

The Relation Between Attention and Memory

Nelson Cowan,¹ Chenye Bao,¹
Brittney M. Bishop-Chrzanowski,¹ Amy N. Costa,¹
Nathaniel R. Greene,¹ Dominic Guitard,²
Chenyuan Li,¹ Madison L. Musich,¹ and Zehra E. Ünal¹

¹Department of Psychological Sciences, University of Missouri, Columbia, Missouri, USA;
email: CowanN@missouri.edu

²School of Psychology, Cardiff University, Cardiff, United Kingdom

Annu. Rev. Psychol. 2024. 75:183–214

First published as a Review in Advance on
September 15, 2023

The *Annual Review of Psychology* is online at
psych.annualreviews.org

<https://doi.org/10.1146/annurev-psych-040723-012736>

Copyright © 2024 by the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See credit lines of images or other third-party material in this article for license information.

ANNUAL
REVIEWS **CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

selective attention, working memory, long-term memory, neuroscience, attention and memory, development

Abstract

The relation between attention and memory has long been deemed important for understanding cognition, and it was heavily researched even in the first experimental psychology laboratory by Wilhelm Wundt and his colleagues. Since then, the importance of the relation between attention and memory has been explored in myriad subdisciplines of psychology, and we incorporate a wide range of these diverse fields. Here, we examine some of the practical consequences of this relation and summarize work with various methodologies relating attention to memory in the fields of working memory, long-term memory, individual differences, life-span development, typical brain function, and neuropsychological conditions. We point out strengths and unanswered questions for our own embedded processes view of information processing, which is used to organize a large body of evidence. Last, we briefly consider the relation of the evidence to a range of other theoretical views before drawing conclusions about the state of the field.

Contents

INTRODUCTION	184
ATTENTION AND MEMORY CONCEPTS IN THE PRESENT REVIEW	185
Attention	185
Memory	185
SUBTLITIES OF THE ATTENTION–MEMORY RELATION	186
CURRENT THEORETICAL FRAMEWORK	186
Mechanisms of Embedded Processes	186
Historical Roots of the Embedded Processes Approach Linking Attention and Memory	188
PRACTICAL CONSEQUENCES OF THE ATTENTION–MEMORY RELATION	188
ATTENTION AND WORKING MEMORY	189
Effects of Attention on Working Memory	189
Effects of Working Memory on Attention	190
Computational Modeling of Attention and Working Memory	191
ATTENTION AND LONG-TERM MEMORY	192
Attention and Explicit Long-Term Memory	193
Attention and Implicit Long-Term Memory	193
Computational Modeling of Attention and Long-Term Memory	193
INDIVIDUAL DIFFERENCES	194
LIFE-SPAN DEVELOPMENT	196
Infant and Child Development	196
Adult Aging	198
ATTENTION, MEMORY, AND THE BRAIN	199
NEUROPSYCHOLOGICAL CONDITIONS	200
THEORETICAL IMPLICATIONS	201
Support for Embedded Processes	202
Evolution of Embedded Processes	202
Relation to Other Theoretical Approaches	203
CONCLUSION	205

INTRODUCTION

The relation between how we attend and what we remember is a fundamental and important relation within the human cognitive system. Attention can be described as the mental processes that select and prioritize some information for further consideration, given limits in human capability. Memory can be described as one's mental record of the past. The term "mental" is important. Being deaf in one ear narrows one's reception of stimuli, but it still is not an act of selective attention because it is not a mental process. Similarly, the loss of hearing from an explosion is a physical record of the event but not a memory—i.e., not a mental representation.

Attention helps determine what will be remembered and, consequently, how we prepare for the future. Conversely, memories influence how we direct our attention. We integrate work on the relation between attention and memory across many subdisciplines to further a theoretical understanding.

We first consider varieties of attention and memory and tools for a useful reading. In the first section, we present the embedded processes theoretical framework (Cowan 1988, 1995, 1999, 2019) to understand the attention–memory relation and illustrate areas of practical importance. We next examine relevant research involving working memory (WM) and long-term memory (LTM), explore individual and life-span developmental differences, and examine normal brain function and neuropsychological cases. The aim is to achieve theoretical coherence among these areas and guide further research. The embedded processes approach is sufficiently specific and evidence based to serve as our theoretical guide but is further tuned here based on the current evidence.

ATTENTION AND MEMORY CONCEPTS IN THE PRESENT REVIEW

Attention

A key aspect of attention is selectivity. There are many concurrent incoming stimuli and ideas from past experiences, but one can only think about a small portion of them at the same time. James [1892 (1985)] famously described selective attention as the mind seizing upon some information at the expense of other information. Selectivity can be further dissected into the scope or capacity of attention, or how much information can be attended at once (Cowan et al. 2005), and the control of attention, or how the target of attention is determined (Cowan et al. 2006). Voluntary control often must struggle against involuntary processes such as mind wandering (Kane et al. 2007) or attention capture (e.g., in the attentional blink, attention becomes briefly unavailable for new targets while still processing a current one; see Petersen & Vangkilde 2022).

Other basic qualities of attention are alertness or arousal, the capability to attend, and its intensity (Unsworth et al. 2022). Alertness depends on one’s physiological and mental state, decreasing, for example, with sleepiness or hunger. It increases gradually when one has coffee and suddenly when one receives an alerting (orienting) signal discrepant from the current neural model of the environment (Sokolov 1963). Maintaining alertness continually during a tedious task is termed vigilance (Davies & Parasuraman 1982) or the consistency of attention (Unsworth & Miller 2021, Unsworth et al. 2022). Attention may be temporarily depleted following even subtle demands, such as comprehending a word low in frequency of occurrence in the language (Popov & Reder 2020). Selectivity and alertness are interdependent, in that high alertness should assist selective attention, and selecting one’s attention wisely should assist alertness. One’s current goals contribute to selectivity and alertness (Madore & Wagner 2022).

Memory

There has been confusion about types of memory. We define WM as the ensemble of mental components that hold limited information temporarily in a heightened state of availability for use in ongoing information processing. Cowan (2017) compared this generic definition to others in the field, as definitions have varied widely. WM, as we define it, includes short-lived sensory information about multiple incoming stimuli, currently activated (primed) semantic concepts, and more integrated information in a limited-capacity, attention-related system holding up to several separate chunks of information concurrently (Cowan 1988, 2019). We make no distinction between WM and short-term storage here. Other views use WM for the attention-dependent part of temporary memory and short-term memory for the attention-independent part (Engle 2002); WM for complex span tasks and short-term memory for simple span tasks (Daneman & Carpenter 1980); or WM for a multicomponent system, with short-term memory probably seen as an outmoded term (Baddeley 1986).

Working memory (WM): the small amount of information that can be held in a temporarily heightened state of availability

Long-term memory (LTM): the brain’s repository of a lifetime of learning, including episodic (representing specific events), semantic, and procedure-based varieties

Embedded processes approach: Cowan’s (1988, 1999, 2019) information processing model with the focus of attention embedded in activated long-term memory

Neural model of the environment: a presumed pattern of neural activity reflecting an individual’s current knowledge about the environment

Activated portion of long-term memory (aLTM):

the part of memory from which information is in a heightened state of accessibility

Focus of attention (FoA):

coherent representation of several separate items or ideas guiding current thoughts and actions

LTM is information acquired over the life span. Explicit LTM is available for conscious recollection, making it generally more attention dependent than implicit memory, which comprises learning effects of which participants may be unaware (Schacter 1990). These types of LTM could exist on a continuum (Dew & Cabeza 2011).

SUBTLETIES OF THE ATTENTION–MEMORY RELATION

Memory can influence attention. For example, skills that at first require attention, such as finding letters from a target set within an array, become automatic after many trials with the same target set (Shiffrin & Schneider 1977). Conversely, attention critically affects explicit memory (Dew & Cabeza 2011). However, the relation between attention and memory is nuanced. In unconscious priming, a briefly presented word followed by a mask so quickly that the word cannot be detected facilitates retrieval of a second word with a similar meaning (Marcel 1983). However, the flashed word that causes priming apparently leaves no long-term trace for later recognition (Balota 1983).

Analysis of an allegedly unattended channel can falsely seem automatic. Eich (1984) used selective listening to one speech channel in one ear, to be repeated (shadowed), while a channel in the other ear was to be ignored and contained word pairs such as *taxi-fare*. Subsequently, participants were to spell one spoken word per trial, and spelling was influenced by the unattended word pairs (e.g., increasing the frequency of *fare* rather than *fair*). However, this effect of allegedly ignored speech disappeared when the shadowing task was presented at a faster rate more typical of such studies (Wood et al. 1997).

