Article:

Superbia-Guimarães, L. and Cowan, N. (2023) Disentangling processing and storage accounts of working memory development in childhood. Developmental Review, 69. 101089. ISSN 0273-2297

https://doi.org/10.1016/j.dr.2023.101089

Developmental Review 69 (2023) 101089

Contents lists available at ScienceDirect

Developmental Review

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

Review

Disentangling processing and storage accounts of working memory development in childhood☆

Luísa Superbia-Guimarães *, Nelson Cowan

University of Missouri, Department of Psychological Sciences, 210 McAlester Hall, Columbia, MO 65211, United States

ARTICLE INFO

Keywords:
Working memory
Cognitive development
Attention

ABSTRACT

Researchers have been asking the question of what drives the development of working memory (WM) during childhood for decades. This question is particularly challenging because so many aspects of cognition develop with age that it is difficult to disentangle them and find out which factors are causal or fundamental. In this review, we first prepare to discuss this issue by inquiring whether increases in storage, processing, or both are the fundamental driving factor(s) of the age-related increase in WM capability in childhood. We contend that by experimentally manipulating either factor and observing changes in the other, it is possible to learn about causal roles in WM development. We discuss research on school-aged children that seems to suggest, by means of such an approach, that the growth of storage is causal for some phases or steps in WM tasks, but that the growth of processing is causal for other steps. In our theoretical proposal, storage capacity of the focus of attention determines earlier steps of information processing by constraining the selective encoding of information into WM, whereas processing dependent on the focus of attention determines later steps, like the detection of patterns that can simplify the effective memory load and adoption of a proactive stance of maintenance in dual-task settings. Future directions for research are discussed.

Introduction

Humans undergo incredible changes in their cognitive abilities from infancy to adulthood (e.g., Cowan, 2022). Perhaps one of the most noticeable changes is the increase in working memory (WM) capability during childhood. WM, as we use the term here, is the ability to hold information in a heightened state of availability, readily accessible to conscious manipulation, during an ongoing task (Baddeley & Hitch, 1974; Cowan, 2017a). WM abilities predict many complex cognitive abilities and educational outcomes, such as reading, reasoning, mental calculation, problem-solving, and school achievement (Cowan, 2014; Daneman & Carpenter, 1980; Formoso et al., 2018; Gathercole, Pickering, Knight, & Stegmann, 2004; Hitch, 1978; Pickering, 2006; Siegel, 1994; Simms et al., 2018; Süß et al., 2002). Therefore, unveiling the sources of WM development in childhood is fundamental to better understanding how children learn and, ultimately, to improving educational practice.

Here we address one of the most difficult issues that has plagued researchers of WM development: that with so many changes taking place during development, it is difficult to figure out which changes are fundamental. In several previous articles, Cowan and

☆ This article resulted from an invitation by the editor for Nelson Cowan to submit a paper on his recent views, and Cowan asked Luísa Superbia-Guimarães to serve as first author.

* Corresponding author.
E-mail addresses: luisa.superbiaguimaraes@missouri.edu (L. Superbia-Guimarães), CowanN@missouri.edu (N. Cowan).

https://doi.org/10.1016/j.dr.2023.101089
Received 15 March 2023; Received in revised form 22 June 2023;
Available online 7 July 2023
0273-2297/

colleagues have addressed aspects of this issue. Cowan et al. (2002) illustrated the usefulness of finding measures that yield invariant results across different levels of task difficulty. Cowan (2016, 2017b) presented research in which capacity could be shown to increase with age even with certain confounding factors held constant (encoding speed, knowledge, attentional filtering ability). Cowan (2022) explored the opposite side of that coin, finding certain processing factors that can help account for increases in capacity. Here, we continue this line of inquiry by asking how we might determine whether the most fundamental basis of developmental growth in WM is a growth in storage, processing, or both. Most of the relevant research from which we have been able to disentangle processes involves tasks that can be administered from the early elementary school years through adulthood, so that is the focus of our review.

In the remainder of the article, we first attempt to summarize what is known about WM capability as measured by the number of items remembered. Then we discuss the debate about the roles of storage and processing in developmental improvements in WM capability. (We use the word capability for performance level in order to reserve the term capacity more specifically for when we wish to discuss storage rather than processing; we assume that WM storage capacity plus relevant mnemonic processing of the stored information, including processes that maintain and/or improve the stored representations, together yield the participant's WM capability.) Next, we examine studies that have focused on several types of potentially relevant mnemonic processing, in turn, as the basis of capacity development; and then we turn to other studies that instead favor storage capacity as the basis of developmental change. We propose, however, that the answer may not be uniform for all stages of a WM task. For example, it is possible that storage plays a critical role at encoding, with the scope of attention determining how well encoding focuses on the relevant stimuli, while processing is critical for maintenance, with the use of strategies to reactivate and elaborate the material critical in the post-encoding period before the memory test. Storage and processing nevertheless may both depend on the focus of attention as a common resource that increases or improves with development.

With that background in place, we describe our program of research in which we have tried to disentangle storage from processing and indicate areas of progress and priorities for future research to help settle this key question of the nature of development. Our efforts focus on development beginning in the elementary school years, using techniques derived largely from the adult literature. However, milestones of earlier infant and child development, where the same techniques cannot be applied, will be briefly described to provide further context to our work.

Describing the WM developmental growth that is to be explained

The developmental changes in WM culminate in an adult capability of around 7 verbal items when it is possible to use mnemonic strategies such as verbal rehearsal and chunking (as described by Miller, 1956, e.g., gathering items together into a remembered structure, such as remembering the letters in the acronym USA). When such strategies are prevented, only 3 or 4 separate objects or chunks are typically remembered (Cowan, 2001), such as in the retention of meaningless visual objects (e.g., Adam et al., 2017; Luck & Vogel, 1997); in the retention of verbal materials in a running span situation, in which the endpoint of the list is unclear (e.g., Broadway & Engle, 2010; Bunting et al., 2008); or when the verbal items are presented in a brief spatial array, and/or with rehearsal suppressed (e.g., Murray, 1965; Ricker et al., 2020). Both the larger and the smaller estimates of WM capability undergo changes from infancy to adulthood, in a quantitative increase in the amount of information that can be held and, as we will see, in a qualitative maturation in mechanisms involved.

WM capability increases markedly until adolescence and peaks around the age of 20 years and then declines gradually in old age (Bopp & Verhaeghen, 2005; Brockmole & Logie, 2013; Dempster, 1981; Hester et al., 2004; Simmering, 2012; Wingfield et al., 1988). Our review focuses on what occurs during the elementary school years, but we mention some of the infant literature to give a fuller picture of development.

A rapid increase in WM capability can be observed even during the first year of life, although the capability estimates are not necessarily comparable to the adult and children's data, given the experimental constraints of testing babies. Infant researchers have suggested that, in some contexts at least, babies demonstrate a WM capability of one item at the age of 6 months (Káldy & Leslie, 2005; Kibbe & Leslie, 2016) and that this effective capability reaches three to four items by the end of the first year of life (Kibbe & Leslie, 2013; Oakes et al., 2006; Oakes et al., 2009; Ross-Sheehy et al., 2003; for a review see Buss et al., 2018). However, as infant researchers may generally realize, tasks used with infants cannot provide capability measures that are fully comparable to child and adult data. The premise that one-year-olds can maintain between 3 and 4 items in WM would be surprising because it implies that they have the same capability observed in adults (Alvarez & Cavanagh, 2004; Awh et al., 2007; Cowan, 2001; Cowan, 2010; Cowan et al., 2004; Luck & Vogel, 1997; Todd & Marois, 2004). School-aged children (7- to 9-year-olds), on the other hand, show capability estimates ranging from 1.5 to 2.5 items in WM (Gilchrist et al., 2009; Cowan et al., 2010; Cowan et al., 2011; Cowan et al., 2015). Instead of implausibly suggesting that infants can already hold about 3 items in mind, but that WM regresses from infancy to childhood and later recovers its capability, Cowan (2016, 2017b) suggested that the infant tasks are necessarily less demanding. For example, infant procedures may not require memory of all features of the items to be retained, and the feature richness of representations appears to increase markedly in infancy. Demonstrating this early developmental change, Zosh and Feigenson (2012) tested the hypothesis of incomplete object representations in WM by surreptitiously swapping objects hidden in a box. If infants can hold three object identities in WM (rather than generic, non-individuated representations), they should continue to search for the original objects in the box after retrieving different objects from it. The results showed that 18-month-old babies continued to search the box only when 1 or 2 objects were hidden but stopped searching for the original objects at set size three. The authors suggest that 18-month-old infants can maintain up to three object representations in WM but, once this limit is reached, they often forget the objects' identities. Many issues are yet to be solved in infant research, but either the number of object files or the completeness of these files must develop in infancy (as discussed further by Cowan, 2016, 2017b).

Early studies on WM development in children focused on memory for verbal material and the acquisition of verbal strategies such as labeling and articulatory rehearsal. A hallmark review by Dempster (1981) showed a constant increase in verbal WM spans from 3 to 12 years, and this observation has been replicated many times. For instance, Cowan et al. (1999) showed that first graders had a mean span of 5.38 digits, fourth graders of 6.4, and adults of 7.4. Such spans were obtained in procedures allowing for verbal rehearsal. Two recent longitudinal studies including samples of about 13,300 (Reynolds et al., 2022) and 3,500 children (Ahmed et al., 2022) showed that verbal spans (forward and backward digits) grow in a curvilinear pattern during childhood. By the end of elementary school years, the function between digit span and age decelerates.

The WM capability for visuospatial information also develops during childhood, with a faster increase for static visual patterns (e.g., two-dimensional arrays) than for spatial sequences (Logie & Pearson, 1997), but both grow monotonically between the ages of 5 and 12 years. WM capability on simple tasks stabilizes at the beginning of adolescence, around age 12 (Luciana & Nelson, 2002), but the efficiency of its functioning is not optimized until at least around the end of adolescence (e.g., Gathercole, Pickering, Ambridge, & Wearing, 2004). Performance in tasks that are more reliant on the prefrontal cortex activity, such as those requiring participants to simultaneously maintain and manipulate spatial information (e.g., the reverse Corsi task), keeps improving until at least 15 years (Conklin et al., 2007). On the other hand, tasks that are supported by more posterior areas in the brain reach adult levels earlier: for instance, the recognition of verbal material and faces have been found to reach optimal performance at age 9 (Conklin et al., 2007), and recall performance of isolated spatial locations has been found to develop until the age of 12 years (Luciana et al., 2005).

Regarding the smaller estimate of WM capability, there is an ongoing debate in the field on whether WM is limited to a fixed number of slots to store chunks or whether its limits are flexible depending on the visual complexity of the items, the salience of its visual features, task goals, and attentional allocation during the task (Bays & Husain, 2008; Bays et al., 2011; Hardman & Cowan, 2015; Lilburn et al., 2019; Ma et al., 2014; Schneegans & Bays, 2016). The precision of visuospatial representations (Lilburn et al., 2019; Roggeman et al., 2014; Xie & Zhang, 2017) and the number of bindings (Oberauer, 2019) have also been pointed out as limiting factors of WM. For items fitting clear categories so that the precision of representations is not an issue, we believe that the pure capacity limit of WM is about 3 to 4 chunks that can be simultaneously represented in the focus of attention (Cowan, 2001).

In school-aged children, the corresponding measures of this capacity have been shown to be lower, with estimates ranging from 1.5 to 2.5 chunks and/or single features (Gilchrist et al., 2009; Cowan et al., 2010; Cowan et al., 2015; Forsberg, et al., 2022). For instance, in memory for unrelated sentences, seven-year-olds were able to access around 2.5 clauses when presented with lists of four shorts sentences, as measured by the correct recall of at least one content word from a sentence (Gilchrist et al., 2009). When presented with colored shapes in spatial arrays, seven-year-olds were found to remember about 1.5 items in a static presentation condition (Cowan et al., 2010) and 1.9 items in a sequential presentation condition (Cowan et al., 2011).

Interestingly, despite their lower capability estimates in a color-recognition task, seven-year-olds have *meta*-cognitive reports of remembering three colors – an estimation that remains roughly constant through elementary school years to adulthood (Forsberg et al., 2021). Although the estimated number of items in WM increased from about one (first graders) to three items (college students and parents), Forsberg et al. found that the *meta*-cognitive reports remained stable around 3 to 4.5 items, depending on the presented set size, across all age group. Such *meta*-cognitive judgments may reflect the sense of knowing how many items were memorized in some way, rather than specifically remembering the probed item in sufficient detail.

The discrepancy between *meta*-cognitive judgements and capability estimates in school-aged children suggest that, although they cannot reliably report memory for all the items in the array, they know that *something* was encoded into WM. Forsberg et al. (2022) found that, when a more lenient criterion is adopted and the calculations were based on an estimate of the number of objects for which at least one feature is known (location, color, orientation, or more than one of these), the capability estimates were 4.2, 5.1, and 5.6 items for, respectively, children (6–7 years), adolescents (11–13 years), and adults (18–25 years). This calculation based on any-one-feature may reconcile the paradoxical results between infant and child studies: it could be that the traditional method of probing one feature (e.g., a color probe) might underestimate children's capability in terms of the number of objects, whereas they know something about the encoded objects, but not necessarily the probed feature that allows the question to be answered correctly. Alternatively, children may be aware of information they fleetingly had in the focus of attention but may be unaware of the effects of interference that may remove some of that information by the time of the test.

Halford et al. (2005) proposed that the number four was a constraint on cognitive growth in another way. They demonstrated that adults can only process a maximum of four variables at once, to be associated with one another, while solving reasoning problems. An example of the inter-association between only two variables is when one keeps in mind that a silver key works for the front door to the office building only, whereas a gold key works for one's office but not the front door. In examining adults' interpretation of statistical interactions between variables, Halford et al. found that the practical limit in understanding was three variables. According to Halford's (1993) theory, human reasoning relies on relational mental models that are learned via experience, and these mental models serve as structures to represent information in the mind (e.g., categorical knowledge, inclusion and exclusion relations, hierarchical ordering). Children integrate more dimensions into mental models as they grow, going from understanding unary concepts (e.g., class membership, Mouse[Mickey]) and binary concepts in early childhood (e.g., larger than[lion, mouse]), to ternary concepts (e.g., addition[$2 + 2 = 4$]) at age 5, and quaternary concepts (e.g., proportion($1/2 = 4/8$) at age 11 (Halford et al., 1998). The interesting observation that the maximum number of variables in a mental model approximately corresponds to the estimated adult capability of WM (Halford et al., 2007) could occur because we remember the current contents of the focus of attention through associations between the items.

In sum, WM capability increases markedly in childhood and even in the first year of life. This increase is observed in the verbal and visuospatial domains, for simpler, unitary features (colors, shapes), but also for more complex, bound representations. During school years, span measures increase monotonically until adolescence and finally reach the adult level of about 7 items when rehearsal and grouping are possible and otherwise only 3–4 items in WM. Next, we will address how the processing of information is optimized during childhood, with a focus on how attention plays a role in such development.

WM development in childhood: A matter of storage, processing, or both?

There is systematic improvement in various measures of WM across development from infancy to young adulthood, (Cowan, 2022; Gathercole, Pickering, Ambridge, & Wearing, 2004). Understanding why WM develops is important both theoretically and practically. An impediment to doing so, however, is that so many different traits improve with development that it is difficult to disentangle them. Knowing which traits change fundamentally with age and which ones are secondary consequences of the fundamental changes is essential to unveil the sources of WM development in childhood. In this review, we ask about the causal pathways between WM storage and the processing of WM information, both of which appear to improve with development. To this end, we adopt the definitions of *storage* as the persistence of information in WM and *processing* as an operation upon the memory representation in some way – even if that operation is just the recurrent activation of a mental representation by means of rehearsal and/or refreshing, without necessarily modifying the features or identity of the representation in itself. Note that our definition of processing accommodates both the use of maintenance strategies that are supposed to simply strengthen or stabilize memory traces (e.g., refreshing, consolidation) without altering the nature of the representation (Camos et al., 2018, Jolicœur and Dell'Acqua, 1998), and those that somehow modify it (e.g., by integrating information into it, as when using semantic elaboration, or by chunking items) (Miller, 1956; Pressley 1982). The allocation of attention and the removal of items from WM are also phenomena that are included in our definition of processing. Operations are made to information in WM in all cases: in the first example, the strength/stability of memory traces are changed; in the second, features of the representation are enriched; in the third, a subset of the content is prioritized.

Through these definitions, we propose that changes in storage and processing subsume all accounts of WM development; and that these two overarching factors underlie all the developmental changes in WM happening during childhood. Table 1 exemplifies different situations that fall within the definitions of storage and processing adopted in this article.

It is theoretically possible that the mechanisms to hold information passively in WM do not change with development, but that the observed WM measures show increases because of improvements in mnemonic *processes*. These processes can be used to increase WM representation of stimuli or ideas by refreshing the representations (Barrouillet et al., 2011), by improving them (Ricker & Vergauwe, 2022), or by protecting them from interference (Oberauer et al., 2012). Alternatively, it is possible that WM *storage* increases for fundamental reasons, for example, due to biological maturation of the brain regions involved in storage, and/or the increase in knowledge and improvement in understanding of information being stored. Then it is further possible that increases in WM storage underlie optimization in processing. Specifically, the storage increases could allow developing children to have better mental tools for such processes, for instance by holding in mind the task goal (Kane & Engle, 2003; Marcovitch et al., 2010), remembering mnemonic strategies that can be used, and executing those strategies. For example, what appears to be an increase in the speed at which attention can be used to refresh items in WM could actually result from refreshing all items simultaneously in a capacity-limited WM, with increases in the storage capacity resulting in a faster effective rate at which the whole set of items can be refreshed (Lemaire et al., 2018). That possibility would be an instance of the view that storage increases promote optimization in WM processing. In research on children from the early elementary school years through adulthood, we see that both storage and processing can be viewed as primary sources of WM development in different situations, and that each one has an effect on the other in particular ways.

Different stages in a WM trial can rely on storage and processing factors. For example, imagine an experimental situation in which a child is given the instruction "When I say go, put the camera, the turtle, and the globe into the box" and waits for the command "Go!". In a first stage, the relevant information must be encoded and stored in the focus of attention. In a second stage, the child might be presented with distractions, which could be accidental events in the testing room, deliberately placed distractions in the task materials, or even information that automatically occurs to the participant (through mind-wandering from the task). These must be suppressed for optimal performance to be achieved. In these two initial stages, storage capacity (here understood as being determined by the size of the focus of attention) might restrict the ability to selectively encode task-relevant information and filter out distractors. In a third stage, the task-relevant information (camera, turtle, globe, wait until you hear "Go!", put into the box), may have to be briefly held in the focus of attention, all at once or several at a time, in order for them to be encoded in a way that can be remembered when the response is required. In a fourth stage, it may be possible for the participant to simplify the materials by combining items, in this case

Table 1
Definitions and examples of storage and processing in WM.

Concept	Definition	Examples			
Storage	The persistence of information in WM.	Storage of items, e.g., sequences of objects, letters, words, locations.	Storage of goals, e.g., task instructions.		
Processing	Operation upon representations in WM in some way.	Allocation of attention, e.g., prioritizing some perceptual stimuli and/or mental representations over others.	Activities for the maintenance of the stored information, e.g., implementation of rehearsal, refreshing, the recoding of the memorized material, chunking items based on long-term knowledge.	Improving the encoding through pattern detection to assist storage, e.g., chunking items.	Carrying out a separate task upon demand, e.g., in a dual-task setting, alternate between two tasks and change the contents of WM accordingly.

creating the possibly novel imagined event in which the turtle is using the camera. This newly imagined event can be memorized well enough so that it can free up the focus of attention, possibly being stored as a newly-learned, still-active entry into long-term memory. With this freed-up memory, the participant can continue to think of other things, such as the last item to be remembered, the globe. Eventually, the imagined event can be built up further, for example, if it is imagined that the turtle with the camera is taking a picture of the globe. Once the material is encoded in the best way that the participant is able, this information may still have to be processed further to prevent it from becoming unavailable during a delay until it is time to use it. The child might continue to imagine the event, refresh the activity level of that memory, or might rehearse it in verbal form. The success of correctly placing the items inside the box upon hearing "Go!" depends both on capacity (the size of the focus of attention) and on the way the information is processed (imagining a novel pattern between the items and maintaining it until the end of the trial).

