quality relates to their average WM accuracy.

4.5.1. Overall relation between stories and order reconstruction

Excluding one participant whose audio recording was unusable, the analysis included 10 report trials per child in the elaboration group, totaling 310 trials. Of these, 226 trials yielded correct order reconstruction, while 84 trials yielded incorrect reconstruction. Complete stories were reported in most trials (50.9 %) with correct order reconstruction, followed by partial stories (24.8 %) and single-item, unrelated, or no-story reported (24.3 %). In trials with incorrect order reconstruction, on the other hand, single-item, unrelated, and no-stories formed most reports (53.6 %), followed by partial stories (26.2 %) and complete stories (20.2 %). Overall, children produced stories including more list items in trials of correct order reconstruction.

It is possible that order reconstruction errors were determined by order errors in the story. For example, given the list *tap – popcorn – airplane – window* (ABCD), the story "*The tap ate popcorn, looked through the window and saw an airplane*" could yield the incorrect order reconstruction *tap – popcorn – window – airplane* (ABDC). We tested this hypothesis by running a correlation between the edit distances of stories and order reconstruction responses. Only a very weak correlation was observed (Pearson's $r = 0.11$), but we found no evidence in favor this hypothesis ($BF_{10} = 0.53$, $p = .044$). It is possible that either stories are sometimes not applied to the reconstruction task, or else transfer from the stories to reconstruction can be erroneous.

4.5.2. Within-subject comparison of story quality in correct vs. incorrect trials

Here we calculated each participant's average story edit distance in correct and incorrect reconstruction trials, separately. We wanted to know whether, for each participant, the story edit distance would significantly differ when they gave correct and incorrect reconstruction responses. We tested this difference with a Bayesian paired *t*-test (one-tailed: edit distance in correct trials < in incorrect trials). Three participants who had either only correct or only incorrect responses were not included in this analysis, resulting in 28 children ($M_{age} = 7$ years and 2 months, $SD = 7$ months). The average story edit distance was 1.69 ($SD = 0.98$) in correct reconstruction trials and 2.73 ($SD = 0.89$) in incorrect reconstruction trials, yielding decisive evidence for a difference ($BF_{10} = 1220$, Cohen's $d = 1.12$, $p < .001$). On average, story edit distances were shorter in correct than incorrect trials, meaning that participants created better stories when they reconstructed perfect sequences.

4.5.3. Individual differences in story quality and WM performance

We adopted two approaches to test the hypothesis that individual differences in creating stories are associated to individual differences in WM performance. First, we calculated each participant's average story edit distance. We ran correlations between the edit distances, the average number of correct items, the percentage of perfect list reconstruction, and the initial measure of digit spans. These correlations inform us whether a child's ability to create stories including all list items in the correct expected serial order is associated to their average WM performance in the order reconstruction task and their initial WM capacity, as measured by the digit spans. Thirty-one participants with available audio transcriptions were included in this analysis, and we used their average WM performance in the whole task, not only in story report trials. We found decisive evidence for strong correlations between the story edit distances and the number of correct items (Pearson's $r = -0.68$, $BF_{10} = 9.46 \times 10^2$, $p < .001$) and the percentage of perfect lists ($r = -0.70$, $BF_{10} = 1.81 \times 10^3$, $p < .001$). However, we found no conclusive evidence for a correlation between story edit distances and children's digit spans ($r = -0.28$, $BF_{10} = 0.70$, $p = .27$). Note that the observed BF_{10} is in the indeterminate range to accept the null hypothesis, i.e., we cannot claim to have found evidence against such a correlation either.

Second, we examined whether using connective words, like "and" and "then", in the stories could reliably predict successful order reconstruction. To investigate this relationship, we computed the number of times the words "and" and "then" were used in each story and ran Bayesian correlations between this measure and the number of correct items in those trials. We found moderate evidence for a weak correlation between the number of connective words and correct reconstruction ($r = 0.16$, $BF_{10} = 8.27$, $p = .004$). In sum, shorter average edit distances were strongly correlated with better serial order reconstruction (with the caveat that no causation can be inferred from this result); nonetheless, using more connective words in the stories was only weakly correlated with serial order reconstruction.

5. Discussion

We investigated whether semantic elaboration supports the development of memory for serial order in children in early elementary school. Our hypothesis was that training children to use story-making, a form of semantic elaboration, would enhance serial order memory, making them outperform same-age peers and near adults without strategy-specific instructions. Our study taps into three important questions that have been disregarded in the literature: first, whether children and adults spontaneously use semantic elaboration in the context of a serial order task; second, whether children can be trained to increase the use of semantic elaboration in the form of story-making to memorize lists of object pictures; and third, if they can be trained in story-making, whether it enhances their serial order memory. Supplementary to those main questions, we explored data from self-reports on strategy use to gain insight about children's strategy development and meta-cognitive comprehension of WM strategies. We will discuss findings relating to each question below.

5.1. Summary of findings

5.1.1. Is semantic elaboration spontaneously used by children and adults to support serial order memory?

This was assessed by looking into strategy reports by control children and adults, who did not receive any strategy-specific instruction. The answer is yes, they do – but children much less so than adults. Only 10 % of control children reported using stories or other type of elaboration during report trials. However, these reports were not reflected in the strategy questionnaire, where only naming, rehearsal, and cumulative rehearsal were mentioned. In contrast, 33 % of adults reported using elaboration and mentioned it in the questionnaire. In sum, strategy reports by control participants suggest that the spontaneous use of semantic elaboration is a prominent difference between children and adults, rarely reported by 7-year-olds. While semantic elaboration is more frequently used by adults, it is not a widely chosen strategy in the context of serial order reconstruction. Adults use a mixture of strategies during the task, with verbal strategies (naming, rehearsal, cumulative rehearsal) being the most cited ones. It is possible that the use of elaboration in addition to verbal strategies in adults plays an important role in promoting superior performance in comparison to children.

5.1.2. Can children be trained to use story-making to maintain serial order?

Children's answers in story report trials clearly confirm that the story-making training changed their approach to the task and their subjective strategy reports. In the questionnaire, 29 % of children in the elaboration group mentioned the use of stories. In report trials, 90.3 % of them were able to create a complete story including all list items, and all of them were able to create partial stories including some of the list items. These results present clear evidence that 7-year-olds can be trained to use elaboration in the form of story-making to memorize lists of object pictures, and that the training matched adult levels of self-reports in the questionnaire. The discrepancy between the number of children reporting stories in report trials and mentioning stories in the questionnaire at the end of the procedure suggests low levels of meta-cognition at that age (Bjorklund and Zeman, 1982, 1983; Forsberg et al., 2021).

5.1.3. Is story-making effective to improve children's serial order memory?

Contrary to our hypothesis, we found no evidence that the story-making training improved children's memory for serial order, despite matching adults' frequency in use of elaboration. Despite a trend towards an elaboration advantage, the numerical difference in memory performance between the two child groups was indeterminate. That result seems analogous to what Cowan et al. (2006) obtained when trying to improve children's serial order recall by speeding up recall. In both cases, the training manipulation succeeded in causing the intended change in behavior concomitant with WM activity, but nevertheless did not improve memory performance itself. Here we make the caveat that we did not find compelling evidence in favor of the null hypothesis either (i.e., BFs $<$ 0.3), so we cannot completely rule out the possibility of training benefits.