Information held in an auditory sensory form decays over a few seconds unless at least some attention is devoted to it (Sperling 1960). Cowan et al. (1990) tested memory for intermittent spoken syllables presented during silent reading of a novel and found dramatic memory loss for the most recent syllable as the silent interval between its presentation and a syllable-recall cue (a light) increased from 1 to 10 seconds. However, when participants were to monitor the acoustic channel for one syllable, /dI/, while reading, memory for the syllables was stable across 10 seconds even though syllable detection was only at 60%. There are discrepancies in the WM literature resolved by the insight that memory decays rapidly when attention to each item during its presentation is insufficient (e.g., concurrent visual arrays; Ricker et al. 2020) but not when attention is higher (e.g., word lists; Oberauer & Lewandowsky 2008). Discrepancies between methods or definitions often underlie discrepancies between results rather than unreliability of evidence.

CURRENT THEORETICAL FRAMEWORK

The relation between attention and memory has been reviewed several times previously (e.g., Chun & Turk-Browne 2007, Cowan 1995, Norman 1968, Oberauer 2019, van Ede & Nobre 2023), but our review includes an especially large scope of areas for this relation, using the embedded processes framework to strive for coherence across areas.

Mechanisms of Embedded Processes

The embedded processes theoretical view emphasizes the attention–memory relation. Other relevant frameworks exist, of course, and are considered in the final section of the article. The key components of WM within our framework, illustrated in **Figure 1**, include the activated portion of LTM (aLTM) and, within it, the focus of attention (FoA). The aLTM is limited by time (for poorly encoded items, generally less than a minute) and interference from similar items, whereas the FoA is limited to about 3–5 unassociated items, or chunks (Cowan 2001).

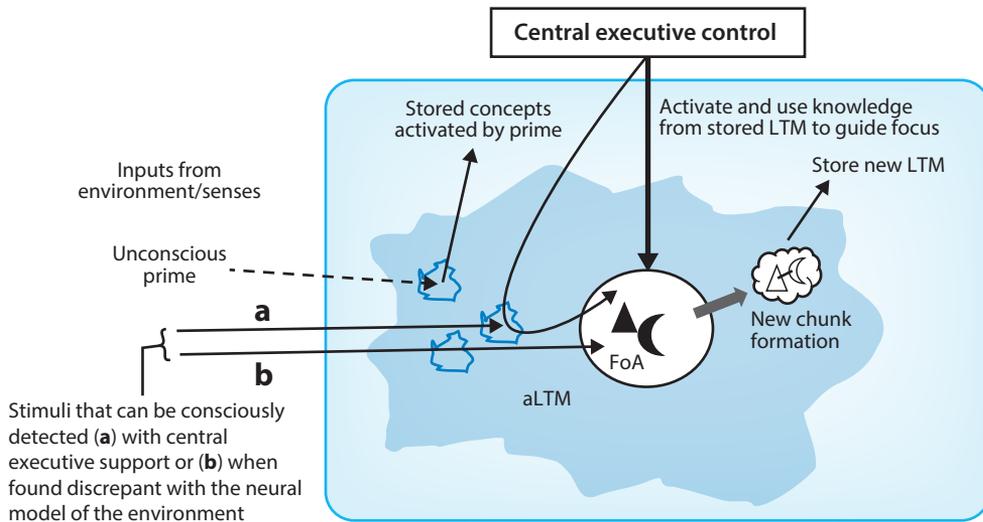


Figure 1

Schematic representation of attention and long-term memory (LTM) in an embedded processes view. Inputs from the environment pass into an activated subset of LTM (aLTM), represented by the large, irregular shape. Some subset of this information passes into the focus of attention (FoA), which is severely limited in capacity. Solid arrows from the environment represent information entering the FoA, represented as two shapes. Knowledge from stored LTM can be used to create structures (e.g., new chunks) from stimuli currently in the FoA, enabling the information to be offloaded out of the FoA into aLTM (cloud with conjoined shapes) and stored as a new LTM. Primes presented either without conscious, explicit awareness (*dashed arrow*, representing input from the environment) or with awareness can activate stored concepts from LTM, which in turn can more easily pass related content to the FoA.

The framework emphasizes several points. (a) The FoA is jointly controlled by abrupt or particularly meaningful changes in the environment and voluntary executive processes; the latter produce goal-directed control and could be influenced through instructions. (b) New, integrated compounds of ideas concurrently in the FoA rapidly form new LTM representations. (c) Outside of the FoA, aLTM (including rapid new learning) serves as a readily accessible store to be accessed by the FoA for cognitive processing. This theoretical framework is in keeping with views in which attention underlies individual differences in both storage and processing (e.g., Kane et al. 2004). The three points outlined above are discussed in turn.

Joint control of the focus of attention. What determines the information entering the central portion of WM? We presumably form a neural model of the environment (a WM not limited to information in awareness), and attention is captured by stimuli discrepant with this neural model (Sokolov 1963). Elliott & Cowan (2001) demonstrated this process with a cross-modal Stroop procedure using a distracting spoken color (e.g., *red*) on labeling of a visually presented color (e.g., a blue spot). Distraction was less potent when there were pre-exposures to the spoken word before the color-naming trial, allowing incorporation of the distractor into the neural model (see also Röer et al. 2015, 2019). Staying on task requires habituation to task-irrelevant stimuli and overcoming dishabituation to new distractions (e.g., baby noises in the lecture hall). Voluntary central executive processes accomplish this. For habituation to occur, distractors may have to enter the FoA for sufficient processing to be included in the neural model.

Habituation: waning of attention as a stimulus is repeated, presumably as the neural model of the environment adapts

New compounds in the focus of attention. Items in focus concurrently can be bound together to form a new concept, which is then maintained as a new representation in aLTM, outside of the FoA (**Figure 1**). For example, if one thinks about green ice, two elements are conjoined in what may be a new concept in memory. The complexity of a concept depends on how many independent ideas are interrelated (Halford et al. 2007). For example, the folk concept of a tiger requires one to keep in mind that it is large (as opposed to a house cat), striped (as opposed to a lion), and a cat (as opposed to a zebra). Young children exhibit overgeneralizations (e.g., calling a horse a dog; Gershkoff-Stowe 2001) and underextensions (e.g., using the word flower only for roses; White 1982), possibly because of the inability to think about all relevant features concurrently. In adults, Jiang & Cowan (2020) showed that the ability to remember which words were presented within the same list was best for items near the end of the list, presumably because they occupied the FoA longer than other items. Fleeting presence in the FoA may not suffice to produce WM (Chen & Wyble 2016).

Most theories tacitly allow for rapid new learning of attended material. For example, Keppel & Underwood (1962) found that on the first few trials, people could recall a consonant trigram after 18 seconds of distraction, but not later in the experiment. Residual memory of each trial's trigram might interfere with retrieval on subsequent trials. Despite this agreement among theorists, the consequences of new learning within aLTM on the current trial are often unappreciated.

Activated long-term memory as an accessible store. A key assertion of the embedded processes framework is that aLTM is a temporary form of memory with activation levels beyond the baseline level in memory. In a simple demonstration of this, McKone & Dennis (2000) presented words or nonwords at intervals of 2 seconds and, for each item, required a word/nonword judgment. The repetition of an item speeded responding. This repetition priming effect was reduced as the number of intervening items increased, but only up to several items. Over time, each word becomes less active in aLTM and, therefore, less effective as a prime. Activation was partly modality specific and partly general across modalities.

Historical Roots of the Embedded Processes Approach Linking Attention and Memory

Wilhelm Wundt, who developed the first laboratory of experimental psychology, was already interested in the relation between attention and temporary memory. To Cowan's surprise, Wundt already had an embedded processes theory (Cowan & Rachev 2018), as illustrated in **Supplemental Figure S1**. James [1892 (1985)], inspired by Wundt, described primary memory as the trailing edge of the conscious present. When Ebbinghaus [1885 (1913)] famously tested himself on previously studied lists of syllables, he found that for short lists, the material could be remembered on the "first fleeting grasp" (p. 33), a phrase suggesting attention and WM. Although experimental psychology has come far since this foundational work, we still follow its trail.

PRACTICAL CONSEQUENCES OF THE ATTENTION–MEMORY RELATION

There are myriad ways in which the attention–memory relation influences everyday activities, as illustrated in **Supplemental Table S1**. In the embedded processes approach, executive processes working with the FoA account for how well WM information is processed and how learning occurs. Higher WM capacity in attention-demanding tasks is associated with better general fluid intelligence (e.g., Conway et al. 2003, Cowan et al. 2005), arithmetic performance (Passolunghi & Siegel 2001), algebraic performance (Ünal et al. 2022), reading comprehension (Arrington et al.