There might be limits in capability that apply to more than one stage of the above example. The main manner in which we suggest that these various capabilities can be examined is by placing extra demands on one factor (storage or processing) and observing alterations in the other one. If some capability to store information is fundamental, then altering the storage demands should have an influence on what happens downstream in processing; with more items to be remembered, the result can be that less of that storage capability is available to support processing. Conversely, if processing is more fundamental, then altering the processing demand may influence how many items can be stored. Last, if some third factor such as neural integrity has an across-the-board influence on both storage and processing, then both directions of influence can take place. We would suggest, however, that the answer need not be uniform across all steps of the task, such as the steps suggested in the above example. It is possible, for example, that for encoding, storage is fundamental whereas, for maintenance, processing is fundamental.

Developmental researchers have asked for a long time what are the sources of cognitive development in children, and much focus has been put on disentangling storage limitations from processing limitations to explain lower WM capability in children (a review by Chi, 1976, illustrates this point). Burtis (1982) noted that one can distinguish between an underlying structure constraining WM capacity (i.e., a short-term storage space) and its manifest performance (i.e., any given measure of WM performance, such as spans), the latter being influenced by factors other than storage space *per se*. The development in WM performance during childhood could be a result of a genuine capacity increase, an optimization in control processes, or a combination of both. The first theory we consider predicts that optimization on processing drives the increase of WM storage. The next theory is the opposite, that storage growth during childhood leads to an improvement in how children process information. The third theory states that there is a general growth in both the processing and storage functions of WM, with children becoming more proficient in allocating resources to these two functions as they grow.

Is the development of processing in WM tasks causal?

Processing is a generic term referring to how information is filtered out, encoded into, selected, prioritized, and maintained in WM. Attention and long-term knowledge play a pivotal role in how information is processed in WM. Specifically, in the embedded-processes view, items consciously represented in WM are maintained in the focus of attention in a prioritized state compared to other mental representations in long-term memory and other sources of information from the environment (Cowan, 1988, 2019; Cowan et al., 2021). The scope and direction of the focus of attention are controlled by a group of processes known as the central executive (as in Baddeley, 1986), which implements voluntary control over the direction of attention and various processes. In a WM task, the focus of attention must contain not only items to be recalled (at least when they are first encoded); it also must contain any task goals and procedures that have not become habitual. Therefore, within the embedded-processes view, the problem of the development of processing is one relating to the limits of attention and its operating mechanisms in WM.

A large stream of work suggests that WM develops because its processing efficiency is optimized throughout childhood. In general terms, the processing account holds that storage capacity (of the focus of attention and otherwise) is constant throughout development and that the efficiency of its inner mechanisms determines how many items can be stored, with children becoming more proficient in a series of control mechanisms in WM as they grow. The term "efficiency" is rather vague in this context and can refer to the speed of mental processes (Case et al., 1982; Kail & Park, 1994), the ability to use attention to select relevant stimuli among distractors (Zukier & Hagen, 1978; Plebanek & Sloutsky, 2019), the control of executive attention to manage tasks (Engle & Kane, 2004; Kane et al., 2007), or even the implementation of maintenance strategies (Camos & Barrouillet, 2011; Chi, 1976; Lehmann & Hasselhorn, 2007). In all cases, it refers to the use of processes to get the most out of a fixed amount of storage. We will break down this section to cover all these purported sources of development in processing efficiency in order to provide the reader with an overview, starting with attentional mechanisms that are a core interest in our research agenda.

Attention and its control

Attention is a multi-faceted human ability that serves many functions. These include maintaining the level of alertness (Davies & Parasuraman, 1982), selecting information from sensory input (Broadbent, 1958), and modulating behavioral responses (Berger et al., 2018; for a review on the properties of attention see Chun et al., 2011). We will discuss exclusively the aspects that are relevant for the argument to be developed later, that is, selectivity and the executive control of attention.

Selective Attention. Selectivity is a property of attention that rapidly develops in infancy and modulates processing in WM. Selective attention is the ability to focus on task-relevant stimuli (or stimulus features) and ignore distractors. From a WM perspective and within the embedded-processes view (Cowan, 1988; Cowan et al., 2021), such operation corresponds to the allocation of mental representations in the focus of attention, either by means of voluntary executive control or by exogenous, stimulus-driven attentional recruiting.

Research on selective attention in infants and young children traditionally used binary decision tasks, speeded classification, search tasks, selective listening (shadowing), and incidental learning paradigms (Lane & Pearson, 1982). These methods share the common rationale that, in the presence of distractors (i.e., irrelevant stimulus features and/or input channels), responses will be more error-prone, slowed down, and subject to intrusions from the irrelevant dimension/channel (that is, failure to selectively ignore distractors will hamper performance).

Typically, studies in selective attention show that children have worse filtering abilities than adults and that such ability improves with age. In one of the first developmental studies on interference, Comalli et al. (1962) showed that reaction times in a classic Stroop task decreased linearly from age 7 and peaked around 17–19 years, providing evidence that the ability to filter out interference from the task-irrelevant dimension (here the printed word) improves during childhood and adolescence despite greater familiarity with the words. Similar developmental effects have been found in a procedure in which irrelevant speech can disrupt memory for a printed list (Elliott, 2002). In speeded classification tasks like sorting cards, reaction times become faster and the interference caused by irrelevant dimensions drastically decreases between ages 6 and 12 years – a reduction of about 5 s when there is only one irrelevant dimension and of 7 s when there are two irrelevant dimensions to filter (Strutt et al., 1975). Moreover, there is a developmental trend toward analytic processing of stimulus dimensions; that is, children get better at selectively attending to individual dimensions of a stimulus as they get older (Thompson, 1994; Ward, 1980). The improvement in selective attention is also observed in the ability to separate channels and sensory modalities: in doing a continuous matching task regarding the posture of a stick figure, kindergarteners were heavily disrupted by distractors, especially when these were separated from the target (tones or a border around the stick figure, worse than by the irrelevant color of the figure itself). Second-graders were less distracted, especially by the tones; and fifth graders are minimally distracted both by any of the distractors (Smith et al., 1975). With age, children improve their ability to attend selectively to input channels and stimulus dimensions.

Increasing selectivity in childhood is accompanied by increases in recall of task-relevant material, whereas the incidental recall of task-irrelevant information remains relatively constant or declines across elementary school years (Doyle, 1973; Hagen, 1967; Hagen & Hale, 1973; Maccoby & Hagen, 1965; Zukier & Hagen, 1978). For instance, in the study by Maccoby and Hagen (1965), children were instructed to memorize the background color (the relevant information) of picture cards in the presence and absence of an auditory distracting task (monitoring musical tones). After being tested on the background colors, they were tested on the identity of the pictures presented in the sets of cards (irrelevant information). Recall of the background colors increased linearly from first grade to seventh grade but recall of picture identities remained constant until fifth grade and decreased from fifth to seventh grade. This pattern suggests an increase in the storage of relevant information until around age 10 (illustrated by better recall of the background colors by older children), followed by an improvement of selectivity at the beginning of adolescence (illustrated by the decrease in the incidental recall of picture identities). The auditory distractor task impaired recall of task-relevant information in all age groups, but this deleterious effect was vastly diminished from third grade onwards. This can suggest either that children improve in protecting task-relevant information from interference (Dempster, 1992; Leslie, 1975) or that they become better in switching attention between a distracting processing demand (i.e., the tone monitoring) and the maintenance of task-relevant mental representations (i.e., the colors) (Barrouillet et al., 2009).

The control of attention is evoked by researchers to explain individual differences in WM capability (Engle & Kane, 2004; Kane et al., 2007), and, therefore, its developmental increase. Research with adults suggests that individual differences in visuospatial WM are accounted for by the ability to filter out irrelevant stimuli at encoding, that is, selective attention (McNab & Klingberg, 2008; Vogel et al., 2005). The argument goes as follows: The filtering ability determines which information will be encoded into WM and allotted to a fixed amount of slots. For instance, given a storage capacity of about 3 slots and while presented with 6 items, the efficiency in filtering out the two irrelevant items according to the task's goals will predict performance. If attentional filtering is flawed and irrelevant items are encoded, then task-irrelevant information will be processed and stored in WM, producing recall errors. Following the reasoning applied to adult studies, one can assume that age-related individual differences in WM stem from the fact that, as they grow, children get better at using attention to prioritize task-relevant items and filter out distractions, therefore saving WM storage space.

Selective attention indeed improves during childhood (Plude et al., 1994), an improvement that accompanies the increase in memory capability estimates (Plebanek & Sloutsky, 2019). The overload of unfiltered information in children's WM could cause task-irrelevant information to be processed (Plebanek & Sloutsky, 2017), which leads to more inter-item interference that is deleterious to recall of relevant information (Oberauer & Kliegl, 2001, 2006; Shipstead & Engle, 2013). The fact that selective attention predicts WM functioning in early childhood (Veer et al., 2017) makes the theory that the improvement of selectivity drives WM development appealing, but adult results have shown that the argument can go both ways (Poole & Kane, 2009), with either selectivity or WM capacity predicting the efficiency of the other function.

Veer et al. (2017) designed a longitudinal study in which children were tested at the age of 2.5 and 3 years in short-term memory WM, selective attention, and response inhibition. The short-term memory task consisted of remembering the location of boxes hiding toys. The WM task consisted of finding toys hidden in boxes, while keeping track of which boxes had already been searched. The selective attention task was a simple computerized visual search (finding elephants among horses and bears). Finally, the response inhibition task was a delay gratification test, in which children had to refrain from touching an attractive gift wrapped in multi-colored paper and ribbon. Path analysis showed that selective attention at age 2.5 years predicted WM and response inhibition at age 3, the correlation between selective attention at age 2.5 and WM at age 3 being 0.28, and between selective attention at age 2.5 and response inhibition at age 3 being 0.17 after controlling for prior levels of WM and response inhibition. The correlation between WM at age 2.5 and selective attention at age 3 was only 0.15. The authors suggest that selective attention in early childhood predicts future WM capability, not the opposite.

Despite findings that attentional selectivity develops with age, many results conflict with the view that selectivity *per se* is the driver of WM development. For instance, the recall of task relevant information (e.g., words from a shadowed auditory message, Doyle, 1973) increases monotonically until around age 10, but the incidental recall of irrelevant information (e.g., words from concurrent speech, Doyle, 1973) remains constant and only starts to decrease after that age. This developmental pattern has been observed in studies using verbal auditory information (Doyle, 1973); visual information (Hagen & Hale, 1973); and cross-modal information (Maccoby & Hagen, 1965). Such a pattern suggests that what improves with age is not necessarily the ability to filter out distractors in a constant capacity WM system – otherwise, incidental recall should also diminish monotonically as the recall of relevant information increases.

To summarize, the view that better attentional allocation drives better WM performance is a popular one in the field. Research on selective attention and WM seems to suggest that, although selectivity is fundamental for the functioning of WM and it does increase with age, it cannot fully account for individual differences, including age.

Executive control. Executive attention is conceptualized as the ability to organize the processing of information around objectives, a set of mental actions that require the intentional maintenance of relevant information, disengagement from task-irrelevant information, and resistance to distractors (e.g., Baddeley, 1986; Miyake et al., 2000; Shipstead et al., 2016). According to an influential view in the literature, individual differences in controlled attention determine both WM capability and fluid intelligence, therefore being a common underlying factor of these two higher-order cognitive functions (Engle & Kane, 2004; Kane et al., 2004; Kane et al., 2007; Mashburn et al., 2020; Shipstead et al., 2015). The executive control of attention also appears to develop with age (Atkinson & Braddick, 2012; Chevalier & Blaye, 2016; Jones et al., 2003), contributing to the development of WM. Some of these authors may consider executive attention to be coextensive to WM capability. In children, studies using the free-play method[1] have shown that the abilities to focus attention and resist distractors start to be observed around the age of 9–10 months and develop during the first four years of life (Kannass et al., 2006; Ruff & Capozzoli, 2003). Controlled attention in some circumstances continues to develop between the ages of 4 and 8 years and stabilizes around age 10 (Rueda, Fan et al., 2004; Rueda, Posner et al., 2004). In other circumstances, such as strategic self-organization, executive function continues to develop into late adolescence (Luciana et al., 2005).

In sum, improvements in selectivity and executive control are observed from infancy through adolescence. The development of such components of attention is intertwined with the optimization of processing in WM, making attention itself a potential candidate to explain the age-related growth of WM capability. It is possible, for instance, that WM capability increases because selectivity improves during childhood, preventing WM from being cluttered with distractors (Veer et al., 2017; Vogel et al., 2005); or that improvement in the executive control of attention prevents interference-based forgetting or goal neglect (Conway et al., 2003; Engle, 2002).

Next, we will address two other factors that also develop in childhood and influence how information is processed in WM. These are knowledge and processing speed.

Application of knowledge

Knowledge advances with age, as children form more extensive long-term memory representations. Here we consider the application of knowledge to organize and simplify the contents of WM to be a process because it reflects some sort of transformation of the material held in WM (see definition and examples in Table 1). For example, if one already knows the acronym USA (United States of America), noticing this knowledge greatly helps in remembering what the sequence of letters is, compared to when one does not know. The until-then meaningless sequence "U-S-A" is modified as soon as one notices that it matches a long-term representation: at this point, it becomes a meaningful chunk in WM. Moreover, the intentional application of knowledge can be used to group and organize the contents of WM, for example when someone tries to memorize the items in a visual display by imagining a familiar pattern (e.g., a square, a triangle, and a house can form a house with the roof and a chimney on top). The application of knowledge to the material in this case also changes the content of WM, and we consider it to be a process.

Long-term representations can bootstrap WM performance by means of familiarity (Jackson & Raymond, 2008; Ngiam et al., 2019), redintegration at recall (Schweickert, 1993), and by providing a resourceful knowledge base for the creation of semantic associations (Forrester & King, 1971; Poirier & Saint-Aubin, 1995). It is therefore not surprising that memory spans are influenced by knowledge acquired by children as they grow. The observation that lexicality effects (i.e., better recall of words compared to non-words) increase with age (Turner et al., 2000; Turner et al., 2004), for instance, provide evidence in favor of the idea that knowledge drives WM development.

It is hard to draw a fine line between effortful versus automatic use of knowledge in WM by children. For example, redintegration (i.e., the reconstruction of verbal items at recall), a process considered to underlie lexicality effects, has been proposed to be a rather automatic influence of long-term phonological representations in speech perception and production during the response phase of a task (Hulme et al., 1997; Schweickert, 1993). One line of evidence for this view is the fact that lexicality effects are greater in tasks requiring speech production at response (e.g., probed recall tasks) than in tasks that do not require it (e.g., serial order recognition) (Gathercole

[1] The free-play method consists of letting babies and toddlers free to play with a toy of their preference, whilst exposed to competing distracting play options. The time spent playing with a single toy is then measured and considered an index of focused attention. Note that such an experimental task is consistent with the aforementioned definition of executive attention in a sense that it requires children to control processing toward a play objective. Interestingly, though, rapid rather than slow infant habituation of gaze toward an object has been shown to be a predictor of later cognitive function (Kavšek, 2004), suggesting that deliberate versus captured attention operate differently, with only the former, reflected in free play, indexing executive function (see Cowan, 1988).

et al., 2001; Turner et al., 2004; but see Jarrold et al., 2008 for a discussion on the sensitivity of these procedures in detecting those effects in younger children). On the other hand, redintegration has also been proposed as the basis of phonological rehearsal (Turner et al., 2004), a maintenance process that is subject to volition. Therefore, it is hard to pinpoint whether the increase in lexicality effects during childhood results from an automatic influence of knowledge in later stages of information processing (i.e., redintegration at recall) or in controlled intermediate stages (i.e., during the use of phonological rehearsal for maintenance purposes). Knowledge is helpful to WM performance in itself, but attention enhances its use during tasks. For the purposes of this review and in agreement with our proposed definition of processing (i.e., changes to the content of WM, in some way), we will focus our discussion only on the effortful use of knowledge in service of WM maintenance.

In an often-cited study, Chi (1978) demonstrated that knowledge has a heavy impact on WM performance, by showing that the span for chess moves was much higher in children (mean age 10.5 years) who were expert chess players than in adults. Memory for digits was nevertheless superior in these adults than in the children, a result that suggests that specific knowledge about the test material positively influences span measures. According to our definition of knowledge as a process, one could assume that the expert chess players benefitted from knowledge because they were somehow applying it to organize the test material. For example, chess experts could be mentally labeling the names of the chess pieces and relying on specific chess rules to reduce the material.

A follow-up of the chess study conducted by Schneider et al. (1993) included experts and novices in both adult and children group; the results showed that the advantage of child experts over adult novices was maintained even when asked to recall random compositions on the chess board (i.e., pieces arranged in violation of chess rules). However, when tested on a control condition with colored geometrical wooden pieces (yellow and blue cubes, cylinders, spheres, cones, etc.) positioned on a colored checkerboard, the effects of expertise (experts > novices) and age (adults > children) on immediate recall were abolished. Interestingly, experts of both age groups improved at a faster rate than novices in subsequent trials, but the improvement was greater among adults nevertheless. This latter finding suggests the existence of maturational effects in WM other than simple knowledge acquisition during childhood, and that adults have greater skills to use knowledge in favor of learning new material. |In sum, the results of Schneider et al. (1993) suggest that knowledge is no longer useful to WM performance once the memoranda are replaced by a novel (but comparable) version of the test material, but it can nevertheless bootstrap the learning of the test material.

In another classic study, Schneider et al. (1989) measured the level of expertise in soccer (knowledge about soccer rules) and reading aptitudes of third, fifth, and seventh graders. The participants were divided into soccer experts and novices, depending on their knowledge, and tested on text comprehension and memory for a story describing a soccer game. Again, the results showed that soccer experts had a better memory than novices regardless of age; critically, the effect of knowledge overcame the influence of reading aptitude, with low-aptitude experts outperforming high-aptitude novices in measures of memory and text comprehension.

However, knowledge alone cannot account for all of the WM development in childhood. Cowan et al. (2015) designed a study in which the influence of long-term memory on recognition was minimized in a condition presenting unfamiliar characters in a visual array. The authors compared memory for arrays of English letters (familiar long-term representations) and unknown characters (unfamiliar stimuli) across different age groups from first grade to college. After the exclusion of children in the youngest group who did not know the English alphabet well, the results showed a steady increase in capability both for English letters and unfamiliar characters, following the same developmental pattern. More importantly, when the measures were normalized (z-score transformed) within conditions but across age groups, the developmental curves for English letters and unfamiliar characters were nearly the same. Although memory for familiar stimuli is overall better than memory for unfamiliar stimuli, the rate of development for both types of information was about the same. This indicates that the increases observed in capability were driven by a factor other than knowledge of the English letters *per se*. In sum, knowledge improves WM performance in children and adults, but it does not seem to be the driving factor of WM development during childhood. Had it been the case, the developmental curve for the familiar stimuli in Cowan et al. should have been more accelerated than the one for unfamiliar characters. Thus, though knowledge develops and clearly influences WM, it cannot be the sole account of WM development.

Processing speed and efficiency

In accounting for WM span, processing speed has had some successes and some failures. For Case et al. (1982), the process was the speed of identifying items to be used in the WM test, and the suggestion was that faster speed produces more efficient memory in some way. Reducing adults' speed by making the items unfamiliar also commensurately reduced their span, so that a single linear function of speed versus span resulted across a wide range of age groups (3 years to adulthood).

For others, the process of interest was verbal rehearsal. The rehearsal speed notion (from Baddeley et al., 1975) was that information in WM decays and is lost within several seconds if it is not reactivated quickly enough through rehearsal. Again, the relation between speed and span was found to be linear across age groups (e.g., Hulme & Tordoff, 1989; Kail & Park, 1994). Flavell et al. (1966) already had shown that span measures increase with age (recently replicated by Elliott et al., 2021, showing that even many 5-year-olds can already use some rehearsal). Between ages 8 and 10 years, children progressively shift from simply labeling the items to adopting cumulative rehearsal, especially for earlier positions in a list (Guttentag et al., 1987; Lehmann & Hasselhorn, 2007). This idea was largely reflected by research in the 1970s and the 1980s suggesting that learning and intellectual disabilities were caused by verbal production deficits that could be remediated by teaching children how to rehearse (Bray & Turner, 1986; Brown, 1974; Flavell, 1970; Leslie, 1980). The extent of use of cumulative rehearsal by children is predicted by their WM spans and children become faster in articulating, which is taken by researchers as evidence that rehearsal efficiency is a driving factor of the increase in WM capability during childhood (Hitch et al., 1989; Kail & Park, 1994; Lehmann & Hasselhorn, 2007).