On the other hand, trained children with overall better order reconstruction performance had average shorter edit distances in stories. Moreover, story edit distances were shorter in trials with correct than incorrect order reconstruction. These results need to be interpreted with caution: despite showing that children who created better stories (i.e., including more list items and preserving the input order) had better memory reconstruction, they can only tentatively suggest an individual benefit of story-making at the trial level. This is because failure to create stories in incorrect trials could be a consequence of incorrect (or absent) memory representations of items and/or their serial order, which prevented children from being able to create better stories. We found only a very weak correlation between edit distances in the stories and the serial order reconstruction scores, suggesting that children may not consistently apply the stories to the reconstruction task, or that transfer from the stories to the reconstruction may sometimes be inaccurate. Moreover, partial or unrelated stories accompanied correct order reconstruction in about half of the report trials (see subsection Overall Relation Between Stories and Order Reconstruction); this suggests that, while children complied with the instructions, they were not necessarily incorporating the memorized material into the stories. Together, these observations suggest two separate steps in the completion of the task: one involving the story creation, and a subsequent step of applying the story to the order reconstruction task. If the story is not effectively used by participants as a device to retrieve serial order at the response phase of the trial, success in the first step will not necessarily imply correct order reconstruction and vice-versa. In sum, the potential individual effects of the story-making did not translate into a memory benefit at the group level.

5.1.4. How do strategy reports develop?

Children's answers in story report trials clearly confirm that the story-making training changed their approach to the task and their subjective strategy reports. Results from the strategy questionnaire show that a variety of strategies are used by children and adults in the context of serial order reconstruction, with a high frequency of mentions to verbal strategies by control groups. Elaborative strategies were never mentioned by control children, while 33 % of adults and 35 % of children in the elaboration group reported using it. Moreover, mentions to various other strategies, such as using fingers, letters, creating new words, and mentions to mental effort were equally frequent with about 30 % of mentions in all three groups. In sum, self-reports suggest a preference for verbal strategies by control adults and children, more frequent use of elaboration by adults than children, and a large variety of co-existing strategies applied to the maintenance of serial order memory by both children and adults.

5.2. Implication of findings

5.2.1. The development of semantic elaboration

Our results indicate that children as young as 7 years are able to use elaboration to memorize serial order and can be trained to do so. However, the absence of a beneficial group effect of the story-making training and rare spontaneous occurrences of elaboration in control children suggest that this strategy is not yet fully developed at this age. These results concur with the literature suggesting that semantic elaboration has a late development, by the end of elementary school (Beuhring and Kee, 1987; Kemler and Jusczyk, 1975; Rohwer and Bean, 1973). Moreover, conflicting reports by children in report trials but not in the questionnaire also concur with the literature showing undeveloped metacognitive skills at this age (Bjorklund & Zeman, 1982, 1983; Forsberg et al., 2021; Schaefer et al., 2023). Finally, the increase in frequency in the use of elaboration by adults replicates observations of a developmental progression toward the use of more meaningful, deeper strategies (Beuhring & Kee, 1987), and extends them to the context of a serial order task.

5.2.2. The effectiveness of the story-making training

To our knowledge, our study is the first to access the effectiveness of story-making to promote the use of semantic elaboration in supporting children's serial order memory. While succeeding in promoting a behavioral change - the adoption of the semantic elaboration strategy - our training failed in promoting memory benefits in children. We will discuss the implication of these findings separately.

5.2.2.1. Strategy production. Our results extend the literature on semantic elaboration training by showing that 7-year-olds can be taught to use story-making to memorize lists of object pictures. Past studies have already been successful in training children in early elementary school to use other forms of semantic elaboration (Begg and Anderson, 1976; Pressley and Levin, 1977b; Varley et al., 1974; Yuille and Catchpole, 1973). These studies, however, were restricted to sentence generation and interactive imagery applied to the learning of paired-associate learning – a much less demanding endeavor than the one entailed by our task.

Not only does memorizing object lists involve a higher WM load than memorizing word pairs but creating narratives with four elements can also be more demanding than associating two items via a mental image or sentence. For example, for the associated pair *frog-purse,* only one association is required (say, an image of a frog jumping into a purse, or the sentence *"The frog has a purse"*). A narrative with four items, on the other hand, requires the creation of either more complex, multi-items associations, or the creation of at least two associated pairs or one associated pair along with single-item, unassociated representations. The first case is exemplified by four list items associated into one single representation in WM: for the list *frog-cake-purse-skate*, the story "*the frog carries cake in a purse while he rides a skate*" (the four items are associated with each other). The second case is exemplified by two two-item associations held concomitantly in WM: for the same list, the story "*the frog was eating cake, then the purse was skateboarding"* (frog is associated with cake, purse is associated with skate). The third case is exemplified by one two-item association and two unassociated representations in WM: *"I jump like a frog, then I eat cake, then I take my purse to go skating"* (frog and cake are single, unassociated items, purse is associated with skate). Given these complexities in the use of story-making, it was even possible that story-making would be a burden and, instead, hamper participants memory performance. This nonetheless did not happen – a case of strategy mediation deficiency.

5.2.2.2. Mediation deficiency of the story-making strategy. This is not the first case in which a specific strategy can be successfully trained in children but remains ineffective in improving WM. Mediational strategy deficits have been extensively reported in the developmental literature (e.g., Flavell, 1970). They occur when children do not benefit from an instructed strategy that is appropriately implemented in a task, as opposed to production deficiencies – when children fail to spontaneously produce a strategy, even though they are capable of effectively using it (e.g., Leslie, 1980). For instance, in order to test rehearsal speed theories of development, Cowan et al. (2006) succeeded in training 8-year-olds to speed up their speech rates to match adults' usual rate of responses in a digit recall task; nonetheless, faster response speeds did not increase children's span. In our case, making children use elaboration as frequently as adults also did not improve their serial order memory.

We can think of at least three, non-mutually exclusive accounts for the mediation deficiency of stories in our study. First, it is likely that the implementation of story-making has drained too much of children's WM capacity, thus leaving fewer resources for storage and thereby obliterating any possible benefit of the strategy use. Indeed, studies specifically tackling the mental effort of strategy use have shown that the costs associated with strategy implementation cause scant improvement in children's performance (Guttentag, 1989; Guttentag & Ornstein, 1990). For example, in a study by Bjorklund and Harnishfeger (1987), adults showed greater recall than children (9-year-olds and 12-year-olds) despite the same mental effort (as measured by dual-task costs incurred by finger-tapping upon word recall). When a strategy was imposed to participants (to group words by semantic category), the mental effort increased to the same extent in all age groups but only resulted in improved recall for adults – for the same cognitive resource, less was gained by children. Specifically on the costs of elaboration, Kee and Davies (1988) showed that elaboration improved memory performance (as compared to rehearsal) in both sixth graders and adults (ages not reported), but the associated mental effort was greater in children. Thus, it is possible that the mental effort associated with the story-making was too high for our participants and prevented memory benefits. This should be clarified in future studies by implementing a dual-task method to the story-making paradigm, thereby measuring the mental effort of using it in different age groups.