Supplemental Material >

2014), and word learning beyond what standardized tests indicate (Gray et al. 2022). Paying attention to instructions from an instructor depends on WM capacity and the control of attention (Jaroslawska et al. 2016).

Attempts to train WM and attention have showed mixed results. According to Demetriou et al. (2014), it could be important to train a child's metaknowledge or conscious awareness of their own memory system. It might be useful to train critical thinking skills that depend on attention and memory rather than training attention and memory directly (see Halpern 1998). Forsberg et al. (2021b) found that children in the early elementary school years overestimate their WM capacity more than older children or adults, which could lead young children to assume that they do not need mnemonic strategies. The success of WM training for transfer to useful skills may depend on how the training increases participants' awareness and control of mental processes (e.g., Chambers et al. 2008).

The embedded processes approach has not often examined emotions or stress, but these are important for cognition. Attention and WM can be impaired by stress, increasing vulnerability to cognitive overload (Matthews et al. 2020). When a crime is committed, memory for it depends partly on how stress is handled. When people focus attention narrowly, they may experience inattentive blindness and, as a result, not even notice unattended events, such as incidentals of the crime scene (Levett et al. 2021).

ATTENTION AND WORKING MEMORY

We distinguish between effects of attention on WM and the converse. We then discuss how computational modeling has been involved in this area.

Effects of Attention on Working Memory

A small amount of attended information is saved for immediate memory tasks, whereas information that is unattended during its presentation is more quickly lost. Broadbent (1958) summarized research on selective listening, showing that people could retain only the last few seconds from speech channels to be ignored, whereas they knew most of what occurred in the attended channel. Sperling (1960) showed rapid loss of characters from a briefly presented visual array of many items, with preserved memory of a row of a few items if the row to be retained was cued within several hundred milliseconds. This indicated a short-lived sensory afterimage, coupled with a small-capacity WM for attended information. Darwin et al. (1972) obtained similar results with a spatiotemporal arrangement of spoken words, but with a longer-lasting estimate of sensory memory (up to several seconds). Treisman & Rostron (1972) obtained the same with tones.

A key question is the extent to which attention plays a role during maintenance in WM. Attention might be used to refresh items in memory (Barrouillet et al. 2011, Raye et al. 2007), prioritize retention of some items (Cowan & Morey 2007, Hu et al. 2016, Lepsien et al. 2011), remove irrelevant information (Oberauer et al. 2012), or enhance the memory representation (e.g., Ricker & Vergauwe 2022). If two different types of information (for example, spatial arrays of visual objects and lists of spoken words) are to be retained in WM concurrently, the role of attention to be expected depends on the degree of modularity. If visual and verbal materials are stored separately, there should not be interference between them, in contrast to the embedded processes view in which the FoA is used for storage in a manner general across sensory modalities and codes. Uittenhove et al. (2019) showed that there is relatively little interference when the task is to recognize an item from one of the sets but a lot of interference when the task is to recall items. Vergauwe et al. (2022) combined a list-recall task with a process after each item (e.g., remembering locations one at a time while answering questions about rhymes between locations

Computational modeling:
use of computers and mathematics to simulate a complex system, allowing quantitative predictions of behavior and brain function

or the symmetry of items, and then recalling the locations), known as a complex span task. They found that the similarity of the kind of materials stored with the kind of processing made no difference at any list length, pointing to WM storage in a general, attention-based store. The distinction between recognition and recall makes sense if temporarily activated representations of the features sometimes suffice for recognition, whereas more attention-based memory must be restored for recall (e.g., Cowan 2019). However, the interference between very different materials is observable even in recognition (e.g., Cowan et al. 2014, Morey et al. 2011).

There may be a special role of attention in maintaining binding (associations between items, between an item and its serial or spatial position, or between features of an item). The embedded processes approach (e.g., Cowan 2019) sets the expectation that binding occurs in the FoA. Consistent with the embedded processes view, studies in which some bindings are prioritized more than others show that more than one item, though probably less than four, can be prioritized concurrently (Allen & Ueno 2018, Hitch et al. 2018, Souza & Oberauer 2016). Note, however, that WM performance levels for items and binding are affected equally by distraction (e.g., recognition of colored shapes; Allen et al. 2012). However, typically the level of performance is lower for binding, which means that the proportion of memory lost through distraction is higher in the case of binding than in the case of items. Items have sources of activation that may not help binding, though still there are item capacity limits (Cowan 2022b).

Guitard and colleagues (Guitard & Cowan 2023a,b; Guitard et al. 2021, 2022) have asked whether encoding and maintaining word lists in WM involves attentional resources allocated selectively to items or to their order. Order information is one type of relational binding between each item and its serial position in the list or between successive items. Clearly, one cannot retain order information without any item information, but one can retain order with only some item information. Guitard et al. (2022) examined the use of attention at encoding. A list was presented, and the participant was encouraged to prepare for an item test (fragment reconstruction; e.g., *s_en_* for spent), an order test (order reconstruction), or the possibility of either kind of test. The need to prepare for either test resulted in a loss of performance relative to the prepare-for-one-task conditions, especially for order. Guitard & Cowan (2023a,b) showed, however, that more time for encoding each item was important not for encoding but for maintenance. Guitard et al. (2021) also supported that conclusion. They presented one or two lists and manipulated whether an item or an order test was expected for each one. There was an effect of having two lists to remember for both items and order. Additionally, the similarity of the tasks for the two lists mattered—with poorer performance when the two lists were of the same type—primarily for order. Overall, order memory requires more commitment of attention during maintenance than does item information.

Effects of Working Memory on Attention

There are several ways in which WM representations affect attention. A neural model of the environment can be compared to the incoming stimulation, and discrepancies can attract attention (Cowan 1988, Elliott & Cowan 2001). That process may serve as the mechanism of an attentional filter, with abrupt changes in stimulation attracting attention, a common phenomenon that an attentional filter concept previously could not explain.

Wolfe (2021) reviewed evidence for the use of WM for guided visual search and concluded that aLTM holds an unlimited number of target templates (e.g., pictures of objects one is looking for), whereas the capacity-limited, attention-demanding part of WM (termed WM by Wolfe) holds up to only a few top-down guiding templates. This theory seems broadly in keeping with the embedded processes framework, though with regard to types of memory activation that do not seem vulnerable to rapid decay. It is still unclear why so much categorized information can stay

activated when it is explicitly needed (as in Wolfe's study) but may diminish rapidly when the participant is not trying to preserve it (e.g., McKone & Dennis 2000).

In sum, attention and WM mutually influence each other (cf. Draheim et al. 2022, van Ede & Nobre 2023). There could be a cycle of causation in which, for example, an attention lapse could cause a search template to be dropped from WM, which then impairs the continuing search process. Conversely, if one is reading a text passage, forgetting a key premise could lead to inattention to important points within the text while reading further, resulting in poorer-quality information in WM.

Computational Modeling of Attention and Working Memory

Unresolved questions about the link between attention and WM might be tackled with more explicit, quantitatively specified theories (Oberauer & Lewandowsky 2019). We welcome computational modeling when feasible. Models vary in scope and in what they can accomplish. At the broadest level of analysis, it is possible to make many assumptions about explicit processes to allow quantitative predictions of diverse sorts of behavior. An example is the adaptive control of thought approach of Anderson et al. (2004). There is an intention module with a goal buffer, a declarative memory module with a retrieval buffer, a production module, and separate modules for different senses. Operations on types of activation in the modules allow quantitative predictions of behavior. What this modeling type accomplishes is the presentation of a plausible set of mechanisms at a holistic level. If some of the assumptions are wrong, the actual processes might differ. Anderson's assumptions seem consistent with the embedded processes approach, except that the capacity-limited construct in Anderson's approach is activation rather than the FoA.

In modeling with a narrower scope, one can look at a single trait of cognition as a numerical process. This requires assumptions about processing but not extensive assumptions across all stages of processing. An example is provided by Cowan et al. (2012), who modeled recognition of singletons, word pairs, or triplets with known associations (e.g., *little black book*) within lists. Given the limited scope of modeling, it was feasible to compare several different models that differed only in a few assumptions: whether capacity was limited by the number of items or by multi-item chunks, whether this chunking was sometimes only partial (e.g., remembering *black book* but forgetting *little*), and whether aLTM supplemented the chunk capacity limit. The model that was most successful allowed for incomplete chunks. In general, about three chunks could be retained, no matter whether they were unrelated words or multiword phrases. One exception, however, was a condition with 18 singletons. Recognition of them was better than expected based on capacity limits, so an additional, rapid long-term learning component had to be introduced. This kind of modeling can sway our preference toward one flavor of model as opposed to another, though some viable processes are omitted (e.g., decay). The model helps shape the embedded processes approach by (a) confirming the importance of chunk capacity limits and (b) emphasizing the necessity of including rapid long-term learning.