Still others (Camos & Barrouillet, 2011; Cowan, 1992; Gaillard et al., 2011) have suggested that a process of using attention to refresh items before they decay may be involved. In this line of studies, attention is thought to alternate between storage and processing functions in WM and used to counteract decay when it is free from processing requirements. In complex span tasks, the experimental manipulation of the interval between storing and processing information – either by means of changing its total duration, or by increasing the number of items to process – has yielded evidence that the efficiency of the attentional switch determines the span, a phenomenon known as the cognitive load effect (Camos & Barrouillet, 2011; Gaillard et al., 2011, but see also Ricker & Vergauwe, 2022, for a different view on the matter). The observation that the cognitive load effect increases with age (i.e., adults are more affected by experimental manipulations hampering the use of attentional refreshing than children) is interpreted as evidence that children become more efficient in using attention to refresh items as they grow (Barrouillet et al., 2009).

From all of these viewpoints, being it rehearsal or refreshing, linear relations have been observed between the mean speed of processing and the mean memory span as a function of age. Cowan et al. (1998) found that rehearsal speed and memory search speed (indicated by the rate of recall) did not correlate with each other but together accounted for the developmental change in digit span; once more, speed and span are found to be closely related (Camos & Barrouillet, 2011). Gaillard et al. (2011) found that equating the rate of refreshing and related processing eliminated age differences. The question about the primacy of speed, however, is whether other processes were affected when speed was adjusted.

Results from adult studies and using the computational modeling approach cast doubt on the idea that rehearsal alone can account for higher spans, and even improve recall performance (Souza & Oberauer, 2018; 2020). In such studies, training adult participants to use cumulative rehearsal increased the rate of rehearsal during the tasks, but it did not promote benefits in recall performance (Souza & Oberauer, 2018; 2020). These results render the proposal of the optimization of rehearsal as being the main drive of WM development problematic, given the absence of an improvement in performance even in mature subjects.

Several studies have attempted to examine the speed-span relation with experimental interventions. Hulme and Muir (1985) attempted to speed up rehearsal during list presentation but were unable to get it to speed up. Cowan, Elliot, et al. (2006) manipulated the speed of spoken recall by children and adults by presenting the stimuli (lists of digits) at different rates (Experiment 1) and by training children to repeat the digits at the same rate as adults (Experiment 2). In Experiment 1, although faster presentation rates yielded faster recall rates (i.e., digits spoken per second), the span measures of both children and adults remained unaffected. In Experiment 2, although children were successfully trained to produce their responses at the same speed as adults, faster recall rates were not accompanied by an increase in span. The results suggest a dissociation between the rate of recall and memory capability and speak against the premise of speed optimization in childhood. Processing speed might be a result, rather than a cause, of developmental change in capacity (Lemaire et al., 2018). Cowan et al. (1998) found two different speeds (retrieval and rehearsal) that together accounted for the development of digit span, but did not correlate with each other, suggesting that speed is at least not a primary, unitary causal trait. Cowan, Elliot, et al. (2006) induced children to recall digit lists at a faster rate equivalent to adults, and yet saw no change in span.

Table 2
Summary of findings suggesting that processing is the causal factor of WM development.

Type of processing viewed as a causal factor	Rationale for the explanation	Summary of evidence observed
Selective attention	The ability to focus on task-relevant material and filter-out distractors at encoding develops in childhood, causing WM to be progressively less cluttered with task-irrelevant information.	The irrelevant-speech effect decreases during elementary school years (Elliott, 2002); recall of task-relevant material increases while the incidental recall of distractors decreases in childhood (Doyle, 1973; Hagen & Hale, 1973; Zukier & Hagen, 1978); the efficiency of attentional filtering predicts WM performance since early childhood years (Plebanek and Sloutsky, 2019; Veer et al., 2017).
Executive control	In this view, the executive control of attention is seen as coextensive with WM functioning. As they get older, children improve their ability to intentionally maintain task-relevant information and resist distraction, therefore improving WM performance.	Individual differences in controlled attention predict WM capacity and fluid intelligence in adults (Engle & Kane, 2004; Kane et al., 2004; Kane et al., 2007; Mashburn et al., 2020; Shipstead et al., 2015); the executive control of attention develops in childhood and stabilizes around age 10 (Atkinson & Braddick, 2012; Chevalier & Blaye, 2016; Jones et al., 2003; Rueda, Fan et al., 2004; Rueda, Posner et al., 2004).
Application of knowledge	Knowledge allows one to organize and simplify the contents of WM; growth in the knowledge base provide children with greater resource to do so.	Specific knowledge about the memorized material overcomes age effects on recall performance (Chi, 1978; Schneider et al., 1989; Schneider et al., 1993).
Processing speed and efficiency	Faster processing speed produces more efficient memory. The processes of interest vary from item identification, to verbal rehearsal, and attentional refreshing.	Reducing adults' processing speed to match children's equalized the memory span between the two age groups (Case et al., 1982). With age, children become faster at using verbal rehearsal, and increases in recall are accompanied by increases in the speech rate (Hitch et al., 1989; Hulme & Tordoff, 1989; Kail & Park, 1994; Lehmann & Hasselhorn, 2007). Children improve in using attention to reactivate memory representations in-between processing intervals in dual-task settings (Camos & Barrouillet, 2011, Gaillard et al., 2011).

Summary of processing findings

In sum, there are clear effects of the control of attention and the application of knowledge on WM performance, and aspects of attention and knowledge clearly grow with child development. These findings are summarized in Table 2. It remains possible, however, that there are effects of storage capacity of the focus of attention on the ability to carry out processes, given that one must often store goals and procedures while carrying out the processes. It is the prospect of the primacy of storage that we examine next.

Is the growth of WM storage causal?

Many theories of WM development stressed the role of an increasing storage capacity of the focus of attention or a similar mechanism throughout childhood as the drive of cognitive development. This idea dates back to Pascual-Leone's (1970) theory, in which he assumes a central processor called "M−space", M being the maximum number of schemes or chunks that the processor can attend, control, or integrate into a single action. The capacity of M is assumed to grow in an all-or-none manner as a function of age in typically developing children. The expression "all-or-none", used by the author, can be understood within a slot-based view of WM, according to which WM has a fixed number of slots to store information. In Pascual-Leone's theory, the discrete growth in M capacity is the precursor of the developmental stages postulated by Jean Piaget in his famous theory of children's cognitive development.

Although not restricted to explaining the development of WM, Pascual-Leone's work was echoed by many WM researchers, who put the assumption of a discrete increase in the mental storage space under test. Evidence favoring this view comes from studies that controlled other factors influencing recall performance, such as the ability to chunk information, and still found an age-related increase in spans. For instance, Burtis (1982) presented children aged 10, 12, and 14 years with pairs of letters for immediate recall and manipulated the facility of chunking the pairs. Some pairs were formed by double letters (e.g., DD, FF, HH), some by familiar pairs of letters (e.g., FM as in radio, BC as Before Christ, PS as at the end of a letter), and others were random combinations of letters. The doubled pairs were expected to be easily chunked by children of all ages, whereas the familiar pairs were expected to be chunked only by older children who can rely on long-term knowledge of the pairs. His rationale was as follows: if the increase in letter spans is driven exclusively by chunking ability and the storage capacity is constant throughout development, then age differences in performance in recalling double and random pairs should not be observed. Burtis (1982) found an age effect that was constant across conditions, even when controlling for the amount of chunking produced by younger and older participants. The author suggested that structural growth in short-term storage capacity in attention can explain a substantial part of WM development.

In a similar vein to Burtis' (1982) approach, a more recent study by Mathy and Friedman (2020) manipulated the ease of forming chunks by presenting classes of visual stimuli with varying degrees of familiarity (letters, colors, everyday objects, kanjis, shaded cubes, and irregular polygons; the assumption is that forming chunks is easier with more familiar stimuli). The authors used a serial recall task and calculated the spans in each class of stimuli. Critically, they calculated the ratio between the span of simple, familiar objects (i.e., letters, colors, and objects) and the span of complex, unfamiliar objects (i.e., kanjis, shaded cubes, and irregular polygons). The rationale was that if WM spans develop as a consequence of the ease of forming chunks, then the ratio between spans for simple (easy to chunk) and complex (hard to chunk) items should diminish with age, suggesting that adults' more effective use of chunking curtails the difficulty in memorizing complex information. The results contradicted this hypothesis. Spans for simpler stimuli were better than memory for complex stimuli across all age groups, and the ratio between them remained unchanged across development. The authors suggest that, although chunking ability probably improves with age, it does not account for increases in the span. It is worth noting, however, that other researchers used the same type of paradigm manipulating the ease of forming chunks and argued otherwise. For instance, Dempster (1978) presented children (7–12 years) with consonants, words, and nonsense syllables, the former two conditions being constructed to produce negligible age differences in chunking ability and the latter to promote them. He found greater age differences in the latter condition, in which chunking was much more difficult for younger participants; they suggested that the increase in the spans was caused by an improvement in forming chunks by older children. The conflicting findings in the examples illustrate our point of the difficulty in differentiating the influence of structural and functional constraints of WM in development.

The ability to combine pieces of information and form chunks in WM is not the only alternative to the hypothesis of an increase in storage capacity of the focus of attention. Some authors have suggested that chunks become larger via passive associative learning when children are exposed to natural language (Jones, 2012; Jones & Macken, 2015; Jones et al., 2020). According to this view, the number of slots in WM should not change with age, but the chunks occupying the slots should become larger.

The hypotheses of growth in the chunk size versus the number of chunks during childhood were specifically tested by Gilchrist et al. (2009). Children (7- and 12-year-olds) and adults were asked to recall lists of simple sentences with a varying number of words and clauses. The number of sentences in a list and the number of clauses in a sentence were manipulated so that the total number of chunks (i.e., the full sentence) and the size of the chunks (i.e., the number of clauses in a sentence) varied across conditions. For instance, a list of four short sentences containing one clause each represents four small chunks in WM. On the other hand, a list of four long sentences containing two meaningfully conjoined clauses represents four large chunks in WM. In all conditions, the sentences were unrelated to one another, so a story across them would not be easy to formulate. To assess the number of chunks kept in WM and the size of the chunks, the authors calculated (1) the number of sentences at least partly reported by participants, and (2) the proportion of words within a sentence that were recalled, provided that at least one content word from it was recalled. Consider a sentence as a chunk; if the chunk size increases with age, then the proportion of words correctly recalled from the sentences reported by the participants should increase with age, indicating that the chunks became larger. If instead, the chunk size remains constant and what increases is the number of chunks in WM, then the proportion of words from the reported sentences should be invariant across age groups and the number of sentences should increase. The latter was the result observed by the authors: younger children recalled words from fewer sentences (i.e., they accessed the content of fewer chunks) than older children and adults, but the proportion of words recalled from the accessed (reported) sentences was not different across the groups (about 80% in children and adults). The result suggests that the amount of information contained in the chunks (i.e., the proportion of words recalled from a sentence) is invariant during childhood, reflecting stability in chunk size, and that WM growth happens via an increase in the number of chunks (i.e., the number of sentences accessed).

In accordance with the account of an increasing number of chunks during childhood, Gilchrist et al. (2008) also found the opposite pattern when comparing younger to older adults. In a study using the same method as Gilchrist et al. (2009), the elderly recalled words from fewer clauses than adults, but both groups recalled the same proportion of words from the accessed sentences. Together, the results from the 2009 and the 2008 studies suggest that the increase in WM capability from childhood to adulthood and the decline in the elderly are greatly reliant on the number of chunks stored in WM. The fact that the number of chunks declines in old age further argues against the view that a growing knowledge basis accounts for the developmental effect in children, inasmuch as linguistic knowledge presumably accumulates throughout life.

The view that knowledge drives the development of a capacity-constant WM system has been present in the field for many years (Bjorklund, 1985; Chi, 1978; Schneider et al., 1989), and has been referred to as the "hardware invariant hypothesis" (Chechile & Richman, 1982). In order to falsify the so-called "hardware-invariance" of the system, one needs to experimentally equate the influence of knowledge across age groups and still observe age-related differences in capability measures. That is, assuming a constant storage capacity, the invariant hardware hypothesis predicts that the performance gap between adults and children should disappear after knowledge is controlled for. If these differences do not disappear, then we are led to conclude that storage capacity increases with age. We saw earlier, however, that a strong developmental growth in WM performance was observed across low and high levels of knowledge (unfamiliar characters vs. English letters: Cowan et al., 2015).

Further evidence favoring the storage account comes from studies that investigated filtering efficiency in childhood. Cowan et al. (2010) used a dual-task paradigm to test whether age differences in WM capability are explained by a more efficient attentional allocation in adults. According to this hypothesis, children's WM gets cluttered with distractors because they cannot filter-out task-irrelevant information at encoding. The authors tested this hypothesis by using a change detection task and varying the number of objects presented in a visual array. Participants (children aged 7–8 years and 12–13 years, and adults) saw a mixture of colored circles and colored triangles and were to remember the objects. In a given trial block, the distribution of trials testing recognition of an object of one shape versus the other were distributed 100%-0%, 80%-20%, or 50% each. The rational strategy is to attend more to the shape that is presented most often. If filtering (or prioritization) efficiency is the fundamental factor driving WM development, then adults' performance should be vastly more hindered than children's when the non-relevant shape is probed, as a result of their more efficient allocation of attention. It was found that filtering efficiency, as measured by the difference between performance on higher- versus lower-priority trials, depended on the memory load. With 2 circles and 2 triangles in the array, prioritization was the same across ages despite lower performance levels in the youngest children. Cowan et al. (2011) replicated this effect using a slower, sequential presentation of items rather than a spatial array. When there were 3 circles and 3 triangles in an array, however, Cowan et al. (2010) found that the youngest children no longer showed evidence of allocating attention according to the presentation frequencies, whereas older participants still did (Fig. 1). The results suggest that the constraining factor for selective attention at encoding was WM storage capacity – not the opposite – and thus attention allocation cannot be an explanation in itself for the WM capability increase in childhood.

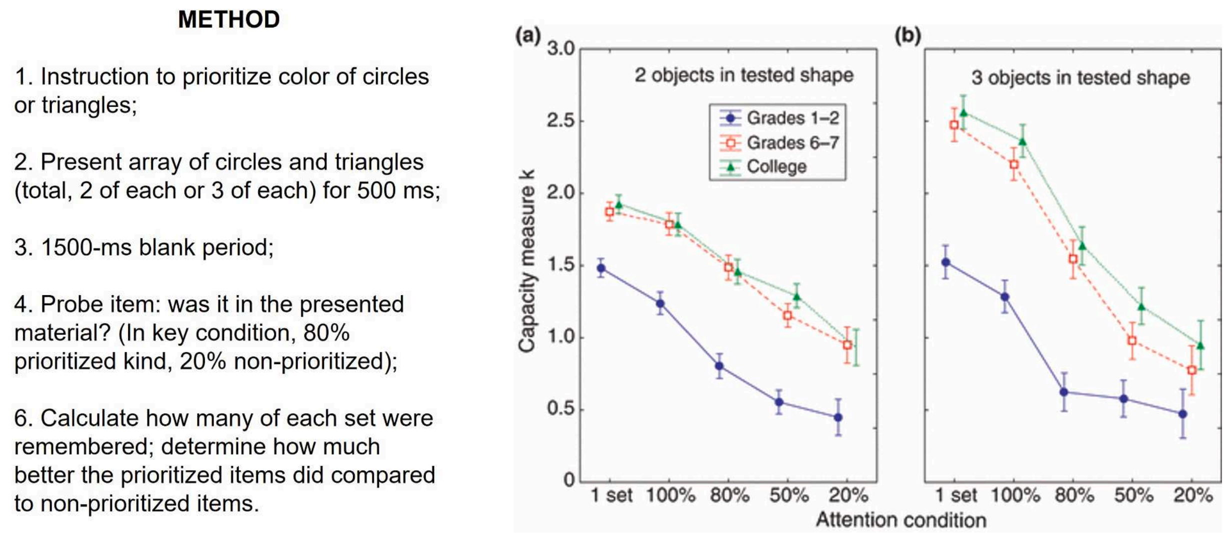

Fig. 1. *WM capacity according to the level of attention allocated to the probe, per age group.* The right half of the figure is reproduced from Cowan et al. (2010, Fig. 4). The *y*-axis represents the capacity estimates of WM capacity (*k* value, corresponding to the number of items stored in WM). The attention conditions in the *x*-axis correspond to a continuum in which increasingly less attention at encoding had been paid to the kind of item being tested on the present trial. Error bars are standard errors.

Note that storage capacity here is assumed by the authors to grow with development because of increases in the scope of the focus of attention, i.e., how many items can be simultaneously in the focus. For adult individual differences, there is a similar finding, strengthened by eye movement data to examine the locus of attention. Specifically, Mall et al. (2014) showed that filtering efficiency was not greater in high- than in low-WM-capability individuals, as evidenced by how long they spent looking at distractors. We will discuss attention prioritization in more depth while presenting our proposed framework.

Evidence for the storage-growth account, in sum, includes studies controlling for selective attention (Cowan et al., 2010, 2011), knowledge (Gilchrist et al., 2009; Cowan et al., 2015), and processing speed (Cowan et al., 2006), and consider that there is either an increase in the number slots in WM or in the size of the stored chunks. Table 3 summarizes the evidence in favor of these views. However, as we have mentioned, opposing evidence also exists. For example, when processing differences were neutralized across age groups, capability differences have been abolished in some instances (Case et al., 1982; Cowan et al., 1998; Gaillard et al., 2011). In the following section, we consider a third possibility before suggestion a resolution of apparently conflicting findings.

Is there an overarching, causal third factor?

Above we raised the possibility that neither storage nor processes (attention control, application of knowledge, etc.) are causal in the developmental increase in WM capability, but rather that there could be a third factor that is responsible for both storage and processing increases. A potential underlying factor is attention, given that both storage and processing are said to depend on the frontal-parietal network of the brain (for a review in relation to WM see Cowan, 2019; Rose, 2020). If storage and processing depend on the same brain mechanisms, perhaps maturation in the function of that system accounts for both storage increases and processing improvements with child development.

Arguing against the theory that storage and processing are based on the same mechanisms, there are several types of evidence suggesting that the behavioral manifestations of storage and processing are partly interrelated and partly separate, and that storage depend on interrelated but separate brain mechanisms. In the behavioral domain, Friedman et al. (2006) examined three kinds of executive function (inhibition, attention-switching, and updating WM) and found a relation to intelligence only of the one executive function that involves storage (i.e., updating). Cowan, Fristoe, et al. (2006) had a different type of attention control measure than Friedman et al. did and found a different result, but again supporting a distinction between storage and processing. They tested 10- to

Table 3
Summary of findings suggesting that growth in storage is the causal factor of WM development.

Type of storage growth viewed as a causal factor	Rationale for the explanation	Summary of evidence observed
Increase in the number of chunks stored in WM.	With development, the number of slots in WM increases, allowing children to store more chunks of information in WM.	When the easiness of forming chunks is manipulated in order to equate children's and adults' ability to chunk information, the age effect in WM capacity remains constant (Burtis, 1982). This suggests that there is an age-related increase in storage that is not explained by chunking ability *per se*. In a task manipulating the length of English sentences, children recalled fewer sentences than adults, but the proportion of words correctly recalled from the accessed sentences was invariant between the age groups (Gilchrist et al., 2008). This suggests that there was an increase in the total number of chunks stored in WM (i.e., the full sentences), not in the size of the chunks (i.e., the proportion of words retrieved from the sentences). The ratio between the span for item that are easy to chunk and difficult to chunk does not change with age, suggesting that what drives differences in the span is not the growing ability of forming chunks (Mathy & Friedman, 2020). The ability to filter out irrelevant information at encoding in children is as efficient as in adults, inasmuch as children's WM is not overloaded. Beyond a capacity limit of 2 items, filtering efficiency broke down in first graders (Cowan et al., 2010, 2011). This suggests that there is an increase in the number of items that can be encoded into WM during childhood.
Increase in the chunk size.	The number of slots in WM remains constant across development, but the size of the chunks stored therein is enlarged to accommodate more pieces of information per chunk.	When items are harder to be chunked, for example, while memorizing nonsense syllables compared to consonants and words, age differences between younger and older children are exacerbated (Dempster, 1978). Computational models of chunking via associative learning of phonemic sequences account well for developmental changes in digit span, word and non-word repetition tasks (Jones, 2012; Jones & Macken, 2015; Jones et al., 2020).