A second explanation for the ineffectiveness of story-making is that, in mature strategy users, semantic elaboration supplements the use of rehearsal in serial order tasks. Here, both strategies are employed but most of the heavy lifting is accomplished through rehearsal. In this scenario, the gain in using elaboration is only incremental to that promoted by rehearsal. Because our procedure specifically instructed children to create stories, it is possible that the use of elaboration came at the expense of reported naming and rehearsal. Our data from strategy reports seem to support this view: the percentage of children mentioning naming and rehearsal dropped by 45 % and 22 % respectively in the elaboration group in comparison to control children. Moreover, while control children never mentioned elaboration, adult controls mentioned it alongside with naming and rehearsal, suggesting some sort of combined strategy use. Together with data showing the variety of other strategies in both age groups, this interpretation is in line with the idea of a developmental diversification of co-existing strategies (Siegler, 2016), rather than the hypothesis of a gradual transition from using simpler forms of rehearsal to more complex, deeper semantic strategies (Rohwer et al., 1977; Rohwer, 1974).

Finally, the effectiveness of semantic elaboration to support serial order memory is still disputed even in adult research. Whilst some studies have successfully found an association between using elaboration with better serial recall (Bailey et al., 2009; Bailey et al., 2011; Bartsch et al., 2024; Loaiza and Lavilla, 2021), others have failed to observe an improvement in memory for the order of items (Thomas et al., 2023). More broadly, the contribution of semantic information (fundamental to elaboration) to serial order memory remains unsettled, with most findings reporting null effects in serial recall tasks (Kowialiewski et al., 2023; Neale and Tehan, 2007; Neath et al., 2022; Poirier and Saint-Aubin, 1995; Saint-Aubin and Poirier, 1999) and only limited evidence supporting a beneficial effect under a narrow set of experimental arrangements (Kowialiewski et al., 2022; Kowialiewski et al., 2024). These involved the presentation of words belonging to the same semantic categories in subgroups within lists, so as to create a list structure where certain serial positions are consistently associated with given categories. This was not the case in our study, in which lists were composed of semantically unrelated words.

In developmental samples, word lists of context-dependent categories (e.g., an event-based schema of dressing up for winter: *coat* – *sweater* – *hat* – *trousers*) support improved serial recall as early as age five, whereas benefits of context-independent, supra-ordinate categories (e.g., clothing: *dress* – *socks* – *pajamas* – *t-shirt*) emerge at around age 9, reflecting the maturation of conceptual knowledge (Monnier and Bonthoux, 2011; Nelson, 1985). However, semantic organization does not seem to reduce the number of order errors in children. In some cases, it can even increase transposition errors despite promoting an overall increase in the verbal span; adults, on the other hand, are not so prone to this increase in transposition errors caused by semantic knowledge, whilst also benefitting from it in terms of an increase in the verbal span (Rosselet-Jordan et al., 2022).

A key difference between the aforementioned studies and ours is that they did not directly examine the strategic use of semantic elaboration, but rather assessed the influence of semantic knowledge using semantically related word lists. Taken together, our results and theirs suggest that semantic knowledge can influence serial recall in children and adults, but previously observed improvements are unlikely to reflect a real advantage of the effortful creation of associations between items. Instead, they are likely to reflect a simpler, non-strategic effect of semantic relatedness upon serial recall. In sum, semantic relatedness per se (which was not manipulated in our design) seems to improve serial recall in both children and adults, but it can also increase order confusions in children – suggesting higher reliance of their WM system on long-term memory due to lower maturational levels and/or immature strategy use (e.g., difficulties in implementing accumulative rehearsal). It is unclear whether using elaboration to encode lists of semantically similar items could mitigate order confusions in children while maintaining the benefits in span.

Together, these findings imply that the effectiveness of semantic elaboration strategies such as story-making may depend critically on strong semantic associations between list items and on list structure. When list items share strong, previously-learned semantic relations (Monnier and Bonthoux, 2011; Rosselet-Jordan et al., 2022) and semantic categories are strongly tied to list positions (Kowialiewski et al., 2022; Kowialiewski et al., 2024), elaboration can indeed enhance serial recall; when items are arbitrary or weakly related, like in our study, the scaffolding effect of semantic knowledge may be insufficient to yield measurable gains when only memory for serial positions is tested in serial order reconstruction – though we observed a numerical trend towards an advantage.

We reiterate that our rationale for a beneficial effect of elaboration relied on the sequential nature of story-making, which we hypothesized to have the capability to reinforce the encoding of serial order. In future research, adding elaboration instructions to semantically similar lists could leverage the limited benefits of semantic knowledge upon serial order memory, as observed in the studies discussed above. This might be especially likely when there is a natural pattern to the sequence to be noticed (e.g., in terms of increasing size: *mouse* – *dog* – *lion* – *elephant*). A delay in responding might allow more time for the semantic organization of a list to be mentally reconstructed, an example of what is termed hypermnesia (Payne, 1987). Alternatively, elaboration might be an ineffective strategy to maintain the order of associations in WM (Thomas et al., 2023), even with strongly associated materials that are conducive of elaboration. If this proves to be true, then the significant presence of spontaneous adult elaborators in our sample spurs new questions on WM strategy choice.

6. Conclusion

We have shown that children in early elementary school can be successfully trained to use story-making in a serial order task, which is a more complex and demanding type of elaboration training than commonly reported in the literature. Moreover, data from participant self-reports show a developmental increase in the use of semantic elaboration within the context of serial order memory.

However, even when children matched adult levels of strategy production, their WM capacity remained inferior. This finding reinforces previous conclusions that adult-like behavior in children does not necessarily translate to adult-like performance in terms of WM capacity (Cowan et al., 2006) and concurs with the adult literature showing little or no benefit of semantic knowledge on serial order memory (Kowialiewski et al., 2022; Kowialiewski et al., 2023; Kowialiewski et al., 2024; Neath et al., 2022; Saint-Aubin and Poirier, 1999). Our results underscore the importance of considering both strategy use and underlying maturational factors, e.g., a growth in storage capacity, when evaluating WM development in childhood.

CRediT authorship contribution statement

Luísa Superbia-Guimarães: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Reese Lavers**: Data curation. **Maya Steiger**: Data curation. **Kellen Hendrix**: Data curation. **Bret Glass**: Supervision, Project administration, Data curation. **Nelson Cowan**: Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Microsoft Copilot for text editing purposes. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This work was supported by the National Institutes of Health [Grant R01-HD021338 to Nelson Cowan]. The authors have no conflict of interest to declare.

Acknowledgements

The authors thank participants and their families for taking part in this study, Dominic Guitard for helpful programming advice, and Gabriel Arantes Tiraboschi for insightful discussions on data analysis.