A simple model suggests that there is not only a chunk capacity limit (about three units in adults) but also a limit in how many features per item can be included. Such a model was used successfully to fit several data sets (Cowan et al. 2013, Hardman & Cowan 2015, Oberauer & Eichenberger 2013).

In a still more focused application of computational modeling, one can examine very specific processes pitted against one another. This type of modeling may also help sharpen the embedded processes approach. For example, several models have been used to account for the effect of cognitive load, a decline in recall appearing as a linear function of the proportion of presentation time that is taken up by a distracting task. Information about the memoranda might decay,

Refreshing: boosting the level of neural activity of an item or idea by holding it in the focus of attention

Supplemental Material >

making necessary free time for the representations to be refreshed (Barrouillet et al. 2011), or there might be interference from the distracting material, making necessary free time for the unwanted representations to be removed (Oberauer et al. 2012). Both possibilities can be represented numerically (see **Supplemental Figure S2**) to understand what is to be expected as a function of the amount of time and the schedule and number of distracting events. Slightly different versions of this sort of model can be assessed. For example, although decay-based theories commonly assume that attentional refreshing occurs in the order in which items are presented, Lemaire & Portrat (2018; see also Lemaire et al. 2018) showed that the least-activated memory items may be refreshed first, regardless of their serial position. Lemaire and colleagues also supported the possibility that multiple items within the FoA can be refreshed concurrently (see also Gilchrist & Cowan 2011).

Evaluation of a particular computational approach can change suddenly with the introduction of new data. The account of cognitive load effects may be altered by the finding (Ricker & Vergauwe 2022) that an effect of cognitive load did not emerge in some circumstances. Ricker & Vergauwe (2022) suggested that memory loss may be prevented through enrichment of representations when time permits, a stabilization process that does not involve either repeated partial loss and refreshing of memory or mental removal of distracting items. In another recent development, a blank interval between two list items appears to assist WM performance for items presented after the interval but not before it as one would expect from either refreshment or removal of interference (Mizrak & Oberauer 2021, Ricker & Hardman 2017). There may be depletion of attentional resources, which recover during these intervals (Popov & Reder 2020). Thus, although computational modeling can sharpen verbal theories like the embedded processes framework, new empirical evidence still plays a key role.

ATTENTION AND LONG-TERM MEMORY

The relation between attention and LTM is likely to be bidirectional, as depicted in the embedded processes model in **Figure 1**. A fundamental assumption of our model is that the FoA acts as an encoding bottleneck for LTM retention. WM capacity limitations constrain how much information becomes available in LTM (e.g., Forsberg et al. 2021a, Fukuda & Vogel 2019).

The assumption that we must attend to learn is common but was sometimes called into question by the idea that learning might happen during sleep. Hugo Gernsback wrote a science fiction novel in 1911 titled *Ralph 124C 41+*, which included a sleep-learning device, the Hypnobioscope [Gernsback 1911 (2014)]. Alois B. Saliger invented the Psycho-Phone in 1927 that played inspirational messages during sleep (Bryan 2009). A recent review (Ruch & Henke 2020) indicates that learning during some phases of sleep is indeed possible, but with severe drawbacks. It is not consciously accessible or explicit learning, and it may even interfere with conscious learning of related material.

Dividing attention at encoding influences conscious recollection and explicit memory but does not prevent a sense of familiarity with the material or implicit memory. For example, Jacoby et al. (1989) found that people tended to judge that a name previously presented under divided-attention conditions was famous because the actual source of familiarity was forgotten.

In the popular procedure developed by Hebb (1961), a particular list is repeated multiple times throughout a recall session, whereas other lists are not repeated. Guérard et al. (2011) found that benefits of repetition occur similarly with or without awareness of the repetition, even though the performance level may be higher in participants who become aware of the repetition. The effect has also been found in a densely amnesic individual without awareness of learning new information (Gagnon et al. 2004).

Attention and Explicit Long-Term Memory

Divided attention procedures (e.g., Craik et al. 2018) suggest that some commitment of attention is necessary during encoding to establish a new episodic LTM. Accurately retrieving LTMs is less attention dependent, though the speed of retrieval can be affected by divided attention (Naveh-Benjamin et al. 1998). Recent research has focused on how attention is used to bind different components of an episode (e.g., an object and where it was encountered). Greene et al. (2021) found that an additional commitment of attention, beyond that needed to encode an item, is required to bind it to its sources during encoding; binding is not automatic. Recent research has also focused on the relationship between attention and the representational quality of episodic LTMs. Inspired by fuzzy-trace theory, which distinguishes between memory for surface form (or verbatim) details of an episode and memory for the meaning or gist of an episode (Brainerd & Reyna 1990, 2015), Greene & Naveh-Benjamin (2022b,c) and Greene et al. (2022) investigated whether attention during encoding is necessary to establish a specific or gist level of representation. Contrary to the prior consensus that attention was not needed for gist (Rabinowitz et al. 1982), divided attention at encoding impaired young adults' memory not only for episodic representations (e.g., "this old man was in this park") but also for gist-like representations (e.g., "this old man was in some type of nature scene"). Resources needed for gist may be less than for verbatim memory but above zero.

Both episodic and semantic memories can orient attention to features of the environment. For instance, a schematic semantic representation could guide attention to salient objects in an environment (Henderson et al. 1999). Alternatively, a specific episode may help direct attention to features of the environment or a specific goal. For example, if an individual has misplaced their keys, retrieving a memory of the last time they had their keys would be useful. Reinhart & Woodman (2014) showed that as a template to be searched for in a visual array became familiar, event-related potential evidence of the WM representation of the template subsided and was replaced by evidence of its retrieval from LTM. WM evidence returned on certain trials designated as high priority. Theeuwes et al. (2022) showed how statistical learning influences the direction of attention.

Attention and Implicit Long-Term Memory

Our model in **Figure 1** also includes a scenario in which a stimulus beyond conscious awareness (i.e., an unconscious prime) passes into aLTM but not into the FoA. Yet it elicits stored knowledge, which may then enter the FoA. This pathway illustrates how priming effects, thought to be implicit in nature and not available for conscious recollection, influence the relationship between attention and LTM.

Divided attention reduces, but does not eliminate, priming benefits (e.g., Keane et al. 2015). There may be a summation of two priming mechanisms: an unconscious, automatic mechanism at short delays and a conscious, attention-demanding mechanism that overrides semantic priming at longer delays. Neely (1977) elegantly demonstrated this dual mechanism by pitting semantic priming (e.g., *bird-robin*) against expectation-based priming (e.g., if the first word is a kind of furniture, expect the second word to be a kind of bird) and varying the time between the prime and target words. Semantic priming occurred quickly, whereas expectation, which should depend heavily on the control of attention, kicked in at longer intervals, overriding priming.

Computational Modeling of Attention and Long-Term Memory

We illustrated computational modeling in WM with different models that were compared for their adherence to the data. For LTM, we illustrate a different way to use computational modeling. In

Event-related potential: electroencephalographic recording synchronized to the stimulus onset, allowing averaging across similar stimuli for a stable neural response observation

Multinomial model: in computational modeling, a method of analyzing behavior into branching tree structures, each branch representing a choice point

Tetrahedral model: Jenkins's (1979) notion that behavior depends on conditions of encoding, retrieval, stimuli, and individual differences

this approach, one only constructs a single model that makes few theoretical assumptions on its own but incorporates parameters that help to indicate what is happening in the data. Different theoretical interpretations map onto different parameter values of the model. Greene & Naveh-Benjamin (2022b,c) used a multinomial model to examine the effects of attention on memory for verbatim and gist information. In this kind of model, a probability is attached to each potential outcome of each particular situation, resulting in a tree structure indicating possible paths of outcomes (see **Supplemental Figure S3**). On some trials (represented in one multinomial tree) there were intact probes, with the same pairing of a person and a scene in the probe as in the encoded material. On other trials (represented in separate trees), the pairing was changed between study and test (e.g., the same person paired with a park, but not the same park, changing the verbatim information but not the gist; or the same person paired with a city scene, changing both verbatim and gist information about the pairing). The parameters of the model are for the probability that the participant (*a*) has verbatim knowledge, (*b*) has only gist knowledge, and (*c*) has neither verbatim nor gist knowledge but still responds “old” on the basis of guessing. The model fit the data well. Moreover, when attention was divided, the probability of both verbatim knowledge and gist knowledge was reduced. Still, gist and implicit memory are less attention dependent than verbatim and explicit information.