11-year-old children and young adults on storage and processing tasks. The storage tasks included a spoken digit span test, a printed letter span test, and a visual array test with colored spots to be remembered in a change-detection task. For attention control, the visual letter task with a 1 item per second presentation rate was combined with the spoken letter task, with the spoken letters starting half-way through the visual list at a rate of 2 items per second. The instructions on a given trial were to remember either the visual or the spoken list but, in conflict with the instructions, half of the tests were on the supposedly unattended list. Children in this situation showed no difference between attended and unattended list memory, but many adults did. Moreover, among adults, there was a correlation of 0.47 between the benefit for the attended items over the unattended items and a measure of intelligence. In regression analyses, storage and span measures picked up considerable variance in intelligence in both children and adults, but the processing measure only picked up variance in intelligence in adults. Storage (based on the visual arrays) and processing (based on the attention benefit in the span tasks) together picked up 37% of the variance in adults' intelligence, and it was partitioned as follows: 15% uniquely from storage, 10% uniquely from processing, and 12% from both factors together. Thus, storage and processing were found to be related, but partly separate.

Results from a study by Bayliss et al. (2003) also provided evidence of the separation of a storage component and a processing component contributing to WM capability in complex span tasks. Complex span tasks require participants to alternate between maintaining items (storage function) and processing information in WM, therefore being a suitable method to test the assumption that these two WM functions rely on a common pool of resources. In their study, Bayliss et al. (2003) separately assessed storage and processing efficiency in children (8- and 9-year-olds) and adults and used factorial analysis to examine how much of performance in complex span tasks loads on domain-specific and domain-general storage/processing factors. They found support for two domain-specific storage factors (one for visual information and other for verbal information) and for one domain-general processing factor, both in children and adults. The same three factors have also been found by Gathercole, Pickering, Ambridge, and Wearing (2004) in children from age 4 to 15 years. Gray et al. (2017) collected a large number of WM measures on 9-year-old children and constructed a structural equation model to examine the factors of WM. In some of the measures, information had to be not only stored, but also processed in some way. These measures included a 1-back task in which the participant had to remain vigilant in comparing each item in the stream to a previous item, and a task in which numbers in a small array in memory had to be updated based on new input. The resulting best-fitting model had three factors: a central executive factor for the tasks that needed processing, a phonological memory factor, and a third factor that seemed to reflect the focus of attention. It loaded highly on not only visual memory tasks, but also a running spoken digit span task in which it was impossible to rehearse or group the information (cf. Bunting et al., 2008) because the endpoint of the list was unpredictable. Thus, storage and processing were correlated but separate in the model. In follow-up work, Gray et al. (2022, Fig. 2) showed that the central executive and focus of attention factors had different implications. Whereas the focus of attention factor loaded on intelligence (0.60) more highly than did the central executive (0.40), the latter also loaded on a measure of expressive vocabulary (0.22), whereas the focus of attention did not (0.18).

In brain function, Cowan (2019) described considerable evidence that an area in the intraparietal sulcus serves as a hub or focus of attention, functionally connected to both the frontal areas that control attention and the posterior areas that represent the stored information. Majerus et al. (2018) manipulated both the load to be remembered (e.g., the number of letters and digits in a series) and the difficulty of processing (instructions to remember all items in a series, or only one kind). Both a higher storage load and a greater processing difficulty activated the frontal-parietal network, but only the storage load also activated the posterior regions representing the information (temporal regions for spoken items, and occipital regions for faces).

Given the evidence that storage and processing are partly separate and partly interrelated, we prefer a model in which there is a cascade of events in which, at key points, either storage or processing is primary and has consequences for the other factor.

A storage-and-processing cascade framework and support for it

One definition of a cascade is "a process whereby something, typically information or knowledge, is successively passed on" (from a Google search of *cascade*, based on Oxford Languages; second definition). The description of a storage-and-processing cascade seems suitable for the hypothetical description of WM task performance given in the section "*WM Development in Childhood: A Matter of Storage, Processing, or Both?*" (the child who is asked to put the camera, the turtle, and the globe into the box). We have seen thus far that increases both in processing and storage seem to be the fundamental sources of WM development in childhood, depending on the different experimental situations. These two factors need not be mutually exclusive explanations for the age differences in WM capability. Each one might be fundamental in different steps of information processing in WM task and in different developmental ages. Here, we propose a tentative theoretical analysis to reconcile the two branches of evidence favoring either storage or processing, by integrating them in a processual description of a trial event in a WM test. The proposed framework is not intended to explain which factor is fundamental at different ages, for the aforementioned hypotheses have not been extensively tested in all age groups. Rather, it aims at elucidating how the age differences in both storage and processing affect each step of information processing in WM in the context of a dual-task, thus causing different capability estimates in children and adults.

Description of the framework

Take as an example the situation we mentioned at the beginning of this article. First, a set of instructions requiring memory for four elements is given to the participant ("When I say 'go', can you put the *camera*, the *turtle*, and the *globe* into the *box*?"). At this step, irrelevant events need to be suppressed and not encoded into WM, and the success of this operation depends on the selective allocation of attention to the targets (the camera, the turtle, the globe, and the box, but not a butterfly flying by or a teddy bear laying near the target objects). The input of information to an attention-demanding store depends on a central capacity limit in WM, as the required filtering information will demand some of this capacity to be carried out. Cowan et al. (2010, 2011) showed that the attentional allocation at encoding by the youngest children (7-year-olds) was as efficient as that of the adults, provided that the number of attended items did not exceed their WM capability. Beyond the capacity limit, the filtering efficiency broke down in children, suggesting that the constraining factor of WM capability at encoding was not the filtering ability in itself, but rather storage capacity of the focus of attention.

Following the successful encoding of the targets, items are briefly maintained in the focus of attention until a second process is triggered. Items in the focus of attention can be chunked to form meaningful patterns and expand the effective storage capacity. By creating larger chunks, one is able to allocate more information in the slots available in WM and reduce load (Thalmann et al., 2019). Research by Gilchrist et al. (2009) has shown that the size of chunks remains stable in childhood, at least for simple information that does not provide an advantage for the older participants, and what changes with development then is the number of chunks in WM. Therefore, we propose that the fundamental difference between children and adults at this step of information processing is whether or not chunks are formed and what is done with the formed chunks. It is possible that adults are better able to create and/or detect inter-item associations and, once the items in WM are associated, they are offloaded to activated long-term memory releasing space in the focus of attention.

The result of such an offloading operation is that adults need less attentional resources to maintain items in WM, and therefore can use the focus of attention to carry out other concurrent operations and/or store more information. This is illustrated in our working example by a child that is already able to chunk the items turtle, camera, and globe, imagining that the turtle is taking a picture of the globe. Such an association demands less storage space in the focus of attention compared to holding each of the three items separately. The chunked items can be temporarily "put on hold" in activated long-term memory until the moment they become relevant for carrying out operations, e.g., response selection.

We assume that greater knowledge and experience levels in older children and adults scaffold the creation of patterns necessary for chunking (Miller, 1956) and the subsequent offloading operations. This aspect of the framework needs to be further developed because we still do not know if the creation of patterns is the means by which information is offloaded from WM in adults. We know, however, that the contribution of activated long-term memory in accounting for capability substantially increases in the late elementary-school years, leaping from holding less than 1.5 item to more than 3 items in adulthood (Cowan et al., 2018). The current working hypothesis in our lab is that adults file information into activated long-term memory by offloading patterns from the central component of WM, that is, the focus of attention. At this purported step of our cascade framework, the fundamental factor constraining children's WM capability is the ability to apply knowledge to the contents of WM, because it determines the efficiency of the pattern creation and its subsequent offloading.

Once out of the focus of attention, the offloaded pattern needs to be revisited by attention in order to be maintained during distraction. Say a competing demand interrupts the task at this step, and the child is asked about where the remote control is. Younger

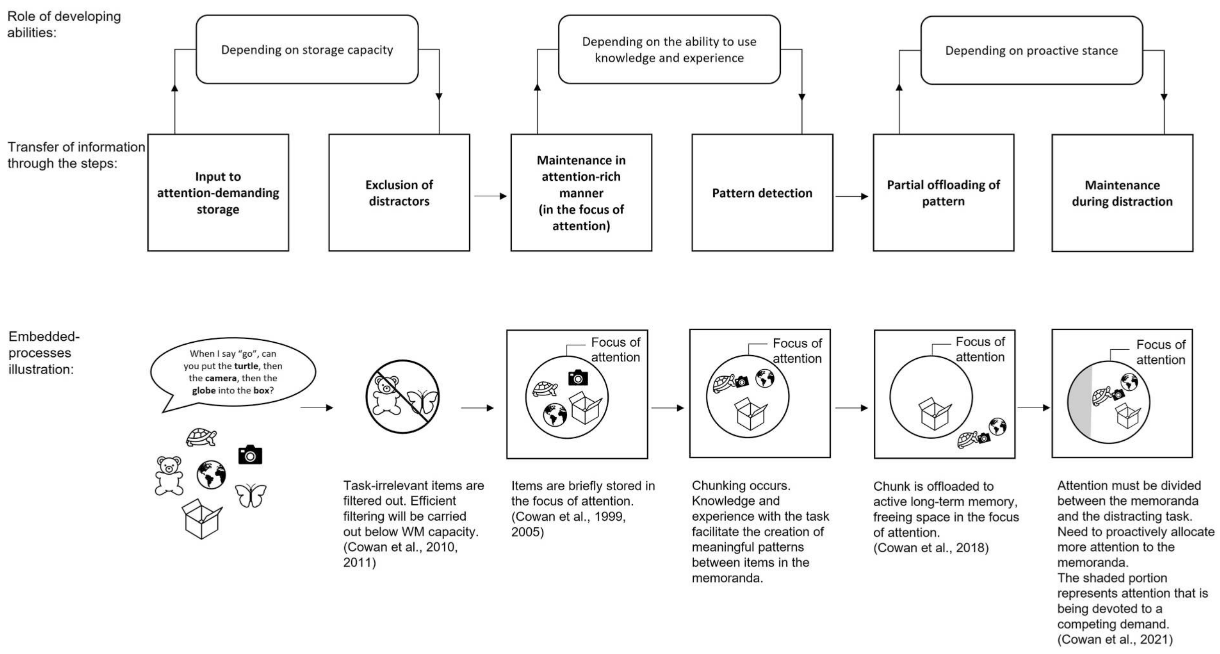

Fig. 2. *Influence of storage and processing in the management of the focus of attention in a dual-task setting.* The upper row represents how the developmental levels of storage and processing skills mediate the transfer of information through the steps in WM, according to the Storage-and-Processing Cascade Framework. The box diagrams in the middle row represents the steps of information processing in WM, during a memory trial. The bottom portion of the figure illustrates, for one task, the contents of WM in each step, and how items are managed by the focus of attention.

children are likely to completely forget the information that was being held in WM because their attention was completely occupied by the secondary task. At this point of information processing, the successful maintenance of the offloaded patterns depends on the ability to share attention between the two tasks and not drop the items previously stored in WM, i.e., the created pattern. In our imaginary situation, this corresponds to not completely forgetting the turtle taking the picture of the globe while attention is diverted to answering the question about the remote control. By adopting a proactive stance of attentional control, one should be able to revisit the imagined pattern that had been offloaded, prevent it from being forgotten during the interruption, and correctly execute the response of putting the objects in the box even after reacting to the interruption.

It has been observed that in dual-task settings, younger children (6–8 years) adopt a reactive stance of control and devote most of their attention to the execution of a secondary task in the detriment of maintaining information in WM. Adolescents (10–14 years) and adults (college students) show the reversed pattern and recall more information from the memoranda, at the cost of making more errors and slowing their reaction times in the secondary task (Cowan et al., 2021). These results suggest that, with age, children develop a proactive stance of attentional control that allows them to strategically share attention between the two tasks and prioritize the maintenance of the items. In our proposed framework, the proactive stance to dual tasks is the fundamental factor determining the success of maintaining information in WM during distraction.

Fig. 2 represents which factor – storage or processing – is fundamental in accounting for age-differences in WM capability at each step of information processing in a dual-task setting, according to the Storage-and-Processing Cascade Framework.

Support for the framework

Encoding into the focus of attention. The first step in our framework is the input of information to an attention-demanding storage system. In this step, information is assumed to be consciously processed and enter the focus of attention before attention is selectively allocated to relevant stimuli. This means that in a simultaneous presentation setting, as the one in our working example, the encoding of information from the surroundings is a pre-requisite for the exclusion of distractors to happen. In the example, the child encodes all information (the turtle, the globe, the camera, the box, the teddy bear, the butterfly, and the verbal command given by the adult) and then selectively allocates attention to the task-relevant items (the turtle, the camera, the globe, and the box). Under this assumption, limitations in the scope of attention would constrain this initial step by limiting the amount (or completeness) of information that can be encoded at once in WM, even before filtering can occur. Incomplete encoding, therefore, could theoretically explain lower capability estimates in younger children, without necessarily requiring the efficiency of filtering mechanisms to explain capability limitations.

Cowan et al. (2011) designed a study to test the hypothesis of incomplete encoding in children. They replicated the method used in a previous study (Cowan et al., 2010, discussed previously), in which participants had to allocate attention differentially to the colors of circles versus triangles, by changing the presentation of a visual array from simultaneous to sequential with a slower presentation rate (1/second rate). This was done to provide children with more time to encode the items. The sequential presentation was intended to spare children from the theoretical disadvantage of not being able to encode the whole visual array during the brief simultaneous presentation that Cowan et al. (2010) used. Cowan et al. (2010) showed that with 2 circles and 2 triangles in each array, 7-year-olds prioritized the shape tested on 80% of trials as well as adults did, but that their performance was at a lower level than adults and older children. Cowan et al. (2011) used this same manageable set size to determine whether encoding differences could explain the age difference in capability. The age differences between 7-year-olds and adults remained with a slow, 1/second presentation rate, and this was the case in each of several different encoding conditions (remain silent, name the color of each object to encourage rehearsal, or say "wait" after each object to discourage rehearsal). More time for encoding did not reduce the difference between 7-year-olds and older groups. However, when verbalization of task-relevant information was encouraged or allowed, children 11–13 years old no longer performed as well as adults; there was an additional factor that came into play, presumably based on verbal rehearsal or elaboration.

Exclusion of distractors. The second step in our framework is the exclusion of distractors from WM by means of selective attention. In the working example, this operation corresponds to the ability to filter out task-irrelevant items, carried out by the focus of attention. It is theoretically possible that children's capability estimates are lower because of an inefficient filtering mechanism that causes WM to be cluttered with task-irrelevant mental representations. According to this view, capability estimates would be lower in children not necessarily because storage space is smaller, but because it is occupied with distractors. Although we have reviewed some evidence of conditions in which selectivity develops with age (Elliott, 2002; Lane & Pearson, 1982; Smith et al., 1975) and children do show effects of admitting too much into WM (e.g., Plebanek & Sloutsky, 2017, 2019), Cowan et al. (2010) showed that capability differences persist even when the conditions are such that the prioritization of information admitted into WM works to a degree similar to adults. This was the case provided that WM was not overloaded (Fig. 1, left data panel). The breakdown of prioritization when WM is overloaded (Fig. 1, right data panel) suggests that the process of being selective and prioritizing WM optimally shares an attentional resource with storage.

In the study of Cowan et al. (2010) shown in Fig. 1, in order to conclude that it is indeed WM storage capacity that limits selectivity and prioritization, and not the speed of encoding, it would be helpful to test this full pattern with more encoding time. One could argue, for instance, that children are slower to encode stimuli presented simultaneously in the visual display, and therefore not able to efficiently prioritize the more-often-tested shape. Although Cowan et al. (2011) repeated the method of Cowan et al. (2010) using more encoding time (i.e., sequential presentation of items), they did not include the larger set sizes of 3 circles and 3 triangles, at which Cowan et al. (2010) found that 7-year-olds were no longer able to prioritize the more-often-tested shape. Thus, it is possible that a slower rate of encoding was the limiting factor, not capacity. However, previous investigations have suggested that, in adults, extending the time for encoding before a mask occurs affects the number of items encoded only until about 200 ms (Vogel et al., 2006). In contrast, Cowan et al. (2010) presented arrays for 500 ms and followed each one not with a mask, but with a 1500-ms blank interval before the comparison. Under these conditions, we believe, children's difficulty to selectively encode items of the more-often-tested shape is most likely to occur not because the encoding time is insufficient, but because the need to store the items using the focus of attention conflicts with the need to use attention to select information.

In sum, prioritizing information that reaches WM to minimize the effects of distractors or low-priority items is a process that seems likely to be limited by storage capacity, and not only encoding time. Studies showing a concomitant increase in WM capability and selectivity often use experimental tasks that mimic complex situations. For instance, in a study by Elliott (2002), participants were supposed to recall serial information (a digit span) while ignoring irrelevant auditory information that varied in terms of its features (speech or musical tones) and state (steady or changing stimuli, e.g., "red-red-red" or "red-yellow-blue", respectively). In another study, by Maccoby and Hagen (1965), participants needed to keep track of location-color bindings (memorize the background color of picture cards and indicate the location of a given color in the array, after the cards had been flipped down), while performing a tone discrimination task. We believe that these complex testing situations might incur a greater WM load and therefore affect filtering in younger children.

The results reported by Elliott (2002) offer an example of how WM load might affect filtering. The author observed an increase in span measures during childhood and a decrease in interference caused by irrelevant sounds. Compared to a silent control condition, the youngest group (second graders) was the most disrupted by the irrelevant sounds conditions. Interestingly, the proportion of items correctly recalled from the list was invariant across age groups in the silent condition. It is worth noting that the length of the lists was tailored to each participant's capability and the lists' length never exceeded their span. According to our framework, it is possible that the sounds were more disruptive to young children's performance because they overloaded their WM capacity at encoding.

Maintenance in the focus of attention. Our framework was designed to reflect the position in the embedded-processes framework that the focus of attention has storage properties. (Information is also stored in activated long-term memory, but we assume that the initial storage in the focus of attention is crucial for good encoding.) At this step, items are briefly maintained in the focus of attention and hence, WM capability will be limited by its size.. The suggestion that WM has a core capacity reliant on attention and that increases in childhood was proposed by Cowan et al. (1999). In this study, children (first and fourth graders) and adults were presented with lists of spoken digits and were tested under different attentional conditions. In one condition (attended speech), participants were instructed to listen closely and recollect each list immediately after its presentation. In a second condition (unattended speech), they were instructed to ignore the spoken digits and play a rhyming game (click on pictures of objects that rhymed with a target object) and were asked to recall the last presented list only occasionally, so that the test was unpredictable. Performance in the attended speech condition was assumed to reflect memory capacity under the influence of attention-demanding mnemonic strategies and verbal rehearsal. Performance in the unattended speech condition was assumed to reflect the transfer of information, when a recall cue is presented, from a short-term auditory sensory memory, which fades within a few seconds but otherwise has no capacity limit itself, into a WM with a fixed limit (Darwin et al., 1972). This measure is called "span of apprehension", because it corresponds to the maximum of sensory information that can be apprehended (or, in our terms, transferred into the focus of attention) without the implementation of mnemonic strategies. For all age groups, performance increased across list lengths in the attended speech condition, but it remained stable in the unattended condition (i.e., the number of correctly recalled digits was the same irrespective of list length). First graders correctly recalled 2.4 items in the unattended condition, fourth graders recalled 3.1 items, and adults recalled 3.5 items. The stability across list lengths in the unattended condition reflects a central limitation in the span of apprehension, and the age differences reflect a developmental increase in WM core capacity. The observation that this core capacity increases as a function of age during childhood is aligned with theories proposing an increase in storage capacity of the focus of attention during childhood. Because participants also recalled items in early serial positions, the authors suggest that the limit observed is more likely to reflect an attention-related faculty rather than the sensory record itself. In the latter case, correct recall should be restricted to later serial positions in the unattended condition. According to the authors, the fixed limit in WM can be understood as the capacity of the focus of attention when revisiting the sensory record of the list at recall.