Appendix A

Lists of Picture Objects and Spoken Words

Table A1
Object Lists Used in the Serial Order Reconstruction Task

Serial position			
1	2	3	4
helmet	sneaker	cracker	door
toaster	onion	zebra	frog
helmet	sneaker	cracker	door
bees	ox	wallet	celery
scooter	package	canoe	knife
wrench	crayons	tire	hotdog
hanger	tin	dresser	chipmunk
towel	screen	dustpan	ruler
domino	peanut	chair	fork
rooster	pacifier	sink	can
birdhouse	broccoli	sheep	pickle
pizza	crocodile	bell	pasta
teeth	mug	fan	pie
perfume	eraser	apple	cake
starfish	brush	sword	boots
giraffe	basket	crib	cushion
tap	popcorn	airplane	window
truck	teddy bear	teapot	baseball bat
tweezers	caterpillar	drink	muffin
chocolate	cat	bird	fox
garlic	desk	igloo	gloves
watch	bucket	rocket	cupcake
stairs	beetle	medicine	toothpaste
boy	drum	clock	elephant
barbecue	stoplight	couch	bulb
bench	penguin	scissors	doll

(continued on next page)

Table A1 (*continued*)

Serial position			
baseball	motorcycle	cage	pool
burger	bananas	goat	tray
cow	sausage	plum	crown
butterfly	nest	bottle	milk
table	ice cream	pan	cheese
school bus	raspberry	cap	ballerina
building	shorts	rock	camel
train	oven	lock	shovel
kettle	comb	gift	ladder
glasses	grapes	ring	jacket
snack	robot	paint	bowl
spider	plate	snail	ice skate
meat	seal	milkshake	bunny
bread	washer	candies	tissue
pants	dragon	shopping cart	ice
spices	corn	axe	crab
backpack	trophy	lollipop	brownie
bed	horse	cane	bowling pin
melon	trash	fish	bicycle
umbrella	purse	jug	guitar
lipstick	plant	pen	shirt
football	bathtub	notebook	dolphin

Note. Lists were visually presented as a sequence of object pictures. During the familiarization phase, participants were shown each object image while simultaneously hearing the corresponding word from the table.

Table A2
Foil Objects and Words Used in the Familiarization Phase

Foil object picture	Non-matching word
blender	lamp
books	sandals
buoy	donut
cookie	globe
eggs	lemon
fire	tree
house	playground
keys	coins
ladybug	bat
leaf	flower
lobster	turtle
mask	sock
pillow	rug
pine cones	pumpkin
pot	hat
skateboard	ball
spoon	hammer
strawberry	carrot
tiger	dog
toilet	toothbrush
turkey	swan
wardrobe	refrigerator

Note. Foil objects were never used during the serial order reconstruction. Each foil object picture was presented simultaneously to the audio presentation of a non-matching word. Upon a keyboard press, the audio with the corresponding word in column 1 was played. E.g., In the first row, the audio *"This is a lamp"* accompanied the picture of a cookie, and the audio *"Sorry, this is a cookie"* was played upon a keyboard press.

Appendix B

Instructions Given to Participants In the Serial Order Reconstruction Task

Table B1
Instructions in Each Phase of the Experimental Procedure, According to Group Condition

Stage of procedure	Instructions	
	Elaboration	Control
Digit Span	*"In this challenge, Tigress will say lists of numbers that you will have to remember. After Tigress finishes a list, you have to repeat it. You must repeat the numbers in the same order as Tigress. Her lists will get longer until you cannot remember them. Let's try an example.* […] *Do you have any questions? Now let's start."*	
Serial Order Reconstruction: General instructions	Elaboration (E): *"Welcome to this memory game! In this game, you have to memorize the order of pictures that will appear on the screen* ***and make up short stories with them.*** *Before we start the game, the computer will show you some pictures of objects and tell you their names. Sometimes the computer gets confused and makes mistakes. I want you to listen carefully and press the SPACEBAR whenever the computer says the wrong name for an object. Are you ready?"*	Control (C): *"Welcome to this memory game! In this game, you have to memorize the order of pictures that will appear on the screen. Before we start the game, the computer will show you some pictures of objects and tell you their names. Sometimes the computer gets confused and makes mistakes. I want you to listen carefully and press the SPACEBAR whenever the computer says the wrong name for an object. Are you ready?"*
Practice trial 1	*"Now that the computer knows the name of the pictures, let's practice the memory game. In this game, you have to memorize the order of pictures that will appear on the screen. Pay attention and try not to forget the order. A few seconds after you see the pictures, they will appear on the screen in the wrong order. Your job is to put them in the correct order, by clicking on them. Let's try an example?"* • Sequential presentation. • Response screen, *"Can you put these in the order in which you saw them? The toaster was first."* • Feedback.	
	E: *"Now let's make up a short story with the pictures. Let's try?"*	C: *"Let's practice some more."*
Practice trial 2 (Repeat up to 3x on error)	E & C: • Simultaneous presentation, items remain on the screen with the text:	
	E: *"The helmet fell and hit the sneaker, then a furious cracker broke through the door."* • New screen with the text:	C: *"Memorize the order of the pictures."*
	E: *"Now let's present these one at a time and you try to remember the order by using the story."*	C: *"Now let's present these one at a time and you try to remember the order."*
	E & C: • Sequential presentation. • Response screen, *"Can you put these in order?*" • Feedback.	
Message in case of error:	E: *"Let's try again. This time, try to use the story to help you."*	C: *"Let's try again."*
Message after 3rd attempt:	E: "*Now try to make up a short story by yourself. Make it the shortest you can and use only the pictures you see. It can be any story you can come up with. Are you ready?"*	C: "*Let's keep practicing with other pictures."*
Practice trial 3 (Repeat up to 3x on error) Message in case of error: Message after 3rd attempt:	• Simultaneous presentation, items remain on screen with the text: E: *"Make up a short story with these pictures"* (free time) E & C: • Response screen, "Can you put these in order?" • Feedback.	C: *"Think about how you can memorize the order of the pictures."* (free time)
	E: *"Almost. Let's try again. Try to change your story so now you can remember the correct order."* *"Let's continue."*	C: *"Almost. Let's try again."*
Practice trial 4 (Repeat 3x on error)	E: *"Now you will see each picture at once. Try to make up a short story with them."* • Sequential presentation of items.	C: "*Now you will see each picture at once. Try to memorize the order."*
	E: *"Make up a story."* (free time)	C: *"Use this time to think about how you can remember the pictures."* (free time)
	• Response screen, *"Can you put these in order?"* • Feedback.	
Message in case of error:	E: *"Almost. Let's practice some more. Try to change your story so now you can remember the correct order."*	C: *"Almost. Let's practice some more."*
Message after 3rd attempt:	*"Let's continue."*	
Practice trial 5 (Repeat 3x on error)	E: *"Now make up your story in silence. I will ask you to tell it to me later."* • Sequential presentation.	C: *"Memorize the order of the pictures."*
	E: *"What story did you make?"* (free time)	C: *"Use this time to think about how you can remember the pictures."* (free time)
	• Response screen, *"Can you put these in order?"* • Feedback.	
Message in case of error	E: *"Almost. Let's practice some more. Try to change your story so now you can remember the correct order."*	C: *"Almost. Let's practice some more. Try to think about better ways of remembering the order."*
Message after 3rd attempt:	"Let's continue."	