INDIVIDUAL DIFFERENCES

Individual differences in attention and memory shed light on the mechanisms of normal functioning. They can influence learning readiness (e.g., the readiness of a young child to begin school), career aptitude, or even behaviors such as social distance compliance during a pandemic (Xie et al. 2020). Individual differences indicate the effects of configurations of processing abilities. If someone has poor attention, will that lead to a situation in which they often fail to encode the most relevant information into memory? Will they more often lose their attention to the goal of a task, forgetting it? Conversely, if someone has poor memory retrieval, will they be more likely to get lost during a movie or play because they cannot keep track of important sequences of events, becoming uninterested and unattentive? Are there separate groups of individuals with attention deficits but not memory deficits, and vice versa, or do these deficits coincide? These are interesting questions.

One tool for the analysis of individual differences in attention and memory is Jenkins's (1979) tetrahedral model, in which four broad factors are considered: encoding conditions (e.g., focused versus divided attention, or foreknowledge of the memory test), retrieval conditions (e.g., the need for familiarity versus declarative knowledge), the stimuli used (e.g., whether the items are emotionally salient), and subject factors (that is, individual differences). These all are relevant to an embedded processes approach to individual differences (see **Supplemental Figure S4**). The basic suggestion from this research is that individuals with better control of attention are the same ones who keep more information in WM and excel at problem solving and comprehension.

In the antisaccade task, a signal appears at one side of the participant and the required response is to move the eyes to the other side (as opposed to same-side looking in prosaccade control trials). The antisaccade task requires suppression of a natural tendency to look at the target. Unsworth et al. (2004) used versions of the task to determine how attention was involved for participants who had high versus low performance on one kind of complex span task (operation span). On a particular trial, participants see a series of math questions with a word to be remembered after each question (e.g., “Is $(9/3) - 1 = 1$? Dog”) and finally recall the words. Those with higher and lower span did not differ in eye movements in a block of prosaccade trials, whereas those with lower spans were slower and less accurate in antisaccade trial blocks. Another difference was the ability to keep the current goal in the FoA. When pro- and antisaccade trials were

Supplemental Material >

intermixed in the same trial block, forgetting the goal on the current trial was an added problem for low-span individuals. Unsworth et al. (2022) further found that individual variation in WM capacity and antisaccade performance depended on both the consistency and the intensity of attention.

In the Stroop task, a participant must resist reading a word aloud, a well-learned task, to quickly say aloud the color of the print (e.g., saying *blue* for the word *red* in blue print). Kane & Engle (2003) showed that low-span individuals were slower to name the color but did not make more errors than high-span individuals. However, if the task included many trials in which there was congruence between the word and print color (e.g., both of them were blue), then those with lower spans started making more errors on the incongruent trials. The explanation is once more that the task goal must be held in the FoA, whereas the prevalence of congruent trials causes those with lower spans to neglect the goal and start responding by relying on the words.

In the flanker task, a participant must identify the central letter in a string and ignore peripheral letters. Heitz & Engle (2007) used compatible strings SSSSS and HHHHH and incompatible strings SSHSS and HSHHH. By making most strings compatible, one can induce lower-span participants to lose the task goal. Individuals with lower spans made more errors than those with higher spans in their faster responses. At the slowest rate of responding, there was no difference between groups. Thus, relatively low-span individuals could still maintain or retrieve the task goal but needed to respond slowly to do so.

Conway et al. (2001) reexamined selective listening, in which people must repeat (shadow) a message in one ear while ignoring a message in the other ear. In this procedure, the participant's name occurs in the message presented in the ignored ear. There was a measure of shadowing errors just after the name occurred, and after the shadowing task questions were asked about whether anything unusual was heard. Interestingly, according to both measures, low-span individuals were much more likely to notice their names than were high-span individuals (for replications, see Naveh-Benjamin et al. 2014, Röer & Cowan 2021). The interpretation was that those with lower spans did not keep attention fixed on the shadowing task as well. Low-span individuals' attention sometimes took in the channel to be ignored when the name was presented. In further support of that interpretation, Colflesh & Conway (2007) found that when participants were to listen for an unusual event in the channel that was not shadowed, it was the high-span individuals, not those with lower spans, who noticed their names more often.

In all of these procedures, executive function seems related to WM capacity but it is unclear how WM is involved. In the case of intelligence, at least, not every type of executive function has the same impact, and WM seems to matter. Friedman et al. (2006) examined three executive functions: shifting of attention, inhibiting irrelevant information, and updating information. Of these, only updating directly involves WM, and it is the only one related to intelligence. Gray et al. (2017) used a battery of WM tasks with 9-year-olds and showed considerable relation between intelligence and tasks that were thought to index the FoA (i.e., visual spans and auditory running digit span, which are tasks that do not promote verbal rehearsal). Unlike Friedman et al. (2006), Gray and colleagues showed less relation between intelligence and tasks emphasizing the executive component of WM—*n*-back tasks, in which the participant receives a stream of items and for each one indicates whether it is the same as what occurred *n* items ago (e.g., a 3-back task), and number updating tasks, in which numbers have to be held in mind while one of them at a time is updated by adding or subtracting an amount indicated. However, Friedman and colleagues did not examine FoA tasks, and Gray and colleagues excluded shifting and inhibiting tasks from their predictive models because these did not cohere into a higher-level (latent) variable. The executive component of WM includes considerable variance that Gray and colleagues found to be shared with the FoA component, so executive function was more predictive of intelligence with the FoA factor omitted

from the predictive model. These studies taken together suggest special relevance to intelligence of both executive function and FoA in WM.

Life-span development: change in the brain and behavior as an individual progresses through infancy, childhood, young adulthood, and old age

LIFE-SPAN DEVELOPMENT

There are challenges to attention control and memory both in child development and in old age. Yet a comparison of these age groups is important. They differ tremendously in knowledge and experience, and the role of such differences might be elucidated by examining life-span development.

Infant and Child Development

Cowan (2016, 2022a) reviewed the transition from infancy to childhood and the progression from childhood to adulthood. Jean Piaget predominated in the field of cognitive development by setting out stages of cognitive organization and describing how concepts or schema develop across stages. Later work showed that infants with a more sensitive response mode (e.g., looking instead of grasping) acquired fundamental concepts like object permanence sooner than Piaget had thought. A general principle that developed to go beyond Piaget theoretically and account for the task dependence of results was termed neo-Piagetian theory. In it, the progression between conceptual stages depends on increases in the capabilities of WM and attention as children's brains mature. Cowan (2016, 2022a) discussed various indications that even after eliminating the effects of age differences in knowledge about the task materials that would facilitate recall, WM capacity increases steadily in childhood. One potential reason is an increase in the ability to carry out mnemonic strategies using executive functions (e.g., Elliott et al. 2021). For example, Camos & Barrouillet (2011) found that, whereas preschool children forgot more when the retention interval increased, older children (7-year-olds) were less susceptible to the passage of time and more susceptible to the difficulty of the activity taking up that time. It appeared as if only the older children adopted a mnemonic, attention-demanding strategy to counteract the loss of information over time, which they could not do during a distracting task.

One hypothesis considered by Cowan et al. (2018) was that the capacity to hold information in the FoA may increase with age, but the data did not support that interpretation. The hypothesis was investigated using a dual task involving memory for a visual source (arrays of colored spots), for an acoustic source (series of spoken digits in one experiment, tones in another), or for both modalities on the same trial. It was considered that attention should be responsible for holding some items of either modality, whereas some items might be held in a manner that does not depend on attention but is specific to the modality. **Figure 2** shows how the issue was investigated. The circle on the left represents the number of items that can be held in WM from the visual modality when only this modality is to be remembered, and the circle on the right shows the results when only the acoustic modality is to be remembered. The overlap between the circles represents the contribution of attention, which must be parceled out to the modalities in the bimodal attention situation. By estimating capacity at each age for unimodal and bimodal situations and subtracting bimodal from unimodal, it was possible to estimate the attention contribution to memory (overlapping areas in the figure). This attention contribution, or central capacity, was about one item and did not increase across the elementary school years or beyond into young adulthood. However, the modality-specific components increased strikingly with age. Cowan and colleagues suggested that older participants learn how to form patterns from meaningless collections of items so that the stimuli can be rapidly memorized without as much further commitment of attention. With age, participants may get better at being efficient with their attention, and there are many implications for educational practices (Cowan 2014).