Cowan et al. (2005) tested the hypothesis of an increase in the scope of the focus of attention during childhood by using two approaches. The first was via change detection tasks. They prevented verbal encoding and chunking strategies by using very short presentation times at encoding, so performance was assumed to reflect a pure measure of the capacity of the focus of attention. Then, it was possible to estimate the size of the focus of attention by varying the size of the stimulus set. In this type of paradigm, the maximum number of items in the focus of attention is considered to be reached when the number of items estimated to be in storage reaches an asymptotic level. The second approach to estimate the growth in the focus of attention was via the running span task. In this type of task, participants are aurally presented with rapid lists of digits, and they are required to recall as many items as they can from the end of the list. The lists vary in length and are interrupted unpredictably, so that participants cannot anticipate when they will be asked to recall the digits. Because of the rapid presentation and the unpredictability of the task, recall is assumed to reflect the capacity of the content of the focus of attention at the moment of test. Indeed, a distraction before recall disrupts the process of drawing information from a passive store into the focus of attention in a running span task (Bunting et al., 2008). Via these tasks, Cowan et al. suggested that the size of the focus of attention is about 2 items around the age of 6–7 years and increases to about 4 items in adulthood. However, later we will qualify this conclusion on the basis of other results (primarily Cowan et al., 2018).

A limited focus of attention with storage functions can explain fast memory loss in young children at this step of information processing. For instance, younger children are known for neglecting goals while performing WM tasks, a phenomenon that correlates with WM capability and that decreases with age (Marcovitch et al., 2010). Most studies on goal neglect use the card sorting paradigm, in which children are asked to sort cards according to a set of pre-determined rules that need to be flexibly adjusted as the task advances (e.g., according to the number of objects, to their colors, etc.). In order to succeed, children need to keep the rules activated in WM while processing the cards. In such a situation, and without the implementation of a maintenance strategy by young children, a limited focus of attention would result in the loss of the set of rules as incoming stimuli are encoded.

Indeed, a study by Bertrand and Camos (2015) showed that preschoolers (ages 4 to 6 years) do not attempt to implement any memory strategy to maintain information in WM in a serial recall task (to memorize a sequence of fruits put in a shopping bag while playing "grocery shop"). Introducing an unfilled delay and doing a high attentional demand task (walking in a sinuous path to the grocery shop) before recall depleted performance to the same extent. This led the authors to suggest that preschoolers do not attempt to implement maintenance strategies to maintain the task goal even when they have no competing demands. Interestingly, when children were instructed to walk directly to the "grocery shop" (low attentional demand), recall performance slightly improved in comparison to the unfilled delay and the high attention condition. The authors suggest that a motor activity that is related to the task goal and does not load too much on attention can actually improve performance by scaffolding goal maintenance. According to our interpretation, the phenomenon of goal neglect in preschoolers can be partly accounted for by the fact that, as they are performing a given task, their WM is filled with mental representations of the incoming stimuli that supplant the representation of the rules. In the case of Bertrand and Camos' (2015) study, walking directly towards the "grocery shop" might have helped to keep the mental representation of the task goal in the focus of attention. It is worth noting, however, that follow-up studies attempting to scaffold cognitive control by providing cues of the task goal have provided mixed or no evidence of WM improvements in younger children (Fitamen et al., 2019a, 2019b, 2022). For example, in Fitamen et al.'s (2022) study, children's performance did not improve in a condition in which a toy shopping stall (a reminder of the task goal) was added to the experimental setting.

To summarize, we propose that after the encoding and selection of relevant items, information is maintained in the focus of attention and is subject to its capacity limits.

Pattern detection or creation. The fourth step in the model is the formation of patterns between items in the focus of attention. In this step, the mental representations form conglomerates of meaningful associations that unburden attention. Note that we speak of a detection and/or creation of patterns because we do not want to commit to any assumption of automaticity versus control at this stage of our theoretical elaboration. Although our working hypothesis is that ability to form patterns between items in WM is improved with age thanks to better application of knowledge to the memoranda, we admit the possibility that this could be an automatic process that improves with age. One could either passively form semantic associations between items, as in "milk-glass", "chair-table", etc.; or one could deliberately create a conglomerate of items that are, at first sight, unprovided of any previous semantic association but then become meaningful by means of such operation, e.g., by mentally combining geometrical shapes to create a pattern that resembles a house. We also speak of "patterns", broadly and loosely defined, because the term can accommodate different processes known for linking items to each other, like chunking (Miller, 1956), clustering (Farrell, 2012), mental imagery (Krinsky & Krinsky, 1996; Sahadevan et al., 2021), rhyming (Bower & Bolton, 1969), etc. For example, chunking digits in a sequence, as in Miller (1956), and using visual imagery to link concrete, imageable words (e.g., Einstein et al., 1989) are undisputedly different mental processes, but both share the commonality of linking mental representations into a more compact structure. All the examples given would result in an inter-item pattern that is less attention-demanding than maintaining its constitutive parts individually.

A lot of research has been done in the 1970′s regarding the use of elaborative mnemonics (e.g., visual imagery, sentence generation) to learn paired associates by children. These are some ways that memorable patterns can be formed and remembered. Globally, studies showed that the use of elaborative strategies increases until adolescence, that it promotes better learning of the material compared to the use of verbal labeling, and that children who can use it to learn the pairs outperform children who cannot (Kerst & Levin, 1973; Kestner & Brokowski, 1979; Turnure et al., 1976). More recently, Sahadevan et al. (2021) have shown that creating inter-item associations is more effective than using the peg-word method in a serial recall task. In their study, participants (young adults) were trained in the use of peg-words,[2] interactive imagery, and the link method. In the interactive imagery condition, the words in a list were presented in pairs, and participants were instructed to create a mental image of the two items interacting with each other (e.g., flame-cabin, drum-tail, brandy-rubber, a cabin in flames, a tailed drum, a rubber-made canteen of brandy). In the link method condition, the words were not paired and participants were told to create images of adjacent items interacting with each other (e.g., flame-cabin-drum-tail-brandy-rubber, a cabin in flames plays a tailed drum, then someone trips in a rubber tire and drops some brandy). The most effective strategy was interactive imagery, followed by the link method. The results suggest that forming conglomerates of items in the memoranda is even more effective than linking items to long-term representations that are extraneous to the list to be memorized (i.e., the peg-words). In another study (AuBuchon & Wagner, 2023), the use of self-generated, elaborative strategies (mental imagery and/or sentence generation) was a leading predictor of word serial recall when used in conjunction with phonological rehearsal and grouping, especially for phonologically similar lists. Together, this ensemble of results suggests that elaboration is a powerful way to organize the material in WM and its use can be a source of age-related increase in WM capability.

In our framework, the creation/detection of patterns is thought to depend on the ability to apply knowledge to the memoranda and experience with the task. Studies by Jones and colleagues (Jones & Macken, 2015; Jones et al., 2020) on chunking and associative memory provide support for the idea that knowledge scaffolds WM maintenance by means of experience on applying it to the memoranda. By analyzing the finding that the serial recall of digits improves in childhood at a much faster rate than the serial recall of

[2] A mnemonic strategy that consists in linking items in serial positions to words that rhyme with cardinal numbers, e.g., one is "bun", two is "shoe", three is "tree", etc. By using the method, the list clock-pencil-bottle could be remembered by imagining a clock sandwich (a clock in a bun), a pencil wearing shoes, and a tree drinking a bottle of water. The peg-word mnemonic has been shown to be very effective in serial recall (Roediger, 1980).

words, the authors suggested that exposure to natural language offers support to chunking operations in WM. Their argument is that, because random sequences of digits are much more frequent in natural language (e.g., zip codes and telephone numbers) than random sequences of words, children become better at associating random digits than associating words. For instance, memory for digit pairs improved more than for word pairs between ages 5 and 10 years (Jones et al., 2020). The authors firmly defend the idea that associative learning and language exposure can account for developmental trends in the digit span without resorting to a storage system that grows with age (Jones et al., 2020). Ultimately, they suggest that span measures for verbal lists increase with age because language exposure increases with age. We would, however, dare to use their argument of greater exposure to digit sequences in the natural language in favor of our idea that the application of knowledge and experience with the task materials improve with age, explaining some of the WM development. According to our interpretation, the faster growth in digit spans could be explained by the fact that children get more proficient in using knowledge to combine and chunk digits, precisely because they are more exposed to this type of association between stimuli in natural language. We also highlight the results by Cowan et al. (2015), already mentioned, showing that knowledge growth cannot account for all of the developmental increases in WM. In this study, the increase in WM capacity for English letters and unfamiliar characters across age groups followed the same pattern (if knowledge was the main driver of WM growth, then the increase in capacity should be much steeper for English letters).

In sum, we propose that a child's ability to use knowledge and their level of experience with a certain task and/or stimuli set will facilitate the formation of meaningful conglomerates between items in the memoranda. In the following step, such patterns will be transferred to activated long-term memory and offload attention from storage demands in favor of our view that increased knowledge and experience with the task's material can unburden WM and partially account for its development during childhood.

Offloading of patterns. The fifth step in our framework is the offloading of patterns from the focus of attention to activated long-term memory, a process that we assume to happen in adults and to improve with age. The idea of a process that offloads attention resulted from a study in our lab investigating the growth of the focus of attention during childhood (Cowan et al., 2018). In this study, Cowan et al. (2018) observed that the number of items stored in the focus of attention, which they call a *central* portion of WM, stays constant or even decreases with age, whereas there is a marked increase in storage within what they term visual and verbal or acoustic *peripheral* portions of WM (not modules, but the part of WM that cannot be flexibly allocated to either visual or acoustic items). Based on these results, we suggest that there is a transfer mechanism that displaces information from the focus of attention to activated long-term memory, freeing central storage space, but nevertheless relying on attention to keep track of the information that has been offloaded. (Presumably, offloading information to be retrieved later frees up some attention for other uses, but some attention to it is still needed to keep track of it.).

It is useful for this argument to explain the assumptions and calculations made by Cowan et al. (2018). Studies on the capacity of the focus of attention often take experimental precautions to fill it with information so as to cause an overload of capacity. We have mentioned several studies that adopted this approach in our article, and most of them traditionally vary the set sizes in order to find an upper capacity limit across different age ranges. In this approach, the capacity measure (k) is based on the difference between hit rates (proportion of correct recognitions) and false alarm rates (proportion of false alarms), weighted by the set size. The underlying assumption is that participants will correctly recognize the probe when its mental representation is stored in WM at the moment of the test. For instance, suppose a participant has perfect performance with a hit rate of 1 (i.e., 100% of correct recognition) and 0 false alarm, and she was presented with sets of three items. In this experimental situation, $k = 3$, meaning that the whole set was in WM when the items were probed, yielding perfect performance. The model used to estimate k is based on the assumption that the correct answer can sometimes be obtained by a lucky guess, which is essentially subtracted from the total to estimate items in WM.

Measures of k are very useful in research (they provide a capacity estimate in recognition or change-detection tests, in which measures of span cannot be applied), but they cannot tell us about the specific contributions of storage and processing in WM, nor about the size of the focus of attention at a given age. This is because k does not necessarily reflect only operations that exclusively took place in the focus of attention. It is possible, for instance, that the scope of attention remains unchanged and that peripheral, non-attentional portions of WM gradually start to store information with development. In that case, one could observe an age-related increase in k that does not necessarily correspond to an increase in the central storage in WM. It is necessary to distinguish central and peripheral storage in WM in order to test the hypothesis of a growing amount of information stored in the focus of attention (central storage) during childhood. (Note, though, that as we will explain later, attention can contribute to peripheral storage too, as it may be used to prepare information for a more efficient offloading out of the focus of attention into activated long-term memory.).

Cowan et al. (2014) developed a bimodal paradigm to distinguish between these two sources of information in WM, and Cowan et al. (2018) later adopted it to examine their development during childhood. In the task, participants are presented with two sets of stimuli, each of them in a different modality (e.g., a series of musical tones and an array of colored shapes, presented one set after another), and asked to remember information of either one or both sets. Because the focus of attention is equally devoted to storing information from different domains (Cowan, 1999; Saults & Cowan, 2007), items from both sets compete for its storage space in the dual condition, but not in the single condition.

An estimate of the central part of WM storage is the difference obtained by taking k when only one set of items is to be remembered, summing it across both types of sets, and subtracting the smaller k obtained when both sets have to be remembered at once. In a typical finding with adults, single-set k estimates might be about 3.5 colors and 3.5 tones, and dual-set k estimates might be about 3.0 colors and 3.0 tones. Then the estimate of the central portion would be $(3.5 + 3.5)\text{-}(3.0\text{–}3.0) = 1.0$. Peripheral visual performance would be calculated as $3.5\text{–}1.0 = 2.5$ because it is assumed that the capacity estimate of 3.5 was obtained when the participant could use all of the central portion for vision in that single-set condition, so that it must be subtracted from the total to obtain the peripheral portion (i.e., single-set k-central = peripheral). The same would be true of the peripheral portion for audition. Note that such distinction between central and peripheral WM does not depend on the assumption of separate storage modules in the peripheral portion; it simply assumes that items' features are activated outside the conscious focus of attention.

Cowan et al. (2018) compared the central and peripheral capacity estimates of children in the age ranges of 6–8, 8–10, and 10–13 years old and adults. The results contradicted the prediction of an age increase of the central portion of WM and showed that the estimated size of the focus of attention only increased a bit until age 8 (from 0.5 to 1.0 item between the first and second age groups), whereas the capacity of the peripheral components increased markedly with age. This finding suggests that, with age, children do not increase in the use of attention for storage, but instead figure out how to store more information without increasing the investment of attention for storage per se. This result is shown in Fig. 3.

Although we do not yet know why the number of items stored in the peripheral components increased with age in Cowan et al. (2018), our current working hypothesis is that older children get better at filing away information into activated long-term memory in an easily retrievable form, offloading it from the focus of attention to free it up for other tasks. Although the offloading process lightens the burden of holding items in the focus of attention, we propose that it is not an attention-free mechanism. We assume that at least some central resource is used to keep track of information put in activated long-term memory. An underlying mechanism assisting this offloading process is thought to be the application of knowledge to create patterns between items in the memoranda, as explained in the previous section.

Alternative to the idea of offloading, it is still theoretically possible that auditory and visual information both improve across ages in terms of the basic dedicated modality storage space for these items; presumably temporal lobe areas for tones or verbal materials and occipital lobe areas for visual items (e.g., Cowan, 2019; Majerus et al., 2018). A difficulty with this alternative interpretation, though, is that there is no evidence that these areas for storage of information undergo a dramatic change in the childhood years, and we can think of no behavioral evidence supporting it. Perhaps the closest is the finding that WM development includes improvements not only in the number of items remembered, but also in the precision of the representations, which has been found using items on a continuum of visual orientation (Sarigiannidis et al., 2016) or tone frequency (Clark et al., 2018). In the present situation (Cowan et al., 2018) of categorical items for which memories do not have to be very precise for a correct response, the precision of the representation may not be relevant. Nevertheless, the question might be addressed by combining a dual-set situation like Cowan et al. (2018) with a task yielding estimates of precision, to determine if precision is important especially for the peripheral component of WM as would be necessary to account for results like Fig. 3.

A study by Akyürek et al. (2017) suggests that the idea of attentional offloading via the creation of patterns is a promising hypothesis. The authors measured electrophysiological activity during a rapid serial visual presentation paradigm in which a varying number of visual targets (corner lines) were presented sequentially and rapidly (100 ms/stimulus) among distractors (printed letters). Critically, the targets could either be or not be spatially integrated into a visual pattern (e.g., upper left corner and bottom right corner, forming a square). At the end of each trial, participants were asked to recall the orientation of the targets seen in the trial (upper left, upper right, bottom left, bottom right). The authors were interested in the magnitudes of the following event-related-potential (ERP) components: the N2pc, a component associated with the attentional processing of visuospatial information; the SPCN (sustained posterior contralateral negativity), a component associated with memory load during retention; and the P3, a component associated with memory consolidation and response selection. Lower amplitudes in these components are assumed to reflect less cognitive effort, either because there was less information to encode and maintain in WM or because these processes were more efficient in one situation than another.

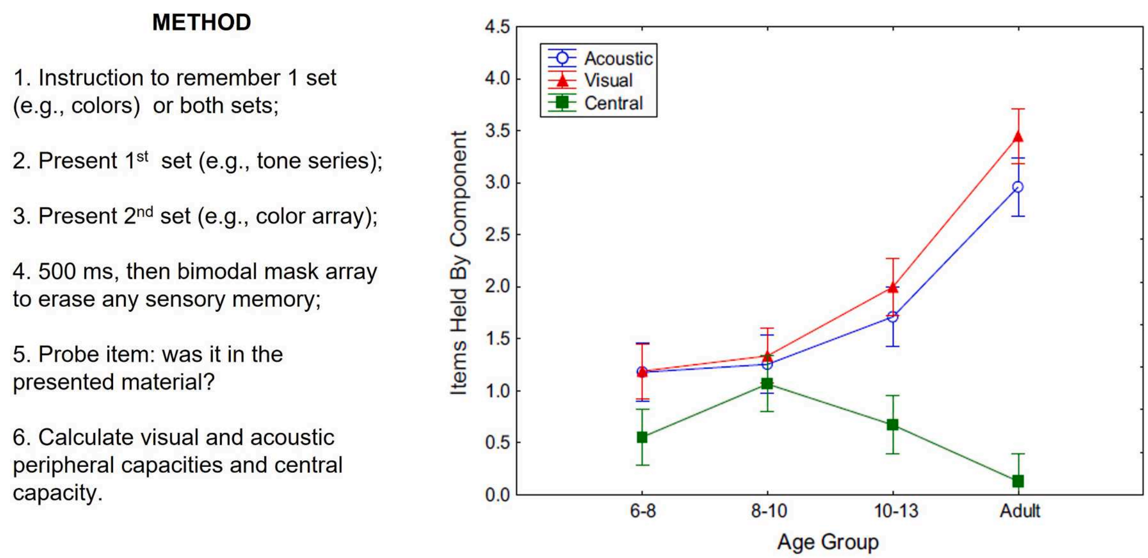

Fig. 3. *Illustration of developmental findings separating central and peripheral portions of WM.* Methods and results from Cowan et al. (2018). The right half of the figure is copied from Cowan et al., Fig. 4. Error bars are standard errors. Older children increase capacity not in the central portion of WM that can be shared between tones and colors, but in the peripheral portion that is not shared.

The results showed that the presentation of two successive integrable targets yielded the same amplitudes of the N2pc, SPCN, and P3 components as in the single target condition. Moreover, the amplitudes provoked by two integrated targets were significantly smaller than that of two non-integrable targets. Akyürek et al. (2017) suggest that the integrability of the two successive targets unburdens attention and WM and recruits the same amount of resource as processing one target. Their results cannot provide information about the possible involvement of activated long-term memory in the offloading process, but they serve as illustration of our proposal that attention is discharged when patterns are created in WM. Further arguments for this hypothesis were summarized by Rhodes and Cowan (2018).

Maintenance during delay. The series of events taking place in the child's WM culminate in the need to maintain information during a delay in which other concurrent demands might get in the way. Maintenance at this step might be accomplished by rehearsal, re-imagination of the patterns, or refreshing. Critically, it is important that attention is not entirely consumed by other processing demands and that the focus of attention is unoccupied by the memoranda. In order to do so, a child must have already acquired enough executive control to proactively share attention strategically between the memoranda and processing any distractors during the delay. The sixth step of our framework, therefore, addresses the issue of the proactive control of attention in dual-task settings during childhood.

Cowan et al. (2021) tested the hypothesis that younger children's lower WM capacities are caused by an undeveloped proactive control of attention. According to this hypothesis, younger children are not yet able to proactively carry out the attentional sharing operation and do not prepare themselves for an upcoming memory test during a delay filled with a distracting task. This causes children to neglect the maintenance of the memoranda (or perhaps the goal of preserving those memoranda) when they are required to share attention between tasks. In contrast, adults proactively anticipate the memory test and share attention between the two tasks at stake, causing them to have smaller dual costs in WM.

In the study, 6–8-year-olds, 10–14-year-olds, and young adults were presented with arrays of colored squares and asked to respond to a speeded reaction task during the retention interval; at the end of each trial, they had to recognize a probe as being part or not of the array. The speeded reaction task consisted of pressing a button on the same or opposite side of a visual target appearing on the screen or an auditory target played in one ear. The authors manipulated the difficulty of the speeded task by varying the instructions to participants. In the easy condition, the side of the motor response (press the button) was congruent to the side of the target (i.e., right or left side of the screen/ear). In the difficult condition, the motor response was to be given on the opposite side of the target. The reason for using both visual and auditory targets in the speeded task was to ensure that any dual-costs observed in WM were not modality-specific and instead were due to attentional sharing. Moreover, the authors manipulated the load of the memory array (0, 1, 2, 3, 4 items) in order to obtain the capacity estimates of participants in each age group. The zero-load condition serves as the baseline in the speeded reaction task, in which participants were not presented to any item to memorize.