(*continued on next page*)

Table B1 (*continued*)

Stage of procedure	Instructions	
	Elaboration	Control
Practice trial 6 (Repeat 3x on error)	E: *"Now make up your story in silence and tell it to me only when I ask you to."* • Sequential presentation. • Response screen, *"Can you put these in order?"* • Feedback. E: *"What story did you create?"*	C: *"Memorize the order of the pictures."* C: *"Could you tell me how you remembered the order of the pictures?"*
Message in case of error:	E: *"Almost. Let's practice some more. Try to change your story so now you can remember the correct order."*	C: *"Almost. Let's practice some more. Try to think about better ways of remembering the order."*
Message after 3rd attempt:	*"Let's continue. Ready?"*	
Practice trial 7 (Repeat 3x on error)	E: *"Now do everything in silence but remember to make up stories, and sometimes I will ask you about the stories."* • Sequential presentation. • Response screen, *"Can you put these in order?"* • Feedback.	C: *"Now do everything in silence and sometimes I will ask you what you did in your head to remember the pictures"*
Message in case of error:	E: *"Almost. Let's try again. Don't forget to make up a short story with the pictures."*	C: *"Almost. Let's try again. Think of better ways of remembering the order."*
Message after 3rd attempt:	*"Let's continue."*	
Block 1 (Naming phase)	*"Before we continue, let's see if you can name some more objects. Tell me the name of the pictures and, if you do not know it, tell me and we will discover it. Are you ready?"* • Object familiarization procedure E: *"Now let's start the memory game. Do not forget to make up stories with the pictures. Are you ready?"* • 20 trials, of which 5 are report trials.	C: *"Now let's start the memory game. Are you ready?"*
Report trials:	E: *"Tell me the story you created."* E: Note to experimenter: If participant verbalizes their story during item presentation, wait until the end of the trial and correct them by saying *"Do not let me hear your stories. I will let you know when you have to share them with me."*	C: *"How did you remember the order?"* C: Note to experimenter: If children rehearse the items out loud during item presentation, wait until the end of the trial and correct them by saying *"Do not let me hear what you are doing to remember the pictures. I will let you know when you have to share this with me."* C: If participant complains about repeatedly reporting the same strategy in report trials, say: *"It is ok to tell me that you are using the same strategy to remember, but I have to keep asking it to you."*
Block 2 (Naming phase)	*"Before we continue, let's check if the computer knows the name of some more objects. Press the spacebar whenever he makes a mistake. Are you ready?"* • Object Familiarization Procedure E: *"Now let's continue the memory game. Do not forget to make up short stories with the pictures. Are you ready?"* • 20 trials, of which 5 are report trials (see instructions above). • Bye-bye emoji and message *"Well done! Thank you and goodbye."*	C: *"Now let's continue the memory game. Are you ready?"*

Note. Instructions were written on the computer screen and read aloud by the experimenter. ISpecific instructions given to children in the elaboration group are signaled with "E:" and specific instructions given to children and adult controls are signaled with "C:".

Appendix C

Strategy Questionnaire and Criteria for Open-Ended Answers

Table C1
Strategy Definitions and Questions in the Questionnaire

Strategy	Definition	Question
_____	_____	1) *"How did you remember the sequences of pictures?"* (open-ended)
Semantic Elaboration	The creation of links between the items, based on previous knowledge.	2) *"Did you try to make a story (or a sentence) involving all the objects? Can you give some examples?"*
Verbal rehearsal	Vocal or subvocal repetition of the words describing the pictured objects.	3) *"Did you repeat the names of the objects in your head?"*
Visualization	Thinking of the appearance and visual features of the pictured objects, e.g., colors or shapes.	4) *"Did you try to visualize the objects in your head, like a picture?"*
Visual imagery	Creation of a mental image combining the pictured objects, based on their meaning.	5) *"Did you try to put together the picture of the objects in your head, for example, a cup inside a box or a shoe kicking a toaster?"*
Acoustic imagery	Thinking of the sounds/noises produced by the pictured objects and focusing on the sound itself.	6) *"Did you try to imagine the sounds produced by the objects, for example, the sound of a clock ticking or sheep saying 'baa'?"*
Grouping	Grouping items in clusters of 2 or 3 items without any reference to prior knowledge.	7) *"Did you create groups of objects, like the first and second pictures together?"*

(*continued on next page*)

Table C1 (*continued*)

Strategy	Definition	Question
Memory reduction	Deliberate attempt to remember only a subset of the presented items, instead of trying to remember them all.	8) *"Did you focus on only a few objects and give up on remembering all of them?"*

Note. All participants were asked the questions in the same order; however, due to a protocol error, not all children in the elaboration group received Question 2. Questions 2 through 8 were answered with Yes or No responses.

Table C2 provides detailed explanations and examples for criteria used to categorize answers in the open-ended question (Question 1).

Table C2
Criteria for Categorizing Answers to the Open-Ended Question

Strategy	Criterion	Example Answer
Semantic Elaboration	Mentions to stories, sentences, or any type of association of items based on their meaning, even if not necessarily including all four items (e.g., "*trash-bike*" = to put the trash outside; "*book-dolphin*" = to throw the book to the dolphin).	*"I imagine someone throwing the book as a frisbee, and the dolphin catching it."* (A)
Naming	Mentions to using the object's name, without necessarily repeating it (e.g., "I said the names when I saw the objects", "I said it in my mind").	*"I said them in my head [...]"* (EC)
Rehearsal	Mentions to repeating or keep saying the names of items in the mind, at least more than once, not necessarily in the order (e.g., "*bed, bed, bed*").	*"Well, mainly I just kept repeating in my head."* (CC)
Cumulative rehearsal	Repeating the names of the items in the same order they were presented, adding the new items in the sequence at every iteration (e.g. "*bed, bed-house, bed-house-tree, bed-house-tree-cat*").	*"Every time an object was flashing on the screen, I was just repeating that word in my mind, and once I had time to memorize it again, I was just, I kept repeating it in a loop in the same order, at least twice [...]"* (A)
Visualization	Mentions to "seeing" the objects in the mind, thinking of them as pictures, or specific description involving their visual features (e.g., mentions to lines, shapes, colors).	*"[...] other times I was, like, thinking of the pictures."* (A)
Visual imagery	Mentions to combining items in the visual domain, emphasizing a pictured impression of such combination (e.g., *lipstick* + *pot* = thick red lines in the pot).	*"Lipstick came before the pot, so I imagine the pot being colored with red, thick red lines with the lipstick."* (A)
Acoustic imagery	Mentions to sounds that would be physically produced in real-life by the objects depicted by the items.	(No occurrence)
Grouping	Mentions to the creation of groups items, for example in pairs.	*"I also tried grouping them. So, I group the first 2 and the last 2 together."* (A).
Memory reduction	Mentions to an elimination process, like memorizing only some of the items and completing the answer based on an exclusion process.	*"Mostly I just concentrated on the first three."* (A)

Note. For simplicity, the examples provided above were chosen to reflect reports matching only one strategy type. The criteria were not mutually exclusive. Answers scored as rehearsal or cumulative rehearsal were mandatorily also scored as naming. A = Adult participant; CC = Control child; EC = Elaboration child.

Data Availability

Materials and data can be found on the OSF page: https://osf.io/nja4c/.