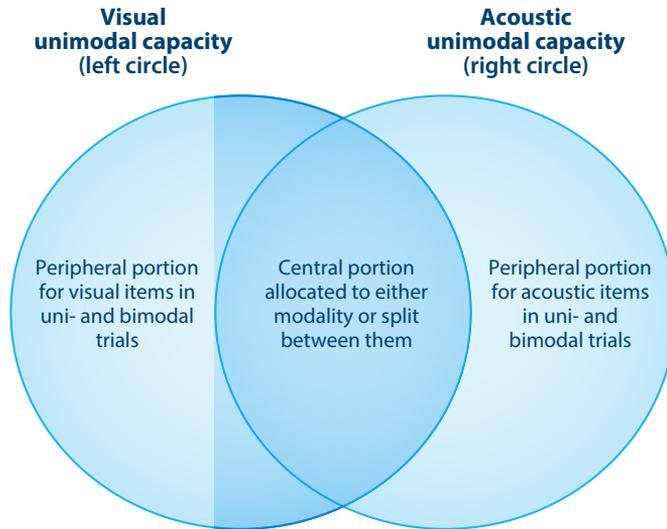


Figure 2

Capacity estimate model for working memory in a dual task. The central portion stays roughly constant at about one item, and the peripheral portions increase markedly during childhood (Cowan et al. 2018) and decrease in older adults (Greene et al. 2020).

As another example of the increasing efficiency with age, Cowan et al. (2021) found that children change from being reactive in their processing style in the early elementary school years to becoming more proactive. On each trial, participants were to remember a variable number of colored spots from an array. They sometimes were to carry out a brief but difficult task during the retention interval (pressing a button on the opposite side from a signal), and they were tested on recognition of an item from the array. The younger children tended to drop the array memory items when they had to carry out the difficult task, devoting all of their attention reactively to that immediate task at great expense to the subsequent color recognition judgment. In contrast, older children and adults showed an increasing tendency to try to maintain the colors that they would have to recognize, to the benefit of that task but at a modest detriment to the button-press task. They learned to maximize their performance overall by not merely reacting to each immediate task demand but proactively distributing attention to encompass all task demands. This proactive stance is useful to ensure that attention is applied when it is needed, as when one does one's homework in a timely manner rather than waiting to react to the imminence of the school day.

It remains unclear whether WM capacity growth with age in childhood is the cause of processing differences, the result of them, or both. Cowan et al. (2010) found evidence tentatively suggesting that capacity is more primary than processing. First- and second-grade children could deemphasize less relevant items in an array as well as adults could when the memory load was small (e.g., two triangles in colors to be tested on most trials, two circles in colors to be tested only occasionally), albeit at a lower span. However, these younger children could not allocate attention well when the memory load was larger (three more relevant and three less relevant items). In the latter case, the prioritization instructions no longer distinguished between trial types occurring on 20%, 50%, or 80% of trials in the block. The process of prioritizing the items may be limited when the same resource is needed to cope with more items to be stored.

There are implications for LTM as well. Forsberg et al. (2022a) utilized an array memory task with common objects, using immediate recognition of one object as a probe that the object was

or was not in the array. After the last immediate memory trial, they tested LTM for other objects that had populated those arrays. The proportion of items loaded into WM increased with age, and that proportion was a good predictor of how many items would be correctly recognized in LTM at each age.

Adult Aging

Cognitive abilities, including reasoning, memory, attention, and processing speed, decline gradually with age (Salthouse 2010, 2021). There are relatively little adult age-related performance changes in memory tasks that make minimal demands on attention, and vocabulary and general knowledge are often preserved (Baltes et al. 1999). There are more pronounced age-related declines in WM and episodic memory, in which more attention resources are involved (e.g., Greene et al. 2020).

Widely influential theories of cognitive aging attribute age-related deficits in WM and LTM to diminished attentional control (e.g., Craik 2020). We have looked at whether the decline with age in attention-related aspects of WM is comparable to what is seen in children. There is some evidence of important similarities, which we highlight by discussing two studies that used comparable methods. Recall that Cowan et al. (2018) used a set of acoustic items along with a set of visual items (colored spots) and found no developmental increase in the central, attention-based component that was shared between acoustic items and colors but found increases in the material-specific components instead. Greene et al. (2020) extended this result to adult aging. As with children, the aging pattern was one in which the developmental change was in the modality-specific components, this time declining with old age. The results from the child development study could have been attributed to the developmental increase in knowledge that might be applied to the stimuli, but the same cannot be said about aging effects. Instead, it appears that there is a biologically determined limit on the ability to use strategies to memorize the items, a limit that increases with child development and then decreases in old age.

Cowan et al. (2021) showed that children in their early elementary school years are reactive in their use of WM in a dual-task situation, whereas with child development they become more proactive. Van Gerven et al. (2016) used a procedure that could assess reactive and proactive processing across the life span between 5 and 97 years. On every trial, a cue indicated whether the participant would have to respond to a signal on the left or the right side or to a neutral cue that did not indicate which side. The informative cues were “anticues” that appeared on the side opposite to where the target would appear. After a preparatory interval that varied between 100 and 850 ms, a target appeared showing which of four fingers was to be used to respond. There was a tendency for the anticue to impair performance compared to the neutral condition, but with longer preparatory intervals, some could shift attention from the anticue to the side it indicated. In young children, behavior was governed by reactive control (responding reflexively in the direction of the anticue). Behavior shifted to proactive control (based on the anticue’s meaning) at progressively shorter preparatory intervals with maturation. In adults over 70 years old the pattern regressed to one most closely resembling that of children 9–12 years old, with more time needed for a proactive response than is found in young adults.

Age-related declines in LTM (e.g., Naveh-Benjamin & Old 2008) seem related to some of the same attentional mechanisms implicated in WM loss. Part of the relation between attention and LTM may stem from the attention-WM connection. Recall that Forsberg et al. (2022a) used arrays of common objects and found that an individual’s LTM for the array objects could be well predicted by that individual’s WM capacity for these objects across age in childhood. Forsberg et al. (2022b) found the same thing for adult aging; the LTM to WM ratio was the same across

adult age groups even though the capacity for both stores declined with age. This result is striking, given that measures of long-term episodic memory typically decline faster than measures of WM.

Not indexed by the procedure of Forsberg et al. (2022b), the most pronounced loss of LTM observed with age is for the precise context of the memory and its verbatim form (e.g., Greene & Naveh-Benjamin 2020, 2022a; Greene et al. 2022; Koutstaal 2003; Koutstaal & Schacter 1997), whereas the gist of past episodes is generally preserved (Greene & Naveh-Benjamin 2020, Brainerd & Reyna 2015). What is the basis of these effects in attention? Divided attention in young adults does not serve as an adequate model of aging effects, inasmuch as the selectivity of the deficit for associations seen in older adults is not mimicked in young adults under divided attention (e.g., Greene & Naveh-Benjamin 2020, 2022c). Some commitment of attention is needed to encode both items and their associations (Greene et al. 2021, Naveh-Benjamin et al. 2003), and older adults may have insufficient time and resources to encode some verbatim and associative information.

ATTENTION, MEMORY, AND THE BRAIN

The purpose of our inclusion of brain research is not simply to learn where in the brain a particular process occurs but also to provide convergent clues to understanding the cognitive processes. There are several reviews of relevant brain evidence (Cowan 1995, 2019; Kamiński & Rutishauser 2020; Postle & Oberauer 2022; Rose 2020). Ekman et al. (2016) found that individuals with high WM capacity had more densely connected lateral prefrontal cortex and posterior parietal cortex than their low-WM counterparts, consistent with the embedded processes approach (Cowan 2019). To avoid oversimplification, note that also connectivity through subcortical regions (thalamus and basal ganglia) was greater for higher-WM participants.

Summarizing across diverse brain evidence, we propose a schematic description of how memory and attention operate in the adult human brain, shown conceptually in **Figure 3a** and in terms of brain anatomy in **Figure 3b**. In this description, there are bottom-up and top-down directions of information flow between neural centers. In the flow of information, the intraparietal sulcus (IPS) plays a special role in indexing information in the FoA, presumably by functional connectivity to the relevant temporal and occipital regions of the posterior cortex in which the information is represented in aLTM (e.g., Li et al. 2014). The FoA is in turn controlled by frontal areas.

Kamiński & Rutishauser (2020) proposed that each component of WM within the embedded processes model corresponds to a different type of neural activity. Activity that is steady over time was said to reflect information in the FoA. Activity-silent representations, which may be based on temporarily heightened synaptic weights, were said to reflect information in aLTM. Dynamic activity that represents information differently at different points during retention was said to reflect executive function. We see considerable merit in this view, though, to avoid oversimplifying, note that Christophel et al. (2018) did find consistent activity for items that were not in the FoA in regions different from those in which activity was found for items in the FoA.