In this type of task, participants with more proactive control of attention are expected to show smaller dual-cost in WM and less accurate, slower responses in the speeded task as the set size increases. The explanation is that they proactively devote attention to maintaining the items in WM during the speeded task, therefore they are less subject to decay during this period. Also, because maintaining larger arrays requires more attentional resources, performance in the speeded task will be penalized as the set size increases. On the other hand, participants with less proactive control will devote most of their attention to one task at a time and

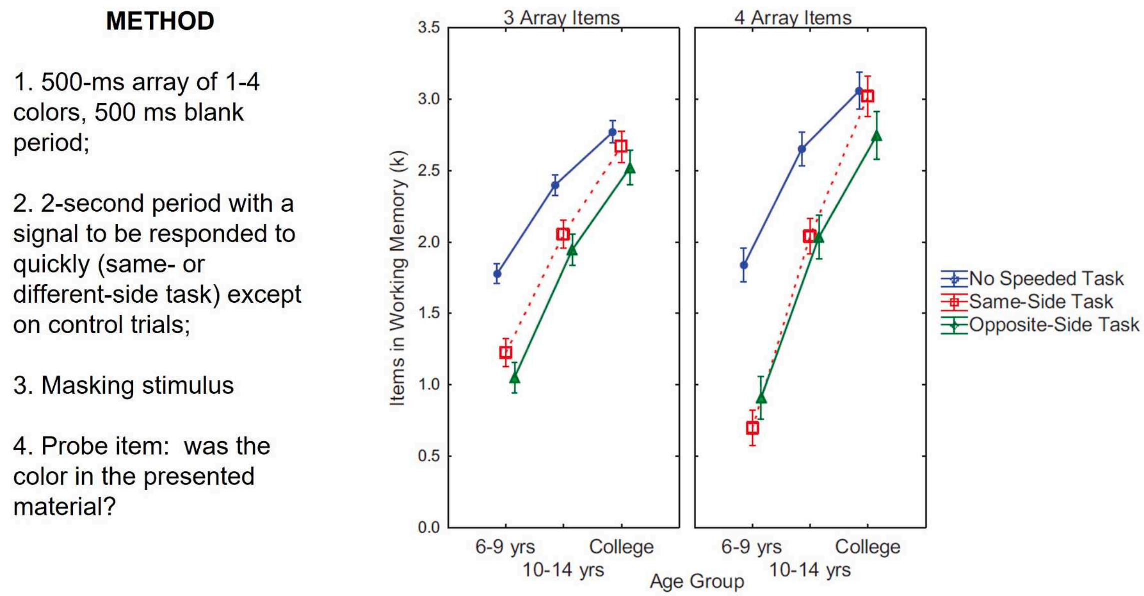

Fig. 4. *WM capacity in different array sizes, speeded-task conditions, and age groups.* The right half of the figure was reproduced from Cowan et al. (2021, Developmental Psychology, Fig. 4). Error bars are standard errors.

therefore will neglect maintenance in WM while doing the speeded task, causing a higher dual cost on WM performance. The authors found a pattern of results congruent with the hypothesis of children using a reactive stance in dual tasks. There was a main effect of age, with older children and adults having a larger WM capability than younger children, and a main effect of the speeded task, with all groups performing better in the WM task when no speeded task was present. However, the cost of the speeded task on WM was different across age groups, with adults' performance being barely affected by it. Fig. 4 illustrates these results.

Cowan et al. (2021) also analyzed the costs of maintaining items in WM upon performance in the speeded task. It, too, was consistent with the switch from a reactive stance in young children to a proactive stance in older children and adults. All participants had poorer performance in the opposite side condition, but the size of the memory arrays affected the age groups differently. Older children and adults showed an increase in reaction time in the speeded task accuracy as the arrays became larger, presumably because more attention was devoted to retaining the arrays, leaving less attention for speedy performance in the speeded task. (The older children only showed this pattern for the easier, same-side-responding task, however.) Younger children's speed was not affected by the array size, and the same was true for the older children in the opposite-side task. These results (Fig. 5) suggest that attention was not being divided much between the maintenance of the items and the speeded task in younger participants, as they apparently did not devote more attention to larger arrays to the detriment of the speeded task the way adults consistently did.

In sum, then, in step 6, a processual change in proactively sharing attention in dual-tasks settings will determine how WM capability grows during childhood.

Unresolved issues

We contend in our framework that knowledge is used to elaborate the material in WM and unload attention (steps 4 and 5), and that storage capacity of the focus of attention is also required to maintain the offloaded material, albeit to a lesser extent than holding items individually in the focus requires. We assume that this is an effortful, controlled process that improves during childhood. Our framework, however, leaves some unresolved issues on the automatic influence of knowledge on children's WM capability. Namely, it does not explain the automaticity of some knowledge effects during earlier steps of information processing nor the influence of cognizance (or *meta*-knowledge) on WM performance. In the following sections, we will address the automatic knowledge effects, discuss the development of *meta*-knowledge, and explain why we favor the view of an age improvement in the effortful application of knowledge. Then we will elaborate on the hypothesis that storage capacity of the focus of attention is divided between storing items individually and keeping track of information filed to peripheral portions of WM, after the detection/creation of patterns occurs.

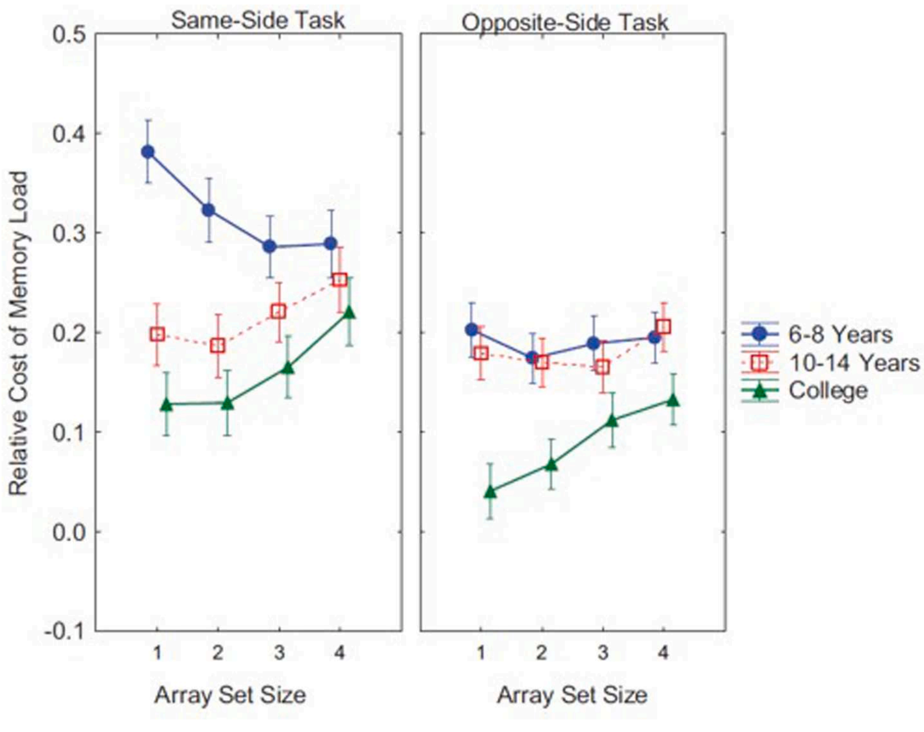

Fig. 5. *Speeded-task performance cost in terms of the proportion increase in reaction time compared to a no-load condition, for different responding conditions, array sizes, and age groups.* Reproduced from Cowan et al. (2021, Developmental Psychology, Figure 8). Error bars are standard errors.

The role of knowledge

The automatic influence of knowledge at encoding

We stressed in step 4 (pattern detection or creation) that our working hypothesis is that the ability to apply knowledge to combine items in the memoranda improves with age. We assume that the influence of knowledge over the creation of patterns is voluntary in the sense that it requires a controlled use of strategy during maintenance. However, we admit the possibility that knowledge also influences the detection of patterns in earlier steps of information processing, for example via perceptual grouping (Jiang et al., 2004) and via the encoding of higher-order structural organization of visual arrays (Brady & Tenenbaum, 2013) – two rather automatic processes.

Regarding the early detection of well-learned patterns at encoding, there is indeed evidence that adults are more hampered than children by violations of the configural pattern of a stimulus, at least in face recognition tasks (Carey & Diamond, 1994, Goldstein, 1975). For example, when asked to recognize familiar and unfamiliar faces, 6-year-olds are less affected by inversion (i.e., presenting the face upside down) than 10-year-olds, who are, in their turn, less affected than adults (Carey & Diamond, 1994). A possible explanation for such age effect in memory for faces is that adults were more exposed to faces during life, making them more reliant on the automatic encoding of a well-learned visual pattern. We contend, however, that although the automatic encoding of patterns might be influenced by knowledge and experience through life, it is likely not the driver of WM development in childhood.

The notion of an improvement in creating patterns in the memoranda came from the results reported by Cowan et al., (2018), already mentioned, in which it was observed that the peripheral components of WM, but not the central storage, increased drastically with age. This gave rise to the idea that adults somehow apply knowledge to elaborate the material and offload attention. Greene et al. (2020) used the same method as Cowan et al. (2018) to estimate the size of the peripheral and central components of WM in younger and older adults. They found the reversed developmental pattern as the one observed from childhood to adulthood: the capacity estimates of peripheral components in older adults are smaller than in younger adults, but the capacity estimate of the central component is the same. The complete developmental pattern, from childhood to old age, seems to indicate that central storage in WM remains stable throughout life, but the ability to file information in the peripheral portion of WM improves as people grow and declines as they age. Children acquire knowledge as they grow, but adults do not lose it as they age, therefore the automatic influence of knowledge does not seem to be the source of WM development in this situation. We believe that a processual change in the application of knowledge in favor of such "filing to the periphery" operation (named offloading in the present framework) is a more plausible mechanism to explain the capability growth in childhood and is most likely based on effortful, attention-demanding processing.

The automatic influence of knowledge at verbal recall. Another automatic effect of knowledge on WM is redintegration at recall, which we briefly discussed in the introductory section "Application of Knowledge". Redintegration is the reconstruction of decaying memory traces based on long-term phonological representations that emerge with speech production at response. We believe that the automaticity of redintegrative processes cannot be the driver of WM development based on the same argument we used to discuss the automatic encoding of visual patterns: that is, knowledge is greater in the elderly and redintegration is a preserved process in this population (Neale & Tehan, 2007). Nevertheless, WM span in verbal recall also declines in old age, including in word, non-word, and digit repetition tasks (Bopp & Verhaeghen, 2005; Grégoire & Van der Linden, 1997; Wingfield et al., 1988). If automatic knowledge processes like redintegration were the main determinant of WM capability in childhood, then we should also not observe a decline in verbal recall tasks in the elderly.

The influence of cognizance. Finally, there is the matter of cognizance, the self-awareness of one's cognitive processes, and its impacts on WM development. Cognizance has been proposed as a core element of cognitive development in childhood, even being called part of the "trinity of mind" along with executive control and fluid intelligence (Demetriou et al., 2018). It is theoretically possible that some steps of our framework are influenced by a child's level of cognizance. For example, knowing that attention has storage functions (e.g., "I remember more when I *pay attention to* the memoranda" could cause children to use their focus of attention more effectively in steps 6 (maintenance during distraction) by adopting a proactive stance of control in the presence of a distracting task. This is not, however, to say that WM develops as a result of enhancements in metacognition. It is more likely that improved metacognition is a result of WM development than the opposite.

Demetriou and colleagues (Demetriou et al., 2018; Spanoudis et al., 2015) investigated the role of cognizance in children's development in relation with measures of many cognitive abilities, such as WM, the executive control of attention, fluid intelligence (Gf), and processing speed. Cognizance was broken down into two different components, perceptual awareness and inferential awareness. The former relates to children's ability to correctly report that sensory processes are the source of a given knowledge (e.g., "I know because I saw it/heard it/hear someone saying it", etc.). The latter relates to children's ability to correctly report that inferential mental processes are the source of a given knowledge, that is, mental operations needed to take part for a conclusion to be reached (e.g., "I know because I though/assumed/I supposed that…"). Through structural equation modelling, Spanoudis et al. (2015) proposed a model in which inferential awareness is an essential mediator between the basic executive functions (WM and attentional control) and the development of Gf. The difference between the two best-fitting models was slight; however, the cognizance factor emerged as a mediator in both. In one of them, both inferential and perceptual awareness are mediators between WM and Gf (WM → inferential and perceptual awareness → Gf); in the other, preferred by the authors, attentional control mediates between WM and inferential awareness, which in turn mediates Gf (WM → attentional control → inferential awareness → Gf). It is germane to our point that the two best-fitting models include cognizance as a mediator between measures of WM and fluid intelligence across childhood, therefore suggesting that *meta*-knowledge is likely not a precursor of WM increase, but rather the opposite.

People's *meta*-judgments in WM tasks are sometimes good indicators of their performance, at least in the adult population (e.g., Adam & Vogel, 2017; Atkinson et al., 2018). Nevertheless, it's usually found in the literature that people overestimate their WM capability and are overconfident in their responses (Conte et al., 2023; Cowan et al., 2016; Sahar et al., 2020). The overestimation of WM capability is especially severe in younger children. Flavell et al. (1970) tested children from nursery school, kindergarten, second and fourth grade (ages not reported) in a serial recall task and asked them to predict how many words they could correctly remember. Younger children (from nursery and kindergarten) predicted they could correctly recall the audacious number of about 7–8 words in a list, but their actual span was about 3.5 words. Older children (second and fourth graders) predicted they could recall about 6 words, but their actual span was 4–5 words. Another example of how children overestimate their WM capability comes from a study by Forsberg et al. (2021), in which children from ages 6 to 13 years and adults were asked to answer how many colors from a visual array they held in WM, before recognizing a color probe. Note that, in this manipulation, participants were asked to estimate their WM capability (i.e., *k,* the number of items currently held in WM), prior to being tested, which makes it prior to any experience of actual success or failure on the trial. Although WM capability increased drastically with age, the *meta*-judgments did not differ across age groups, causing younger children to overestimate their WM capability more than older children or adults. Regardless of age and actual capability, all participants estimated it in about 3 to 4 items. The developmental effect, therefore, was not to improve the accuracy of the *meta*-judgments *per se* but rather increase WM capability while *meta*-judgments remained stable. It is not clear whether younger children are just over-optimistic or are correct in some way; for example, they may have conscious access to the items but not in a way that is sufficient or long-lasting enough to allow subsequent success in probe recognition.

In sum, it is theoretically possible that enhancements in automatic knowledge effects and in cognizance drive some of WM capability increase during childhood, but it is unlikely that these two factors account for all of its development.

The involvement of the focus of attention in maintaining offloaded patterns

In our framework, we assume that offloading is used to save attentional resources in situations in which some attention needs to be divided between tasks or sets of items, as was the case of the bimodal condition in the study by Cowan et al. (2018). The transfer of created patterns to activated long-term memory (step 5) is thought to unload storage from the focus of attention, but nonetheless it still relies on attention to operate successfully. It is possible that central storage capacity (storage in the focus of attention) is also used to keep track of the off-loaded set of items in the periphery, a suggestion made by Rhodes and Cowan (2018) and Cowan (2019) and currently being investigated in our laboratory. According to this proposal, the focus of attention is used to hold individual items in the memoranda while also holding "pointers" to groups of items or chunks held in activated long-term memory. The pointers presumably allow the focus of attention to keep track of the offloaded material and, in the appropriate time (e.g., at response), retrieve this group and "unpack" it again into individual items. For example, in encoding a list of 9 items (say, words), an adult might hold the last three words in the focus of attention along with pointers that keep track of the first two groups of three words that have been elaborated and offloaded to activated long-term memory (not always perfectly, of course). Given the list "*tree-cup-bus-fire-cart-donut-bug-pen-glasses*", the items *tree*, *cup*, *bus* would perhaps be kept in the focus of attention with pointers to the groups *fire-cart-donut* (e.g., a shopping cart full of donuts on fire) and *bug-pen-glasses* (e.g., an intellectual bug wearing glasses and holding a pen nib), both of them kept in activated long-term memory. Attentional resources need to be used at first to encode information (steps 1 to 3), and then what is left can be used to elaborate the material (step 4), offload it to activated long-term memory, and keep track of what has been offloaded (step 5), the whole process being reliant on the capabilities of the focus of attention. Ongoing work in our lab lends support to the idea that the storage space in the focus of attention is used in the "keeping track" operation, but we have not yet tested this hypothesis

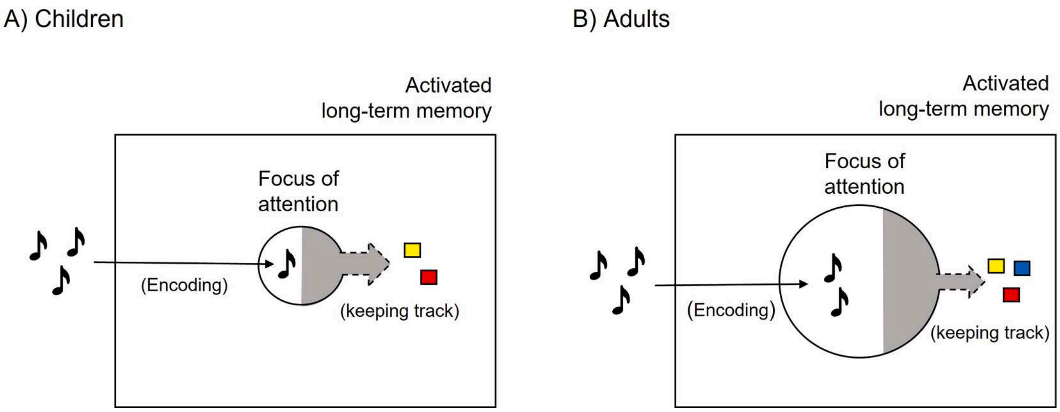

Fig. 6. *Hypothetical partition of attention between encoding new information and offloading items in the memoranda during input.* The figure illustrates our proposal for how the offloading process develops throughout childhood in a two-set WM procedure. The focus of attention needs to be divided between encoding items in a second-presented stimulus set (here, musical tones) and maintaining items in a prior set (here, colored squares). Offloading of the first set helps to minimize the load on attention during encoding of the second set. A developmental increase in attention (shown as a larger circle in Part B) allows both for holding more items in the focus of attention and for keeping track of offloaded items more efficiently. The shaded zones represent attentional resources devoted to maintaining the offloaded pattern in activated long-term memory, and the dashed arrows mean that the focus of attention is keeping track of it.

developmentally.

If it is indeed the case that the focus of attention is used to hold "pointers" to information held in activated long-term memory, lower capabilities of the focus in children could also explain why their peripheral *k* estimates were smaller in Cowan et al. (2018). Therefore, we do not consider that an age-related improvement in the creation and offloading of patterns (steps 4 and 5) eliminates the possibility of a "true" developmental storage increase in the focus of attention. Instead, we propose that there may well be a storage increase in the focus of attention taking place during childhood and that, with increasing knowledge and experience, this storage space also can be strategically spared by offloading meaningful conglomerates of items to the periphery – but keeping these offloaded items is not an operation that completely dispenses with central attention. An analogy is useful to illustrate our argument: with development, adults are able to carry more objects because their hands (storage capacity in the focus of attention) have grown bigger, but also because they learn how to use bags as an aid and partially free their hands (optimization of processing, also carried out by the focus of attention); some of the transportation job has been transferred to bags (offloaded in bundles to activated long-term memory), but it still relies on one's grasp so it is not dropped (the focus is used to keep track the offloaded material). Fig. 6 illustrates our developmental hypothesis of the offloading operation.

Finally, we reiterate that results from methods that prevent the implementation of strategies (e.g., span of apprehension, running span) – and therefore are assumed to reflect only the content of the focus of attention at test time – have shown an increase during childhood (Cowan et al., 2005), hence in our view tending to rule out the possibility of zero storage-growth in WM. Typically, adults perform reliably better than children in running span tasks (e.g., 3.87 versus 2.44 items, respectively, Cowan et al., 2005), but the age difference is enlarged in tasks allowing for the implementation of strategies that unload central attention, such as in Cowan et al. (2018) (see Fig. 1 in their article). In our proposal, the increased storage capacity of the focus of attention (detected by running spans measures) can be strategically employed to optimize processing and act as a multiplying factor of capacity in adults. This larger capacity is detected in tasks allowing for strategy use.