References

Ahmed, S. F., Ellis, A., Ward, K. P., Chaku, N., & Davis-Kean, P. E. (2022). Working memory development from early childhood to adolescence using two nationally representative samples. *Developmental Psychology, 58*(10), 1962–1973. https://doi.org/10.1037/dev0001396

Allik, J. P., & Siegel, A. W. (1976). The use of the cumulative rehearsal strategy: A developmental study. *Journal of Experimental Child Psychology, 21*(2), 316–327. https://doi.org/10.1016/0022-0965(76)90045-X

Attout, L., & Monnier, C. (2023). Evidence of a positive effect of verbal cumulative rehearsal on serial order working memory, as early as 4 years old. *Developmental Psychology, 59*(10), 1771–1783. https://doi.org/10.1037/dev0001621

Bailey, H., Dunlosky, J., & Hertzog, C. (2009). Does differential strategy use account for age-related deficits in working-memory performance? *Psychology and Aging, 24*(1), 82–92. https://doi.org/10.1037/a0014078

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

Bartsch, L. M., Souza, A. S., & Oberauer, K. (2024). The benefits of memory control processes in working memory: Comparing effects of self-reported and instructed strategy use (Advance online publication) *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001370.

Begg, I., & Anderson, M. C. (1976). Imagery and associative memory in children. *Journal of Experimental Child Psychology, 21*(3), 480–489. https://doi.org/10.1016/0022-0965(76)90075-8

Belletier, C., Doherty, J. M., Graham, A. J., Rhodes, S., Cowan, N., Naveh-Benjamin, M., Barrouillet, P., Camos, V., & Logie, R. H. (2023). Strategic adaptation to dual-task in verbal working memory: Potential routes for theory integration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 49*(1), 51–77. https://doi.org/10.1037/xlm0001106

Bender, B. G., & Levin, J. R. (1976). Motor activity, anticipated motor activity, and young children's associative learning. *Child Development*, 560–562. https://doi.org/10.2307/1128822

Bernbach, H. A. (1967). The effect of labels on short-term memory for colors with nursery school children. *Psychonomic Science, 7*, 149–150. https://doi.org/10.3758/BF03328509

Beuhring, T., & Kee, D. W. (1987). Developmental relationships among metamemory, elaborative strategy use, and associative memory. *Journal of Experimental Child Psychology, 44*(3), 377–400. https://doi.org/10.1016/0022-0965(87)90041-5

Bjorklund, D. F., & Harnishfeger, K. K. (1987). Developmental differences in the mental effort requirements for the use of an organizational strategy in free recall. *Journal of Experimental Child Psychology, 44*(1), 109–125. https://doi.org/10.1016/0022-0965(87)90025-7

Bjorklund, D. F., & Zeman, B. R. (1982). Children's Organization and Metamemory Awareness in Their Recall of Familiar Information. *Child Development, 53*(3), 799–810. https://doi.org/10.2307/1129394

Bjorklund, D. F., & Zeman, B. R. (1983). The Development of Organizational Strategies in Children's Recall of Familiar Information: Using Social Organization to Recall the Names of Classmates. *International Journal of Behavioral Development, 6*(3), 341–353. 10.1177/016502548300600304.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2013). Real-world objects are not represented as bound units: independent forgetting of different object details from visual memory. *Journal of Experimental Psychology General, 142*(3), 791–808. https://doi.org/10.1037/a0029649

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*(3), 551–581. https://doi.org/10.1037/0033-295X.106.3.551

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55*(4), 627–652. https://doi.org/10.1016/j.jml.2006.08.005

Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.006.

Cowan, N., Elliott, E. M., Saults, J. S., Nugent, L. D., Bomb, P., & Hismjatullina, A. (2006). Rethinking speed theories of cognitive development: Increasing the rate of recall without affecting accuracy. *Psychological Science, 17*(1), 67–73. https://doi.org/10.1111/j.1467-9280.2005.01666.x

Cowan, N., Li, Y., Glass, B. A., & Scott Saults, J. (2018). Development of the ability to combine visual and acoustic information in working memory. *Developmental Science, 21*(5), Article e12635. https://doi.org/10.1111/desc.12635

Cowan, N., Morey, C.C., & Naveh-Benjamin, M. (2021). An embedded-processes approach to working memory: How is it distinct from other approaches, and to what ends?', in R. Logie, V. Camos, & N. Cowan (eds), Working Memory: The state of the science (pp. 44-84). Oxford.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89*(1), 63–100. https://doi.org/10.1037/0033-2909.89.1.63

Elliott, E. M., Morey, C. C., AuBuchon, A. M., Cowan, N., Jarrold, C., Adams, E. J., & Voracek, M. (2021). Multilab direct replication of Flavell, Beach, and Chinsky (1966): Spontaneous verbal rehearsal in a memory task as a function of age. *Advances in Methods and Practices in Psychological Science, 4*(2), Article 25152459211018187. https://doi.org/10.1177/25152459211018187

Flavell, J. H. (1970). Developmental studies of mediated memory. *Advances in Child Development and Behavior, 5*, 181–211. https://doi.org/10.1016/S0065-2407(08)60467-X

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

Ford, S., & Silber, K. P. (1994). Working memory in children: A developmental approach to the phonological coding of pictorial material. *British Journal of Developmental Psychology, 12*(2), 165–175. https://doi.org/10.1111/j.2044-835X.1994.tb00626.x

Forsberg, A., Blume, C. L., & Cowan, N. (2021). The development of metacognitive accuracy in working memory across childhood. *Developmental Psychology, 57*(8), 1297. https://doi.org/10.1037/dev0001213

Gonthier, C. (2023). An easy way to improve scoring of memory span tasks: The edit distance, beyond "correct recall in the correct serial position. *Behavior Research Methods, 55*(4), 2021–2036. https://doi.org/10.3758/s13428-022-01908-2

Guitard, D., Saint-Aubin, J., & Cowan, N. (2022). Tradeoffs between item and order information in short-term memory. *Journal of Memory and Language, 122*, Article 104300. https://doi.org/10.1016/j.jml.2021.104300

Guitard, D., & Cowan, N. (2023). Attention allocation between item and order information in short-term memory. *Quarterly Journal of Experimental Psychology, 76*(6), 1391–1409. https://doi.org/10.1177/17470218221118451

Guttentag, R. E. (1984). The mental effort requirement of cumulative rehearsal: A developmental study. *Journal of Experimental Child Psychology, 37*(1), 92–106. https://doi.org/10.1016/0022-0965(84)90060-2

Guttentag, R. E. (1989). Age differences in dual-task performance: Procedures, assumptions, and results. *Developmental Review, 9*(2), 146–170. https://doi.org/10.1016/0273-2297(89)90027-0

Guttentag, R. E., & Ornstein, P. A. (1990). *Attentional capacity and children's memory strategy use. In: Enns, J.T. (Ed.), The development of attention: Research and theory* (pp. 305–320). Elsevier. https://doi.org/10.1016/S0166-4115(08)60463-X

Hagen, J. W., & Kingsley, P. R. (1968). Labeling Effects in Short-Term Memory. *Child Development, 39*(1), 113–121. https://doi.org/10.2307/1127363

Hasher, L., & Clifton, D. (1974). A developmental study of attribute encoding in free recall. *Journal of Experimental Child Psychology, 17*(2), 332–346. https://doi.org/10.1016/0022-0965(74)90077-0

Hentschel, M., Averbeck, B. B., & Lange-Küttner, C. (2023). The difference an explanation makes: Verbal debrief effects when children cope with uncertainty. *Cognitive Development, 66*, Article 101317. https://doi.org/10.1016/j.cogdev.2023.101317

Hitch, G. J., & Halliday, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society of London B, Biological Sciences, 302*(1110), 325–340. https://doi.org/10.1098/rstb.1983.0058