The brain evidence can address issues in the relation between attention and memory. Functional magnetic resonance imaging (fMRI) can use multivoxel pattern analysis (MVPA) to classify stimuli. This approach suggests that the classification represents items in the FoA that are currently needed, but not items needed later in the trial (Lewis-Peacock et al. 2012). A part of the IPS that responds to a WM load of either nonverbal visual or acoustic verbal stimuli (Cowan et al. 2011) is likely an FoA hub. In that area, MVPA may not distinguish between different types of stimuli, but it reflects the memory load regardless of the type of stimuli (Majerus et al. 2016). Moreover, that area is involved not only in preserving items in WM but also in distinguishing between similar items, such as three directions of movement presented in succession (Gosseries et al. 2018)

Prefrontal cortex: an area in the front of the brain that is essential for normal decision making and regulation of behavior

Intraparietal sulcus (IPS): a brain area thought to represent a hub of the focus of attention connected to sensory areas

Functional magnetic resonance imaging (fMRI): a technique identifying brain locations of increased neural activity within the most recent few seconds

Multivoxel pattern analysis (MVPA): a machine learning method for classifying the pattern of participants' brain activity to elucidate their thoughts

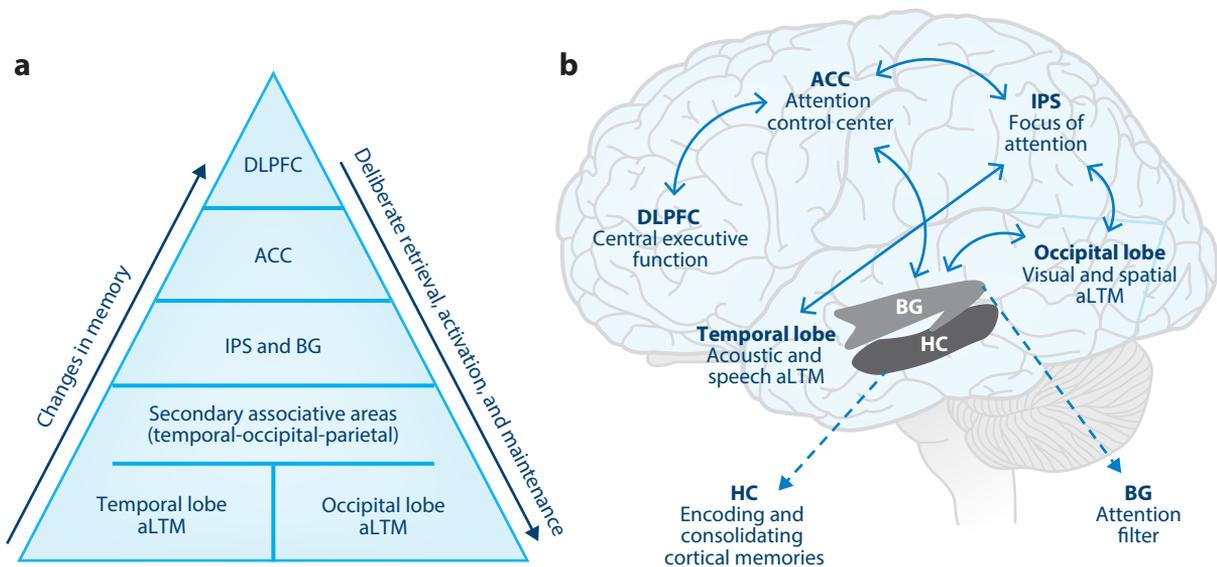


Figure 3

A simplified illustration of a theoretical neural framework consistent with the embedded processes model approach to attention and memory. This figure incorporates elements of former proposals (e.g., Chai et al. 2018; Cowan, 1995, 2019; Ekman et al. 2016; Postle & Oberauer 2022). (a) A schematic illustration of how attention relates to memory hierarchically, with a bottom-up and a top-down transfer of information along the same routes. (b) A brain map of this information flow. Solid bidirectional arrows depict the major neural routes of information transfer. The dorsolateral prefrontal cortex (DLPFC) is involved executive decisions; the anterior cingulate cortex (ACC) is involved in attention control; the intraparietal sulcus (IPS) serves as a hub of activity or focus of attention; the basal ganglia (BG) is a subcortical region involved in channeling attention; and the hippocampus (HC) is a key structure among subcortical regions involved in consolidating new explicit memories. Abbreviation: aLTM, activated long-term memory. The brain outline was constructed using images in the public domain.

or several bars at different orientations in an array (Cai et al. 2020). In these studies, the activity was less when the three stimuli presented on a trial were dissimilar to one another (e.g., a direction of movement intermixed with two colored objects).

There seems to be a trade-off between attention to WM and visual search in behavior and in the IPS (Panichello & Buschman 2021), presumably reflecting limits of the FoA. Majerus et al. (2018) showed that although the neural signatures of WM storage and processing differed, both impinged on activity in the IPS. Using methods sensitive to rapid changes in the brain (magneto- and electroencephalography), Palva et al. (2010) found that frontal-parietal synchrony increased with WM load, as it is expected to do if the executive processes direct the FoA (see **Figure 1**); but they also found that the IPS was a hub indicative of WM capacity (also related to consciousness; see the sidebar titled A Brain Region for the Focus of Attention).

Electroen-

cephalography: scalp recording of electrical potentials caused by neural activity, elucidating the nature and precise timing of brain function

Neuropsychological

condition: abnormal cognition or behavior resulting from defects in at least one brain area from malformation, injury, or disease

NEUROPSYCHOLOGICAL CONDITIONS

Many neuropsychological conditions shed light on the attention–memory relation. The close relation between disorders of attention and memory suggests that they come from related mechanisms (Moscovitch & Umiltà 1990). Consider, for example, research on the well-known densely amnesic patient H.M., who had much of the bilateral temporal lobe removed as protection against the effects of severe epilepsy (Scoville & Milner 1957). H.M. had deficits in the formation of new explicit memories but not of new implicit memories (e.g., he took less time to complete a puzzle with

A BRAIN REGION FOR THE FOCUS OF ATTENTION

Focused attention on perception and on items in working memory may share the elusive quality of conscious awareness. Although consciousness is difficult to study, one way to do so is using binocular rivalry. When the visual displays presented to the two eyes conflict rather than allowing fusion, one eye will predominate for a while in what is seen, suppressing the other image. Then the dominance will switch to the other eye, according to what participants report and what the measured brain activity shows. Zaretskaya et al. (2010) found that the verbal report of what image participants were aware of could be indexed by activity in the IPS, an area that others have found to be a hub of focused attention (e.g., Cai et al. 2020, Cowan et al. 2011, Majerus et al. 2016, Palva et al. 2010). Zaretskaya and colleagues further found that transcranial magnetic stimulation of the right IPS prolonged the period of stable perception before the experienced image switched. Putting the studies together suggests exciting ways in which the focus of attention in working memory could be empirically related to signs of consciousness.

repetition despite having no conscious recollection of the puzzle). Additionally, MacKay (2019) showed that this patient had difficulty assembling elements into new patterns for comprehension or production. For example, in one task, he was shown a picture of a man with two young boys and a stop light and was asked to talk about the situation using the words *first*, *cross*, and *before*. Whereas most adults respond with sentences like “When the light turns green, look first before you cross,” H.M. said “Before at first you cross across” (MacKay 2019, p. 26). Based on considerable evidence, MacKay concluded that binding elements together to construct a new representation is needed not only for learning new memories about events in context but also for aspects of language in which familiar phrases won’t do. These are attention-intensive aspects of forming new representations.

Attention may be needed to remove interference. Strikingly, many amnesic patients who usually retain nothing new in explicit memory are able to recall considerably more when interference is removed from the periods before and after learning (Dewar et al. 2010, McGhee et al. 2020), even after an unfilled retention interval of an hour (Cowan et al. 2004). Attention is also a factor in memory deficits from various types of dementia (e.g., Finke et al. 2013, Silveri et al. 2007).

Conversely, in attention deficit and hyperactivity disorder (ADHD), memory is also a factor (Alderson et al. 2013). A meta-analysis of adults with ADHD showed deficits on verbal, but not visual-figural, LTM (Skodzik et al. 2017). This result is the opposite of what would be expected if ADHD directly affected the use of attention in memory: Because verbal encoding can rely more on knowledge, visual encoding into memory typically depends more heavily on attention (Gray et al. 2017). However, ADHD might affect the use of executive function to carry out verbal mnemonic strategies. There was a similar finding for alcohol intoxication (Saults et al. 2007), which, unexpectedly, impaired performance on visual and auditory sequences but not on visual or auditory concurrent arrays, consistent with the notion that alcohol impaired strategies used to retain sequences. Subsequent research confirmed deficits in executive functions with alcohol intoxication (Bartholow et al. 2018, Cofresí et al. 2021).