Future directions

Not all fundamental sources of WM development have been tested across all relevant age groups, and neither have all of them been oppositely tested as causal explanations in every study. What seems to be a fundamental driving factor of WM development at a certain age might as well not be the main factor years later – and such age comparisons have not been thoroughly tested out. For example, storage capacity seems to impose a constraint on the selective encoding of information during early school years (Cowan et al., 2010), favoring a storage account of development during childhood. However, it appears that there is some processing optimization taking place at encoding during adolescence (Cowan et al., 2011), probably from rehearsal and elaboration, at a time when capability measures in simple tasks that do not allow these strategies have already reached adult levels and supposedly reflect an asymptotic level of storage growth. Our framework aims at reconciling these and other seemingly contradictory evidence of WM development, by proposing a step-by-step description of how information is managed by the focus of attention in dual-task settings and assuming that either factor, storage or processing, has a decisive role at different steps.

We encourage researchers to broaden their age groups while testing hypotheses of WM growth in order to provide a more complete developmental panorama of the causal factors. Ideally, the field should aim at determining not only the order of magnitude of the factors at each step of information processing but also across different ages. This is essential to improving educational practices inasmuch as it allows us to tailor pedagogical materials to each age group. For instance, insisting on activities requiring shared attention between tasks may be an unfruitful strategy for younger children who still do not have a proactive stance of attention, but could yield good results in late elementary school years. The observation that attentional filtering is as effective in young children and adults provided that WM is not overloaded offers an important argument for why it is preferable to present them with fewer items to be processed at once.

An important avenue to explore concerns the partial offloading of patterns in WM, a process that we assume to happen in adults. There is a speculative gap in our model that needs to be tested and we are currently working on it. It concerns the possibility that patterns created in the focus of attention are offloaded to activated long-term memory, and that the focus of attention is used to maintain the offloaded patterns. The question of whether the size of the focus of attention increases or whether it used more efficiently with age – and in our framework, including to keep track of items in activated long-term memory – is one that has been revisited many times in the history of developmental science. The matter is controversial: even leading proponents of processing theories of WM development have acknowledged that it is hard to explain residual increase in WM measures after controlling for strategy use and processing speed without admitting growth in a core storage capacity (Case, 1995). As presented in this article, the exact nature of the role of the focus of attention in WM development, whether by storing more information and/or by optimizing processing, remains an open question.

Finally, testing how knowledge and experience can influence the creation of patterns in the memoranda is vital for the prospects of the theoretical framework here proposed. We are currently working on this theme by means of developing a paradigm that induces the creation of patterns by child participants. If offloading of attention indeed occurs when patterns are created, saving storage space that can be devoted to more items, then we should expect to see age-differences diminished when comparing children who were trained to create patterns to untrained adults. Then, of course, we should also test the hypothesis that adults (voluntarily or involuntarily) rely on pattern creation to offload attention.

The variation of experimental designs to tackle the offloading of attention are multiple and we are currently exploring them in our laboratory. If indeed we observe that pattern creation frees central storage space, a derived line of inquiry regards how many offloaded patterns the focus of attention can simultaneously keep track of, and by which means this operation is carried out. There must be a limitation in how much offload can be made by the focus of attention without incurring forgetting, and it is also possible that these limitations are minimized with age – in which case they would become another purported source of WM development in childhood and possibly another target of future investigation. We look forward to seeing how research in the field will develop, and we hope that the proposed framework encourages researchers to keep exploring the entanglements between storage and processing in the development of WM.

Conclusion

Through this article, we have addressed the important and long-standing debate on the influences of storage and processing in the development of WM during school years. By considering evidence favoring both accounts of WM development, we have proposed an integrative framework in which storage limitations fundamentally constrain WM capability in the early steps of information processing, whereas the efficiency of processing mechanisms fundamentally influences later steps. Critically, our framework assumes that the influence of both factors upon WM development is not mutually exclusive and that they can coexist at different steps of information processing. Our framework is still in its infancy and some of the proposed mechanisms remain speculative, but we believe it can provide developmental researchers with fruitful insight into how to reconcile opposing results in the literature – an aim that we will keep seeking in our own future research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

Funding: This research was supported by NIH Grant R01-HD021338 awarded to Cowan.

References

Adam, K. C., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive Psychology, 97*, 79–97. https://doi.org/10.1016/j.cogpsych.2017.07.001

Adam, K. C., & Vogel, E. K. (2017). Confident failures: Lapses of working memory reveal a metacognitive blind spot. *Attention, Perception, & Psychophysics, 79*, 1506–1523. https://doi.org/10.3758/s13414-017-1331-8

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

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Atkinson, A. L., Baddeley, A. D., & Allen, R. J. (2018). Remember some or remember all? Ageing and strategy effects in visual working memory. *Quarterly Journal of Experimental Psychology, 71*(7), 1561–1573. https://doi.org/10.1080/17470218.2017.1341537

Atkinson, J., & Braddick, O. (2012). Visual attention in the first years: Typical development and developmental disorders. *Developmental Medicine & Child Neurology, 54*(7), 589–595. https://doi.org/10.1111/j.1469-8749.2012.04294.x

Akyürek, E. G., Kappelmann, N., Volkert, M., & van Rijn, H. (2017). What you see is what you remember: Visual chunking by temporal integration enhances working memory. *Journal of Cognitive Neuroscience, 29*(12), 2025–2036. https://doi.org/10.1162/jocn_a_01175

AuBuchon, A. M., & Wagner, R. L. (2023). Self-generated strategies in the phonological similarity effect. *Memory & Cognition*, 1–19. https://doi.org/10.3758/s13421-023-01418-2. Advance online publication.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

Baddeley, A. D. (1986). *Working memory*. Oxford Psychology Series, Voume 11. Oxford: Clarendon Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 48–79). Academic Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior, 14*(6), 575–589. https://doi.org/10.1016/S0022-5371(75)80045-4

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental Psychology, 45*(2), 477–490. https://doi.org/10.1037/a0014615

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review, 118*(2), 175–192. https://doi.org/10.1037/a0022324

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General, 132*(1), 71–92. https://doi.org/10.1037/0096-3445.132.1.71

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321*(5890), 851–854. https://doi.org/10.1126/science.1158023

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia, 49*(6), 1622–1631. https://doi.org/10.1016/j.neuropsychologia.2010.12.023

Berger, S. E., Harbourne, R. T., & Horger, M. N. (2018). Cognition–action trade-offs reflect organization of attention in infancy. *Advances in Child Development and Behavior, 54*, 45–86. https://doi.org/10.1016/bs.acdb.2017.11.001

Bertrand, R., & Camos, V. (2015). The role of attention in preschoolers' working memory. *Cognitive Development, 33*, 14–27. https://doi.org/10.1016/j.cogdev.2014.10.002

Bjorklund, D. F. (1985). The role of conceptual knowledge in the development of organization in children's memory. *Basic Processes in Memory Development: Progress in Cognitive Development Research, 103–142*. https://doi.org/10.1007/978-1-4613-9541-6_3

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 60*(5), 223–233. https://doi.org/10.1093/geronb/60.5.P223
Bower, G. H., & Bolton, L. S. (1969). Why are rhymes easy to learn? *Journal of Experimental Psychology, 82*(3), 453–461. https://doi.org/10.1037/h0028365
Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychological Review, 120*(1), 85–109. https://doi.org/10.1037/a0030779
Bray, N. W., & Turner, L. A. (1986). The rehearsal deficit hypothesis. *International Review of Research in Mental Retardation, 14*, 47-71. Academic Press. https://doi.org/10.1016/S0074-7750(08)60074-9.
Broadbent, D. E. (1958). Perception and communication. *Pergamon Press*. https://doi.org/10.1037/10037-000
Broadway, J. M., & Engle, R. W. (2010). Validating running memory span: Measurement of working memory capacity and links with fluid intelligence. *Behavior Research Methods, 42*, 563–570. https://doi.org/10.3758/BRM.42.2.563
Brockmole, J. R., & Logie, R. H. (2013). Age-related change in visual working memory: A study of 55,753 participants aged 8–75. *Frontiers in Psychology, 4*, 12. https://doi.org/10.3389/fpsyg.2013.00012
Brown, A. L. (1974). The role of strategic behavior in retardate memory. *International review of research in mental retardation, 7*, 55–111. https://doi.org/10.1016/S0074-7750(08)60050-6
Bunting, M. F., Cowan, N., & Colflesh, G. H. (2008). The deployment of attention in short-term memory tasks: Tradeoffs between immediate and delayed deployment. *Memory & Cognition, 36*, 799–812. https://doi.org/10.3758/MC.36.4.799
Burtis, P. J. (1982). Capacity increase and chunking in the development of short-term memory. *Journal of Experimental Child Psychology, 34*(3), 387–413. https://doi.org/10.1016/0022-0965(82)90068-6
Buss, A. T., Ross-Sheehy, S., & Reynolds, G. D. (2018). Visual working memory in early development: A developmental cognitive neuroscience perspective. *Journal of Neurophysiology, 120*(4), 1472–1483. https://doi.org/10.1152/jn.00087.2018
Camos, V., & Barrouillet, P. (2011). Developmental change in working memory strategies: From passive maintenance to active refreshing. *Developmental Psychology, 47*(3), 898–904. https://doi.org/10.1037/a0023193
Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory?. *Annals of the New York Academy of Sciences*, 1424(1), 19-32. https://doi.org/10.1111/nyas.13616.
Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition, 1*(2–3), 253–274. https://doi.org/10.1080/13506289408402302
Case, R. (1995). Capacity-based explanations of working memory growth: A brief history and reevaluation. In F. E. Weinert, & W. Schneider (Eds.), *Memory performance and competencies: Issues in growth and development* (pp. 23–44). Lawrence Erlbaum Associates Inc.
Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*(3), 386–404. https://doi.org/10.1016/0022-0965(82)90054-6
Chechile, R. A., & Richman, C. L. (1982). The interaction of semantic memory with storage and retrieval processes. *Developmental Review, 2*(3), 237–250. https://doi.org/10.1016/0273-2297(82)90013-2
Chevalier, N., & Blaye, A. (2016). Metacognitive monitoring of executive control engagement during childhood. *Child Development, 87*(4), 1264–1276. https://doi.org/10.1111/cdev.12537
Chi, M. T. (1976). Short-term memory limitations in children: Capacity or processing deficits? *Memory & Cognition, 4*(5), 559–572. https://doi.org/10.3758/BF03213219
Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73–96). Lawrence Erlbaum Associates Inc.
Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62*, 73–101. https://doi.org/10.1146/annurev.psych.093008.100427
Clark, K. M., Hardman, K., Schachtman, T. R., Saults, J. S., Glass, B. A., & Cowan, N. (2018). Tone series and the nature of working memory capacity development. *Developmental Psychology, 54*, 663–676. https://doi.org/10.1037/dev0000466
Comalli, P. E., Wapner, S., & Werner, H. (1962). Interference Effects of Stroop Color-Word Test in Childhood, Adulthood, and Aging. *The Journal of Genetic Psychology, 100*(1), 47–53. https://doi.org/10.1080/00221325.1962.10533572
Conklin, H. M., Luciana, M., Hooper, C. J., & Yarger, R. S. (2007). Working memory performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology, 31*(1), 103–128. https://doi.org/10.1080/87565640709336889
Conte, N., Fairfield, B., Padulo, C., & Pelegrina, S. (2023). Metacognition in working memory: Confidence judgments during an n-back task. *Consciousness and Cognition, 111*, Article 103522. https://doi.org/10.1016/j.concog.2023.103522
Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*(12), 547–552. https://doi.org/10.1016/j.tics.2003.10.005
Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*(2), 163–191. https://doi.org/10.1037/0033-2909.104.2.163
Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language, 31*, 668–684. https://doi.org/10.1016/0749-596X(92)90034-U
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.
Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922
Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19*(1), 51–57. https://doi.org/10.1177/0963721409359277
Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review, 26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y
Cowan, N. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science, 11*, 239–264. https://doi.org/10.1177/1745691615621279
Cowan, N. (2017a). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review, 24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6
Cowan, N. (2017b). Mental objects in working memory: Development of basic capacity or of cognitive completion? *Advances in Child Development and Behavior, 52*, 81–104. https://doi.org/10.1016/bs.acdb.2016.12.001
Cowan, N. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin, 145*, 822–847. https://doi.org/10.1037/bul0000199
Cowan, N. (2022). Working memory development: A 50-year assessment of research and underlying theories. *Cognition, 224*, Article 105075. https://doi.org/10.1016/j.cognition.2022.105075
Cowan, N., AuBuchon, A. M., Gilchrist, A. L., Blume, C. L., Boone, A. P., & Saults, J. S. (2021). Developmental change in the nature of attention allocation in a dual task. *Developmental Psychology, 57*(1), 33–46. https://doi.org/10.1037/dev0001134
Cowan, N., AuBuchon, A. M., Gilchrist, A. L., Ricker, T. J., & Saults, J. S. (2011). Age differences in visual working memory capacity: Not based on encoding limitations. *Developmental Science, 14*(5), 1066–1074. https://doi.org/10.1111/j.1467-7687.2011.01060.x
Cowan, N., Chen, Z., & Rouder, J. N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science, 15*(9), 634–640. https://doi.org/10.1111/j.0956-7976.2004.00732.x

Cowan, N., Elliott, E.M., & Saults, J.S. (2002). The search for what is fundamental in the development of working memory. In R. Kail & H. Reese (Eds.), Advances in Child Development and Behavior (Vol. 29, pp. 1-49. https://doi.org/10.1016/S0065-2407(02)80050-7.

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*(1), 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

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., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition, 34*, 1754–1768. https://doi.org/10.3758/bf03195936

Cowan, N., Hardman, K., Saults, J. S., Blume, C. L., Clark, K. M., & Sunday, M. A. (2016). Detection of the number of changes in a display in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 169–185. https://doi.org/10.1037/xlm0000163

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), e12635.

Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science, 13*(1), 120–133. https://doi.org/10.1111/j.1467-7687.2009.00864.x

Cowan, N., Nugent, L. D., Elliott, E. M., Ponomarev, I., & Saults, J. S. (1999). The role of attention in the development of short-term memory: Age differences in the verbal span of apprehension. *Child Development, 70*(5), 1082–1097. https://doi.org/10.1111/1467-8624.00080

Cowan, N., Ricker, T. J., Clark, K. M., Hinrichs, G. A., & Glass, B. A. (2015). Knowledge cannot explain the developmental growth of working memory capacity. *Developmental Science, 18*(1), 132–145. https://doi.org/10.1111/desc.12197

Cowan, N., Saults, J. S., & Blume, C. L. (2014). Central and peripheral components of working memory storage. *Journal of Experimental Psychology: General, 143*(5), 1806–1836. https://doi.org/10.1037/a0036814

Cowan, N., Wood, N. L., Wood, P. K., Keller, T. A., Nugent, L. D., & Keller, C. V. (1998). Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General, 127*, 141–160. https://doi.org/10.1037//0096-3445.127.2.141

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Darwin, C. J., Turvey, M. T., & Crowder, R. C. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology, 3*(2), 255–267. https://doi.org/10.1016/0010-0285(72)90007-2

Davies, & Parasuraman, R. (1982). *The psychology of vigilance / D.R. Davies, R. Parasuraman.* Academic Press.

Demetriou, A., Makris, N., Kazi, S., Spanoudis, G., & Shayer, M. (2018). The developmental trinity of mind: Cognizance, executive control, and reasoning. *Wiley Interdisciplinary Reviews: Cognitive Science, 9*(4), e1461.

Dempster, F. N. (1978). Memory span and short-term memory capacity: A developmental study. *Journal of Experimental Child Psychology, 26*(3), 419–431. https://doi.org/10.1016/0022-0965(78)90122-4

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

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12*(1), 45–75. https://doi.org/10.1016/0273-2297(92)90003-k

Doyle, A. B. (1973). Listening to distraction: A developmental study of selective attention. *Journal of Experimental Child Psychology, 15*(1), 100–115. https://doi.org/10.1016/0022-0965(73)90134-3

Einstein, G. O., McDaniel, M. A., & Lackey, S. (1989). Bizarre imagery, interference, and distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(1), 137–146. https://doi.org/10.1037/0278-7393.15.1.137

Elliott, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory & Cognition, 30*, 478–487. https://doi.org/10.3758/bf03194948

Elliott, E. M., Morey, C. C., AuBuchon, A. M., Cowan, N., Jarrold, C., Adams, E. J., Attwood, M., Bayram, B., Beeler-Duden, S., Blakstvedt, T. Y., Büttner, G., Castelain, T., Cave, S., Crepaldi, D., Fredriksen, E., Glass, B. A., Graves, A. J., Guitard, D., Hoehl, S., … 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.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 44, pp. 145–199). Elsevier Science.

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review, 119*(2), 223–271. https://doi.org/10.1037/a0027371

Fitamen, C., Blaye, A., & Camos, V. (2019a). Five-year-old children's working memory can be improved when children act on a transparent goal cue. *Scientific Reports, 9*(1), 15342. https://doi.org/10.1038/s41598-019-51869-4

Fitamen, C., Blaye, A., & Camos, V. (2019b). The role of goal cueing in kindergarteners' working memory. *Journal of Experimental Child Psychology, 187*, Article 104666. https://doi.org/10.1016/j.jecp.2019.104666

Fitamen, C., Blaye, A., & Camos, V. (2022). Do goal cue and motor activity impact preschoolers' working memory? *British Journal of Developmental Psychology, 40*(1), 1–16. https://doi.org/10.1111/bjdp.12376

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, 283–299*. https://doi.org/10.2307/1126804

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology, 1*(4), 324–340. https://doi.org/10.1016/0010-0285(70)90019-8

Formoso, J., Jacubovich, S., Injoque-Ricle, I., & Barreyro, J. P. (2018). Resolution of arithmetic problems, processing speed and working memory in children. *Trends in Psychology, 26*, 1249–1266. https://doi.org/10.9788/TP2018.3-05En

Forrester, W. E., & King, D. J. (1971). Effects of semantic and acoustic relatedness on free recall and clustering. *Journal of Experimental Psychology, 88*(1), 16–19. https://doi.org/10.1037/h0030653

Forsberg, A., Adams, E. J., & Cowan, N. (2022). Why does visual working memory ability improve with age: More objects, more feature detail, or both? A registered report. *Developmental Science, e13283*. https://doi.org/10.1111/desc.13283

Forsberg, A., Blume, C. L., & Cowan, N. (2021). The development of metacognitive accuracy in working memory across childhood. *Developmental Psychology, 57*(8), 1297–1317. https://doi.org/10.1037/dev0001213

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Gaillard, V., Barrouillet, P., Jarrold, C., & Camos, V. (2011). Developmental differences in working memory: Where do they come from? *Journal Of Experimental Child Psychology, 110*, 469–479. https://doi.org/10.1016/j.jecp.2011.05.004

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 54A*(1), 1–30. https://doi.org/10.1080/02724980042000002

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working Memory Skills and Educational Attainment: Evidence from National Curriculum Assessments at 7 and 14 Years of Age. *Applied Cognitive Psychology, 18*(1), 1–16. https://doi.org/10.1002/acp.934

Gilchrist, A. L., Cowan, N., & Naveh-Benjamin, M. (2008). Working memory capacity for spoken sentences decreases with adult ageing: Recall of fewer but not smaller chunks in older adults. *Memory (Hove, England), 16*(7), 773–787. https://doi.org/10.1080/09658210802261124

Gilchrist, A. L., Cowan, N., & Naveh-Benjamin, M. (2009). Investigating the childhood development of working memory using sentences: New evidence for the growth of chunk capacity. *Journal of Experimental Child Psychology, 104*(2), 252–265. https://doi.org/10.1016/j.jecp.2009.05.006

Goldstein, A. G. (1975). Recognition of inverted photographs of faces by children and adults. *The Journal of Genetic Psychology: Research and Theory on Human Development, 127*(1), 109–123. https://doi.org/10.1080/00221325.1975.10532361

Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of WM in young school-age children and its relation to intelligence. *Journal of Memory and Language, 92*, 183–201. https://doi.org/10.1016/j.jml.2016.06.004

Gray, S., Levy, R., Alt, M., Hogan, T. P., & Cowan, N. (2022). WM predicts new word learning over and above existing vocabulary and nonverbal IQ. *Journal of Speech, Language, and Hearing Research, 65*, 1044–1069. https://doi.org/10.1044/2021_JSLHR-21-00397

Greene, N. R., Naveh-Benjamin, M., & Cowan, N. (2020). Adult age differences in working memory capacity: Spared central storage but deficits in ability to maximize peripheral storage. *Psychology and Aging, 35*(6), 866–880. https://doi.org/10.1037/pag0000476

Grégoire, J., & Van der Linden, M. (1997). Effect of age on forward and backward digit spans. *Aging, Neuropsychology, and Cognition, 4*(2), 140–149. https://doi.org/10.1080/13825589708256642

Guttentag, R. E., Ornstein, P. A., & Siemens, L. (1987). Children's spontaneous rehearsal: Transitions in strategy acquisition. *Cognitive Development, 2*, 307–326. https://doi.org/10.1016/S0885-2014(87)80010-2

Hagen, J. W. (1967). The effect of distraction on selective attention. *Child Development, 38*(3), 685–694. https://doi.org/10.2307/1127246

Hagen, J. W., & Hale, G. A. (1973). The development of attention in children. In A. D. Pick (Ed.), *Minnesota symposium on child psychology* (Vol. 7). University of Minnesota Press.