Hitch, G. J., Halliday, M. S., Dodd, A., & Littler, J. E. (1989). Development of rehearsal in short-term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology, 7*(4), 347–362. https://doi.org/10.1111/j.2044-835X.1989.tb00811.x

Huttenlocher, J., & Burke, D. (1976). Why does memory span increase with age? *Cognitive Psychology, 8*(1), 1–31. https://doi.org/10.1016/0010-0285(76)90002-5

Jarrold, C., & Citroën, R. (2013). Reevaluating key evidence for the development of rehearsal: Phonological similarity effects in children are subject to proportional scaling artifacts. *Developmental Psychology, 49*(5), 837–847. https://doi.org/10.1037/a0028771

Jarrold, C., & Hall, D. (2013). The development of rehearsal in verbal short-term memory. *Child Development Perspectives, 7*(3), 182–186. https://doi.org/10.1111/cdep.12034

Jeon, M., & De Boeck, P. (2017). Decision qualities of Bayes factor and *p* value-based hypothesis testing. *Psychological Methods, 22*(2), 340–360. https://doi.org/10.1037/met0000140

Kee, D. W., & Davies, L. (1988). Mental effort and elaboration: A developmental analysis. *Contemporary Educational Psychology, 13*(3), 221–228. https://doi.org/10.1016/0361-476X(88)90022-7

Kee, D. W., & Davies, L. (1991). Mental effort and elaboration: A developmental analysis of accessibility effects. *Journal of Experimental Child Psychology, 52*(1), 1–10. https://doi.org/10.1016/0022-0965(91)90002-A

Kemler, D. G., & Jusczyk, P. W. (1975). A developmental study of facilitation by mnemonic instruction. *Journal of Experimental Child Psychology, 20*(3), 400–410. https://doi.org/10.1016/0022-0965(75)90115-0

Kolinsky, R., Gabriel, R., Demoulin, C., Gregory, M. M., de Carvalho, K. S., & Morais, J. (2020). The influence of age, schooling, literacy, and socioeconomic status on serial-order memory. *Journal of Cultural Cognitive Science*, 1–23. https://doi.org/10.1007/s41809-020-00056-3

Kowialiewski, B., Krasnoff, J., Mizrak, E., & Oberauer, K. (2022). The semantic relatedness effect in serial recall: Deconfounding encoding and recall order. *Journal of Memory and Language, 127*, Article 104377. https://doi.org/10.1016/j.jml.2022.104377

Kowialiewski, B., Krasnoff, J., Mizrak, E., & Oberauer, K. (2023). Verbal working memory encodes phonological and semantic information differently. *Cognition, 233*, Article 105364. https://doi.org/10.1016/j.cognition.2022.105364

Kowialiewski, B., & Majerus, S. (2020). The varying nature of semantic effects in working memory. *Cognition, 202*, Article 104278. https://doi.org/10.1016/j.cognition.2020.104278

Kowialiewski, B., Majerus, S., & Oberauer, K. (2024). Does semantic similarity affect immediate memory for order? Usually not, but sometimes it does. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 50*(1), 68–88. https://doi.org/10.1037/xlm0001279

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*, 978–990. https://doi.org/10.3758/s13428-012-0210-4

Lee, C. L., & Estes, W. K. (1981). Item and order information in short-term memory: Evidence for multilevel perturbation processes. *Journal of Experimental Psychology: Human Learning and Memory, 7*(3), 149–169. https://doi.org/10.1037/0278-7393.7.3.149

Lehmann, M., & Hasselhorn, M. (2007). Variable memory strategy use in children's adaptive intratask learning behavior: Developmental changes and working memory influences in free recall. *Child Development, 78*(4), 1068–1082. https://doi.org/10.1111/j.1467-8624.2007.01053.x

Leproult, I., Lemaire, B., & Portrat, S. (2024). Does the extension of free time trigger spontaneous elaborative strategies in working memory? *Memory Cognition, 52*, 2022–2052. https://doi.org/10.3758/s13421-024-01615-7

Leslie, L. (1980). Mediation or production deficiency in disabled readers? *Perceptual and Motor Skills, 50*(2), 519–530. https://doi.org/10.1177/003151258005000230

Levin, J. R., Davidson, R. E., Wolff, P., & Citron, M. (1973). A comparison of induced imagery and sentence strategies in children's paired-associate learning. *Journal of Educational Psychology, 64*(3), 306–309. https://doi.org/10.1037/h0034385

Levin, J. R., McCabe, A. E., & Bender, B. G. (1975). A note on imagery-inducing motor activity in young children. *Child Development, 46*(1), 263–266. https://doi.org/10.2307/1128862

Loaiza, V. M., & Lavilla, E. T. (2021). Elaborative strategies contribute to the long-term benefits of time in working memory. *Journal of Memory and Language, 117*, Article 104205. https://doi.org/10.1016/j.jml.2020.104205

Majerus, S. (2019). Verbal working memory and the phonological buffer: The question of serial order. *Cortex, 112*, 122–133. https://doi.org/10.1016/j.cortex.2018.04.016

McCabe, A. E., Levin, J. R., & Wolff, P. (1974). The role of overt activity in children's sentence production. *Journal of Experimental Child Psychology, 17*(1), 107–114. https://doi.org/10.1016/0022-0965(74)90009-5

McFarland, C. E., & Kellas, G. (1975). Category similarity effects in children's semantic memory retrieval. *Journal of Experimental Child Psychology, 20*(3), 369–376. https://doi.org/10.1016/0022-0965(75)90112-5

Miller, S., McCulloch, S., & Jarrold, C. (2015). The development of memory maintenance strategies: Training cumulative rehearsal and interactive imagery in children aged between 5 and 9. *Frontiers in Psychology, 6*, 524. https://doi.org/10.3389/fpsyg.2015.00524

Monnier, C., & Bonthoux, F. (2011). The semantic-similarity effect in children: Influence of long-term knowledge on verbal short-term memory. *British Journal of Developmental Psychology, 29*(4), 929–941. https://doi.org/10.1111/j.2044-835X.2010.02024.x

Morey, C. C., AuBuchon, A. M., Attwood, M., Castelain, T., Cowan, N., Crepaldi, D., Fjerdingstad, E., Fredriksen, E., Jarrold, C., Koch, C., Lelonkiewicz, J. R., Lupyan, G., Mendenhall, W., Moreau, D., Schonberg, C., Tamnes, C. K., Vlach, H., & Elliott, E. M. (2024). Is verbal rehearsal strategic? An investigation into overt rehearsal of nameable pictures in 5- to 10-year-old children. *Journal of Cognition and Development*. https://doi.org/10.1080/15248372.2024.2389123

Morra, S., Patella, P., & Muscella, L. (2024). Modelling working memory capacity: Is the magical number four, seven, or does it depend on what you are counting? *Journal of Cognition, 7*(1), 60. https://doi.org/10.5334/joc.387

Nairne, J. S., & Kelley, M. R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language, 50*(2), 113–133. https://doi.org/10.1016/j.jml.2003.09.005

Naveh-Benjamin, M., & Jonides, J. (1984). Maintenance rehearsal: A two-component analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(3), 369–385. https://doi.org/10.1037/0278-7393.10.3.369

Neale, K., & Tehan, G. (2007). Age and redintegration in immediate memory and their relationship to task difficulty. *Memory Cognition, 35*(8), 1940–1953. https://doi.org/10.3758/BF03192927

Neath, I., Saint-Aubin, J., & Surprenant, A. M. (2022). Semantic relatedness effects in serial recall but not in serial reconstruction of order. *Experimental Psychology, 69*(4), 196–209. https://doi.org/10.1027/1618-3169/a000557

Nelson, K. (1985). *Making sense: The acquisition of shared meaning.* Academic Press.

Ornstein, P. A., & Naus, M. J. (1978). Rehearsal processes in children's memory. In In. P. A. Ornstein (Ed.), *Memory Development in Children* (pp. 69–99). Erlbaum.

Payne, D. G. (1987). Hypermnesia and reminiscence in recall: A historical and empirical review. *Psychological Bulletin, 101*(1), 5. https://doi.org/10.1037/0033-2909.101.1.5

Poirier, M., & Saint-Aubin, J. (1995). Memory for Related and Unrelated Words: Further Evidence on the Influence of Semantic Factors in Immediate Serial Recall. *The Quarterly Journal of Experimental Psychology Section A, 48*(2), 384–404. https://doi.org/10.1080/14640749508401396

Pressley, G. M. (1976). Mental imagery helps eight-year-olds remember what they read. *Journal of Educational Psychology, 68*(3), 355–359. https://doi.org/10.1037/0022-0663.68.3.355

Pressley, M. (1982). Elaboration and Memory Development. *Child Development, 53*(2), 296–309. https://doi.org/10.2307/1128972

Pressley, M., & Levin, J. R. (1977a). Developmental differences in subjects' associative-learning strategies and performance: Assessing a hypothesis. *Journal of Experimental Child Psychology, 24*(3), 431–439. https://doi.org/10.1016/0022-0965(77)90089-3

Pressley, M., & Levin, J. R. (1977b). Task parameters affecting the efficacy of a visual imagery learning strategy in younger and older children. *Journal of Experimental Child Psychology, 24*(1), 53–59. https://doi.org/10.1016/0022-0965(77)90019-4

Reynolds, M. R., Niileksela, C. R., Gignac, G. E., & Sevillano, C. N. (2022). Working memory capacity development through childhood: A longitudinal analysis. *Developmental Psychology, 58*(7), 1254–1263. https://doi.org/10.1037/dev0001360

Rohwer, W. D., Jr, & Bean, J. P. (1973). Sentence effects and noun-pair learning: A developmental interaction during adolescence. *Journal of Experimental Child Psychology, 15*(3), 521–533. https://doi.org/10.1016/0022-0965(73)90102-1

Rohwer, W. D., Jr (1974). Elaboration and learning in childhood and adolescence. *Advances in Child Development and Behavior, 8*, 1–57. https://doi.org/10.1016/S0065-2407(08)60492-9

Rohwer, W. D., Raines, J. M., Eoff, J., & Wagner, M. (1977). The development of elaborative propensity in adolescence. *Journal of Experimental Child Psychology, 23*(3), 472–492. https://doi.org/10.1016/0022-0965(77)90040-6

Rosselet-Jordan, F. L., Abadie, M., Mariz-Elsig, S., & Camos, V. (2022). Role of Attention in the Associative Relatedness Effect in Verbal Working Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(11), 1571–1589. https://doi.org/10.1037/xlm0001102

Saint-Aubin, J., Ouellette, D., & Poirier, M. (2005). Semantic similarity and immediate serial recall: Is there an effect on all trials. *Psychonomic Bulletin Review, 12*, 171–177. https://doi.org/10.3758/BF03196364

Saint-Aubin, J., & Poirier, M. (1999). Semantic similarity and immediate serial recall: Is there a detrimental effect on order information? *The Quarterly Journal of Experimental Psychology Section A, 52*(2), 367–394. https://doi.org/10.1080/713755814

Schaefer, S., Riediger, M., Li, S.-C., & Lindenberger, U. (2023). Too easy, too hard, or just right: Lifespan age differences and gender effects on task difficulty choices. *International Journal of Behavioral Development, 47*(3), 253–264. https://doi.org/10.1177/01650254231160126

Siegler, R. S. (2016). Continuity and change in the field of cognitive development and in the perspectives of one cognitive developmentalist. *Child Development Perspectives, 10*(2), 128–133. https://doi.org/10.1111/cdep.12173

Souza, A. S., & Oberauer, K. (2018). Does articulatory rehearsal help immediate serial recall? *Cognitive Psychology, 107*, 1–21. https://doi.org/10.1016/j.cogpsych.2018.09.002

Souza, A. S., & Oberauer, K. (2020). No evidence that articulatory rehearsal improves complex span performance. *Journal of Cognition, 3*(1), 11. https://doi.org/10.5334/joc.103

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Superbia-Guimarães, L., & Cowan, N. (2023). Disentangling processing and storage accounts of working memory development in childhood. *Developmental Review, 69*, Article 101089. https://doi.org/10.1016/j.dr.2023.101089

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). Revisiting the attentional demands of rehearsal in working-memory tasks. *Journal of Memory and Language, 105*, 1–18. https://doi.org/10.1016/j.jml.2018.10.005

Thomas, J. J., Ayuno, K. C., Kluger, F. E., & Caplan, J. B. (2023). The relationship between interactive-imagery instructions and association memory. *Memory Cognition, 51*(2), 371–390. https://doi.org/10.3758/s13421-022-01347-6

Tse, C.-S. (2010). A negative semantic similarity effect on short-term order memory: Evidence from recency judgements. *Memory, 18*(6), 638–656. https://doi.org/10.1080/09658211.2010.499875

Tse, C.-S., Li, Y., & Altarriba, J. (2011). The effect of semantic relatedness on immediate serial recall and serial recognition. *Quarterly Journal of Experimental Psychology, 64*(12), 2425–2437. https://doi.org/10.1080/17470218.2011.604787

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133*(6), 1038–1066. https://doi.org/10.1037/0033-2909.133.6.1038

Varley, W. H., Levin, J. R., Severson, R. A., & Wolff, P. (1974). Training imagery production in young children through motor involvement. *Journal of Educational Psychology, 66*(2), 262–266. https://doi.org/10.1037/h0036277

Wagenmakers, E. J., Verhagen, J., Ly, A., Matzke, D., Steingroever, H., Rouder, J. N., & Morey, R. D. (2017). The need for Bayesian hypothesis testing in psychological science. In S. O. Lilienfeld, & I. D. Waldman (Eds.), *Psychological science under scrutiny: Recent challenges and proposed solutions* (pp. 123–138). https://doi.org/10.1002/9781119095910.ch8

Wolff, P., Levin, J. R., & Longobardi, E. T. (1972). Motoric mediation in children's paired-associated learning: Effects of visual and tactual contact. *Journal of Experimental Child Psychology, 14*(2), 176–183. https://doi.org/10.1016/0022-0965(72)90042-2

Yuille, J. C., & Catchpole, M. J. (1973). Associative learning and imagery training in children. *Journal of Experimental Child Psychology, 16*(3), 403–412. https://doi.org/10.1016/0022-0965(73)90003-9

Zhang, Z., & Yuan, K.-H. (2018). *Practical Statistical Power Analysis Using Webpower and R. Granger*. ISDSA Press.