Halford, G. S. (1993). In *Children's Understanding: The Development of Mental Models* (1st ed.). Psychology Press. https://doi.org/10.4324/9781315801803.

Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science, 16*(1), 70–76. https://doi.org/10.1111/j.0956-7976.2005.00782.x

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences, 11*(6), 236–242. https://doi.org/10.1016/j.tics.2007.04.001

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21*(6), 803–831. https://doi.org/10.1017/S0140525X98001769

Hardman, K. O., & Cowan, N. (2015). Remembering complex objects in visual WM: Do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(2), 325. https://doi.org/10.1037/xlm0000134

Hester, R. L., Kinsella, G. J., & Ong, B. (2004). Effect of age on forward and backward span tasks. *Journal of the International Neuropsychological Society, 10*(4), 475–481. https://doi.org/10.1017/S1355617704104037

Hitch, G. J. (1978). The role of short-term WM in mental arithmetic. *Cognitive Psychology, 10*(3), 302–323. https://doi.org/10.1016/0010-0285(78)90002-6

Hitch, G. J., Halliday, M. S., & Littler, J. E. (1989). Item identification time and rehearsal rate as predictors of memory span in children. *The Quarterly Journal of Experimental Psychology Section A, 41*(2), 321–337. https://doi.org/10.1080/14640748908402368

Hulme, C., & Muir, C. (1985). Developmental changes in speech rate and memory span: A causal relationship? *British Journal of Developmental Psychology, 3*(2), 175–181. https://doi.org/10.1111/j.2044-835X.1985.tb00969.x

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(5), 1217–1232. https://doi.org/10.1037/0278-7393.23.5.1217

Hulme, C., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology, 47*(1), 72–87. https://doi.org/10.1016/0022-0965(89)90063-5

Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual WM for faces. *Journal of Experimental Psychology: Human Perception and Performance, 34*(3), 556–568. https://doi.org/10.1037/0096-1523.34.3.556

Jarrold, C., Cocksey, J., & Dockerill, E. (2008). Phonological similarity and lexicality effects in children's verbal short-term memory: Concerns about the interpretation of probed recall data. *The Quarterly Journal of Experimental Psychology, 61*(2), 324–340. https://doi.org/10.1080/17470210701202210

Jiang, Y., Chun, M. M., & Olson, I. R. (2004). Perceptual grouping in change detection. *Perception & Psychophysics, 66*, 446–453. https://doi.org/10.3758/BF03194892

Jolicœur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36*(2), 138–202. https://doi.org/10.1006/cogp.1998.0684

Jones, G. (2012). Why chunking should be considered as an explanation for developmental change before short-term memory capacity and processing speed. *Frontiers in psychology, 3*, 167. https://doi.org/10.3389/fpsyg.2012.00167

Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition, 144*, 1–13. https://doi.org/10.1016/j.cognition.2015.07.009

Jones, G., Justice, L. V., Cabiddu, F., Lee, B. J., Iao, L. S., Harrison, N., & Macken, B. (2020). Does short-term memory develop? *Cognition, 198*, Article 104200. https://doi.org/10.1016/j.cognition.2020.104200

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6*(5), 498–504. https://doi.org/10.1111/1467-7687.00307

Kail, R., & Park, Y. S. (1994). Processing time, articulation time, and memory span. *Journal of Experimental Child Psychology, 57*(2), 281–291. https://doi.org/10.1006/jecp.1994.1013

Káldy, Z., & Leslie, A. M. (2005). A memory span of one? Object identification in 6.5-month-old infants. *Cognition, 97*(2), 153–177. https://doi.org/10.1016/j.cognition.2004.09.009

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in WM capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane (Eds.) & A. Miyake & J. N. Towse (Ed.), Variation in WM (pp. 21–46). Oxford University Press.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*(1), 47–70. https://doi.org/10.1037/0096-3445.132.1.47

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of WM capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Kannass, K. N., Oakes, L. M., & Shaddy, D. J. (2006). A longitudinal investigation of the development of attention and distractibility. *Journal of Cognition and Development, 7*(3), 381–409. https://doi.org/10.1207/s15327647jcd0703_8

Kavšek, M. (2004). Predicting later IQ from infant visual habitation and dishabituation: A meta-analysis. *Journal of Applied Developmental Psychology, 25*, 369–393. https://doi.org/10.1016/j.appdev.2004.04.006

Kerst, S., & Levin, J. R. (1973). A comparison of experimenter-provided and subject-generated strategies in children's paired-associate learning. *Journal of Educational Psychology, 65*(3), 300–303. https://doi.org/10.1037/h0035545

Kestner, J., & Brokowski, J. G. (1979). Children's maintenance and generalization of an interrogative learning strategy. *Child Development, 50*(2), 485–494. https://doi.org/10.2307/1129427

Kibbe, M. M., & Leslie, A. M. (2013). What's the object of object WM in infancy? Unraveling 'what' and 'how many'. *Cognitive Psychology, 66*(4), 380–404. https://doi.org/10.1016/j.cogpsych.2013.05.001

Kibbe, M. M., & Leslie, A. M. (2016). The ring that does not bind: Topological class in infants' WM for objects. *Cognitive Development, 38*, 1–9. https://doi.org/10.1016/j.cogdev.2015.12.001

Krinsky, R., & Krinsky, S. G. (1996). Pegword mnemonic instruction: Retrieval times and long-term memory performance among fifth-grade children. *Contemporary Educational Psychology, 21*(2), 193–207. https://doi.org/10.1006/ceps.1994.1018
Lane, D. M., & Pearson, D. A. (1982). The Development of Selective Attention. *Merrill-Palmer Quarterly, 28*(3), 317–337. http://www.jstor.org/stable/23086119.
Lehmann, M., & Hasselhorn, M. (2007). Variable memory strategy use in children's adaptive intratask learning behavior: Developmental changes and WM influences in free recall. *Child Development, 78*(4), 1068–1082. https://doi.org/10.1111/j.1467-8624.2007.01053.x
Lemaire, B., Pageot, A., Plancher, G., & Portrat, S. (2018). What is the time course of WM attentional refreshing? *Psychonomic Bulletin & Review, 25*(1), 370–385. https://doi.org/10.3758/s13423-017-1282-z
Leslie, L. (1975). Susceptibility to interference effects in short-term memory of normal and retarded readers. *Perceptual and Motor Skills, 40*(3), 791–794. https://doi.org/10.2466/pms.1975.40.3.791
Leslie, L. (1980). Mediation or production deficiency in disabled readers? *Perceptual and Motor Skills, 50*(2), 519–530. https://doi.org/10.1177/003151258005000230
Lilburn, S. D., Smith, P. L., & Sewell, D. K. (2019). The separable effects of feature precision and item load in visual short-term memory. *Journal of Vision, 19*(1), 2. https://doi.org/10.1167/19.1.2
Logie, R. H., & Pearson, D. G. (1997). The inner eye and the inner scribe of visuo-spatial WM: Evidence from developmental fractionation. *European Journal of Cognitive Psychology, 9*(3), 241–257. https://doi.org/10.1080/713752559
Luciana, M., & Nelson, C. A. (2002). Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-year-old children. *Developmental neuropsychology, 22*(3), 595–624. https://doi.org/10.1207/S15326942DN2203_3
Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal WM and executive control processes in adolescents. *Child Development, 76*(3), 697–712. https://doi.org/10.1111/j.1467-8624.2005.00872.x
Luck, S. J., & Vogel, E. K. (1997). The capacity of visual WM for features and conjunctions. *Nature, 390*(6657), 279–281. https://doi.org/10.1038/36846
Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of WM. *Nature Neuroscience, 17*(3), 347–356. https://doi.org/10.1038/nn.3655
Maccoby, E. E., & Hagen, J. W. (1965). Effects of distraction upon central versus incidental recall: Developmental trends. *Journal of Experimental Child Psychology, 2*(3), 280–289. https://doi.org/10.1016/0022-0965(65)90031-7
Majerus, S., Péters, F., Bouffier, M., Cowan, N., & Phillips, C. (2018). The dorsal attention network reflects both encoding load and top-down control during WM. *Journal of Cognitive Neuroscience, 30*, 144-159. https://doi.org/10.1162/jocn_a_01195.
Mall, J. T., Morey, C. C., Wolf, M. J., & Lehnert, F. (2014). Visual selective attention is equally functional for individuals with low and high working memory capacity: Evidence from accuracy and eye movements. *Attention, Perception, & Psychophysics, 76*, 1998–2014. https://doi.org/10.3758/s13414-013-0610-2
Marcovitch, S., Boseovski, J. J., Knapp, R. J., & Kane, M. J. (2010). Goal neglect and working memory capacity in 4-to 6-year-old children. *Child development, 81*(6), 1687–1695. https://doi.org/10.1111/j.1467-8624.2010.01503.x
Mashburn, C. A., Tsukahara, J. S., & Engle, R. W. (2020). Individual differences in attention control. In R. H. Logie, N. Cowan, & V. Camos (Eds.), *Working memory: State of the science* (pp. 175–211). Oxford University Press. https://doi.org/10.1093/oso/9780198842286.003.0007.
Mathy, F., & Friedman, O. (2020). Working memory develops at a similar rate across diverse stimuli. *Journal of Experimental Child Psychology, 191*, Article 104735. https://doi.org/10.1016/j.jecp.2019.104735
McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience, 11*(1), 103–107. https://doi.org/10.1038/nn2024
Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100. https://doi.org/10.1006/cogp.1999.0734
Murray, D. J. (1965). Vocalization-at-presentation and immediate recall, with varying presentation-rates. *The Quarterly Journal of Experimental Psychology, 17*(1), 47–56. https://doi.org/10.1080/17470216508416407
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
Ngiam, W. X. Q., Khaw, K. L. C., Holcombe, A. O., & Goodbourn, P. T. (2019). Visual working memory for letters varies with familiarity but not complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(10), 1761–1775. https://doi.org/10.1037/xlm0000682
Oakes, L. M., Messenger, I. M., Ross-Sheehy, S., & Luck, S. J. (2009). New evidence for rapid development of color-location binding in infants' visual short-term memory. *Visual Cognition, 17*(1–2), 67–82. https://doi.org/10.1080/13506280802151480
Oakes, L. M., Ross-Sheehy, S., & Luck, S. J. (2006). Rapid development of feature binding in visual short-term memory. *Psychological Science, 17*(9), 781–787. https://doi.org/10.1111/j.1467-9280.2006.01782.x
Oberauer, K. (2019). Working memory capacity limits memory for bindings. *Journal of Cognition, 2*(1), Article 40. https://doi.org/10.5334/joc.86
Oberauer, K., & Kliegl, R. (2001). Beyond resources: Formal models of complexity effects and age differences in working memory. *European Journal of Cognitive Psychology, 13*(1–2), 187–215. https://doi.org/10.1080/09541440042000278
Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language, 55*(4), 601–626. https://doi.org/10.1016/j.jml.2006.08.009
Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review, 19*(5), 779–819. https://doi.org/10.3758/s13423-012-0272-4
Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32*(4), 301–345. https://doi.org/10.1016/0001-6918(70)90108-3
Pickering, S. J. (2006). *Working memory and education*. San Diego: Academic.
Plebanek, D. J., & Sloutsky, V. M. (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science, 28*(6), 723–732. https://doi.org/10.1177/0956797617693005
Plebanek, D. J., & Sloutsky, V. M. (2019). Selective attention, filtering, and the development of working memory. *Developmental Science, 22*(1), e12727.
Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: A life-span overview. *Acta Psychologica, 86*(2–3), 227–272. https://doi.org/10.1016/0001-6918(94)90004-3
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, 48*(2), 384–404. https://doi.org/10.1080/14640749508401396
Poole, B. J., & Kane, M. J. (2009). Working-memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *Quarterly Journal of Experimental Psychology, 62*(7), 1430–1454. https://doi.org/10.1080/17470210802479329
Pressley, M. (1982). Elaboration and memory development. *Child Development, 53*(2), 296–309. https://doi.org/10.2307/1128972
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
Rhodes, S., & Cowan, N. (2018). Attention in working memory: Attention is needed but it yearns to be free. *Annals of the New York Academy of Science, 1424*, 52–63. https://doi.org/10.1111/nyas.13652f
Ricker, T. J., Sandry, J., Vergauwe, E., & Cowan, N. (2020). Do familiar memory items decay? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*, 60–76. https://doi.org/10.1037/xlm0000719
Ricker, T. J., & Vergauwe, E. (2022). Boundary conditions for observing cognitive load effects in visual working memory. *Memory & Cognition, 1–17*. https://doi.org/10.3758/s13421-022-01320-3
Roediger, H. L. (1980). The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology: Human Learning and Memory, 6*(5), 558–567. https://doi.org/10.1037/0278-7393.6.5.558

Rose, N. S. (2020). The dynamic-processing model of working memory. *Current Directions in Psychological Science, 29*(4), 378–387. https://doi.org/10.1177/0963721420922185

Roggeman, C., Klingberg, T., Feenstra, H. E. M., Compte, A., & Almeida, R. (2014). Trade-off between capacity and precision in visuospatial working memory. *Journal of Cognitive Neuroscience, 26*(2), 211–222. https://doi.org/10.1162/jocn_a_00485

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development, 74*(6), 1807–1822. https://doi.org/10.1046/j.1467-8624.2003.00639.x

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia, 42*(8), 1029–1040. https://doi.org/10.1016/j.neuropsychologia.2003.12.012

Rueda, M. R., Posner, M. I., Rothbart, M. K., & Davis-Stober, C. P. (2004). Development of the time course for processing conflict: An event-related potentials study with 4 year olds and adults. *BMC Neuroscience, 5*(1), 1–13. https://doi.org/10.1186/1471-2202-5-39

Ruff, H. A., & Capozzoli, M. C. (2003). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology, 39*(5), 877–890. https://doi.org/10.1037/0012-1649.39.5.877

Sahadevan, S. S., Chen, Y. Y., & Caplan, J. B. (2021). Imagery-based strategies for memory for associations. *Memory, 29*(10), 1275–1295. https://doi.org/10.1080/09658211.2021.1978095

Sahar, T., Sidi, Y., & Makovski, T. (2020). A metacognitive perspective of visual working memory with rich complex objects. *Frontiers in Psychology, 11*, 179. https://doi.org/10.3389/fpsyg.2020.00179

Sarigiannidis, I., Crickmore, G., & Astle, D. E. (2016). Developmental and individual differences in the precision of visuospatial memory. *Cognitive Development, 39*, 1–12. https://doi.org/10.1016/j.cogdev.2016.02.004

Saults, J. S., & Cowan, N. (2007). A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology: General, 136*(4), 663. https://doi.org/10.1037/0096-3445.136.4.663

Siegel, L. S. (1994). Working memory and reading: A life-span perspective. *International Journal of Behavioral Development, 17*(1), 109–124. https://doi.org/10.1177/016502549401700107

Schneegans, S., & Bays, P. M. (2016). No fixed item limit in visuospatial working memory. *Cortex, 83*, 181–193. https://doi.org/10.1016/j.cortex.2016.07.021

Schneider, W., Gruber, H., Gold, A., & Opwis, K. (1993). Chess expertise and memory for chess positions in children and adults. *Journal of Experimental Child Psychology, 56*(3), 328–349. https://doi.org/10.1006/jecp.1993.1038

Schneider, W., Körkel, J., & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology, 81*(3), 306–312. https://doi.org/10.1037/0022-0663.81.3.306

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition, 21*(2), 168–175. https://doi.org/10.3758/BF03202729

Shipstead, Z., & Engle, R. W. (2013). Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(1), 277. https://doi.org/10.1037/a0028467

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2015). Working memory capacity and the scope and control of attention. *Attention, Perception, & Psychophysics, 77*, 1863–1880. https://doi.org/10.3758/s13414-015-0899-0

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working Memory Capacity and Fluid Intelligence: Maintenance and Disengagement. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science, 11*(6), 771–799. https://doi.org/10.1177/1745691616650647

Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology, 111*(4), 695–707. https://doi.org/10.1016/j.jecp.2011.10.007

Simms, N. K., Frausel, R. R., & Richland, L. E. (2018). Working memory predicts children's analogical reasoning. *Journal of Experimental Child Psychology, 166*, 160–177. https://doi.org/10.1016/j.jecp.2017.08.005

Smith, L. B., Kemler, D. G., & Aronfreed, J. (1975). Developmental trends in voluntary selective attention: Differential effects of source distinctness. *Journal of Experimental Child Psychology, 20*(2), 352–362. https://doi.org/10.1016/0022-0965(75)90109-5

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

Spanoudis, G., Demetriou, A., Kazi, S., Giorgala, K., & Zenonos, V. (2015). Embedding cognizance in intellectual development. *Journal of Experimental Child Psychology, 132*, 32–50. https://doi.org/10.1016/j.jecp.2014.12.003

Strutt, G. F., Anderson, D. R., & Well, A. D. (1975). A developmental study of the effects of irrelevant information on speeded classification. *Journal of Experimental Child Psychology, 20*(1), 127–135. https://doi.org/10.1016/0022-0965(75)90032-6

Süß, H. M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence, 30*, 261–288. https://doi.org/10.1016/S0160-2896(01)00100-3

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory? *Journal of experimental psychology. Learning, memory, and cognition, 45*(1), 37–55. https://doi.org/10.1037/xlm0000578

Thompson, L. A. (1994). Dimensional strategies dominate perceptual classification. *Child development, 65*(6), 1627–1645. https://doi.org/10.1111/j.1467-8624.1994.tb00839.x

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428*(6984), 751–754. https://doi.org/10.1038/nature02466

Turner, J. E., Henry, L. A., & Smith, P. T. (2000). The development of the use of long-term knowledge to assist short-term recall. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 53A*(2), 457–478. https://doi.org/10.1080/027249800390565

Turner, J. E., Henry, L. A., Smith, P. T., & Brown, P. A. (2004). Redintegration and lexicality effects in children: Do they depend upon the demands of the memory task? *Memory & cognition, 32*(3), 501–510. https://doi.org/10.3758/BF03195842

Turnure, J., Buium, N., & Thurlow, M. (1976). The effectiveness of interrogatives for promoting verbal elaboration productivity in young children. *Child Development, 47*(3), 851–855. https://doi.org/10.2307/1128206

Veer, I. M., Luyten, H., Mulder, H., van Tuijl, C., & Sleegers, P. J. (2017). Selective attention relates to the development of executive functions in 2, 5-to 3-year-olds: A longitudinal study. *Early childhood research quarterly, 41*, 84–94. https://doi.org/10.1016/j.ecresq.2017.06.005

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*(7067), 500–503. https://doi.org/10.1038/nature04171

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual WM. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1436–1451. https://doi.org/10.1037/0096-1523.32.6.1436

Ward, T. B. (1980). Separable and integral responding by children and adults to the dimensions of length and density. *Child Development, 676–684*. https://doi.org/10.2307/1129452

Wingfield, A., Stine, E. A. L., Lahar, C. J., & Aberdeen, J. S. (1988). Does the capacity of working memory change with age? *Experimental Aging Research, 14*(2), 103–107. https://doi.org/10.1080/03610738808259731

Xie, W., & Zhang, W. (2017). Dissociations of the number and precision of visual short-term memory representations in change detection. *Memory & Cognition, 45*(8), 1423–1437. https://doi.org/10.3758/s13421-017-0739-7

Zosh, J. M., & Feigenson, L. (2012). Memory load affects object individuation in 18-month-old infants. *Journal of Experimental Child Psychology, 113*(3), 322–336. https://doi.org/10.1016/j.jecp.2012.07.005

Zukier, H., & Hagen, J. W. (1978). The development of selective attention under distracting conditions. *Child Development, 870–873*. https://doi.org/10.2307/1128259