In hemispatial neglect, patients fail to be aware of visual space on the side contralateral to their lesion (Parton et al. 2004), which is most typically in the right parietal lobe, leading to ignoring the left half of space. Individuals with this impairment experience disruption in memory and especially in memory for order, which may depend on spatial imagery (Antoine et al. 2019).

THEORETICAL IMPLICATIONS

Here we have assembled evidence from many subdisciplines on the relation between attention and memory. We have done so within the theoretical framework of the embedded processes approach

(e.g., Cowan 1988, 1999, 2019), which includes extensive-enough connections between attention and memory to be evaluated based on a broad range of evidence, and yet is general enough to be fine-tuned based on that evidence. Here we discuss it and how it is evolving and then compare it to several other approaches.

Support for Embedded Processes

We have shown support for aspects of the embedded processes approach, including (a) pervasive relations between attention and memory, (b) a distinction between function of executive processing and the FoA, (c) some generality of WM storage across modalities, and (d) also some modality- or code-specific storage that is presumed to be feature specific (e.g., tonal, tactile, taste, semantic, orthographic, and lexical features in aLTM). It is on the basis of this last point, and of the notion that a stimulus may activate multiple kinds of features, that a feature-based storage system seems to us preferable to a simpler taxonomy based on verbal and visual modules.

Evolution of Embedded Processes

The research is also useful to improve the embedded processes approach. Several new conclusions can be drawn. First, the neural model of the environment that serves as an attention filter does not seem to include semantic information except when it is attended. Thus, there is the finding that young adults who notice their name in an unattended acoustic channel tend to have low WM span (Conway et al. 2001, Naveh-Benjamin et al. 2014, Röer & Cowan 2021). Mind wandering (Kane et al. 2007) to the channel to be ignored might be the cause (see also Wood et al. 1997) in place of automatic semantic memory.

Second, spare time at least sometimes seems more useful in WM proactively rather than retroactively. In particular, ongoing mnemonic processing of the items already presented requires spare time, and, if enough time it is not available, processing of those items is carried out nevertheless, at the expense of the quality of encoding of any subsequent items (Kowialiewski et al. 2022, Ricker & Hardman 2017). This finding is at odds with the expectation from attentional refreshing or distraction removal accounts of a retroactive benefit to memory for items that occurred before the spare time.

Third, although the trace decay across seconds proposed by the refreshing account is observed clearly for items that are hard to categorize or are presented rapidly (Ricker & Vergauwe 2022, Ricker et al. 2020), no decay has been directly observed for lists of easily categorized items, such as those in verbal lists (Oberauer & Lewandowsky 2008).

Fourth, an asymmetry has been found in which shared, cross-modal attention typically has a greater effect on visual than on verbal retrieval (Morey & Mall 2012, Morey et al. 2013, Vergauwe et al. 2010). The same has been found for sets of nonverbal tones combined with colors (Li & Cowan 2021).

Fifth, it is only possible to account for WM based on a capacity-limited system such as the FoA if it is complemented by rapid learning of the material (Cowan et al. 2012). This rapid learning may make use of grouping of the stimuli to form new manageable chunks and patterns that achieve information compression (e.g., Brady & Tenenbaum 2013, Chekaf et al. 2016). For example, performance often benefits from the participant being able to choose grouping flexibly to match the pattern in the current list. The participant's grouping takes into account their WM span (Cowan & Elliott 2022). When there are multiple repetitions of items in a list, as in most serial numbers used for practical purposes, it is surprisingly advantageous for grouping not to be imposed on the list, so as to allow the participant to find a grouping that matches the structure of repetitions and other patterns (Cowan & Hardman 2021).

Sixth, whereas Cowan (1988, 2001) thought that the capacity limit of WM might reflect how much can be continually held in the FoA, it may be instead that the capacity limit is related to the fleeting use of attention to encode and consolidate the stimulus set in a memorable way to free up attention for other uses (Rhodes & Cowan 2018). This change in view may be needed to explain life-span evidence that what changes is the ability to offload information out of the FoA in a memorable form, with little change in the amount maintained in the FoA (Cowan et al. 2018, Greene et al. 2020).

Relation to Other Theoretical Approaches

The embedded processes approach was designed with an emphasis on the relation between attention and memory. Investigators have considerable intellectual and emotional investment in their own theoretical approaches (see Cowan et al. 2020, Watkins 1984). However, the approach taken here can complement and improve other approaches too, without abandoning them.

Baddeley model. The behavioral data do not distinguish very well between the more modular approach of Baddeley and colleagues and the more feature-based approach of the embedded processes model. Baddeley & Hitch (1974) included attention as storage in their model, if one reads carefully, but Baddeley (1986) removed it for the sake of parsimony. Baddeley (2000) added it back again in the form of the episodic buffer, with relations to attention still under investigation. When one finds double dissociations in which verbal material interferes more with other verbal material and visual material interferes more with other visual material, they can be accounted for by separate verbal and visual storage modules or, alternatively, by feature-specific interference in the embedded processes approach. The differences between approaches probably depend most on neuropsychological and brain investigations (e.g., Buchsbaum & D'Esposito 2019, Cowan et al. 2011, Li et al. 2014, Majerus et al. 2016, Morey 2018, Morey et al. 2020, Shallice & Papagno 2019, Yue et al. 2018). These have been variously interpreted to show specific modules from verbal and visuospatial processing or overlapping sets of features headed by general storage in the FoA. The result favoring modules in neuropsychological special cases warrants further research in which investigators of opposing views work together.

Modular views with no central executive component. Logie (2016) and Vandierendonck (2016) both claim support for models in which there is no central executive, but rather central executive function emerges from the ensemble of more specific processes. This approach runs against the notion of a general attention mechanism, and the claim is that interference between tasks occurs when very specific processes are in conflict between two tasks. The complexity of results on dual-task effects, which we have reviewed, keeps the two views alive (e.g., see Cowan et al. 2020). One key issue is whether a cognitive model will eventually be able to deal with the conscious impression one has of having a unified view of the world, which could stem from a global workspace notion of consciousness (Baars & Franklin 2003), in which the purpose of WM is to assemble relevant information to be used in thinking and decision making. Consciousness could be viewed as off-limits because the data are private for each of us, or it could be viewed as eligible for consideration on the basis of subjective reports. The claim would not be that people are aware of all cognitive processes going on in their brains, but rather that (a) there is a general attention function, (b) people are aware of the subset of processing that is going on within the FoA, and (c) they are capable of modifying that subset of processing. Although the central executive would be formed from various mechanisms that could be examined separately, a claim of the embedded processes approach is that attention affects any part of central executive functioning and that there is a trade-off between these functions competing for attention. This is an important avenue for further research.

Adaptive control of thought models. Anderson et al. (2004) described an evolving computational model of the mind, tied to brain regions that include modules that result in productions based on capacity-limited activation. There is an intention module with a goal buffer, a declarative memory module with a retrieval buffer, a processing module (involving the basal ganglia) leading to productions, and separate visual and manual modules. Other sensory modalities are not explicitly represented in this version of the model, but they presumably could be. This model seems consistent with the embedded processes approach except that activation in the latter is not capacity limited; capacity limits apply to only a subset of the activated information that is in the FoA. That is an interesting difference to explore in future work.

Time-based resource-sharing model. Barrouillet et al. (2011) and Barrouillet & Camos (2021) have summarized studies also discussed above in the section titled Effects of Attention on Working Memory, indicating that the way attention is used in WM is to refresh items one at a time to counteract decay. There is no contradiction with the embedded processes approach except perhaps that the distinction between individuals might not be in the rate at which items can be refreshed one at a time, but rather in the number of items that can be refreshed together (Gilchrist & Cowan 2011, Lemaire et al. 2018). Further work is needed also to determine whether the information is merely refreshed, which might not be necessary if the items do not actually decay over time (Oberauer & Lewandowsky 2008), or whether the critical process taking up attention is instead removal of distractors from an episodic record (Oberauer et al. 2012) or perhaps encoding of patterns that assist memorization of the items (Rhodes & Cowan 2018).

Interference-based models. Several models (e.g., Oberauer & Lin 2017, Oberauer et al. 2012) claim that there are two bases of capacity limits: a one-item FoA and limits due to the mutual interference between items. This model is not as antithetical to the embedded processes approach as it might appear, because the allowed interference is not entirely feature specific, and general interference between items could in effect result in what looks like a capacity limit (cf. Davelaar et al. 2005). More work is needed to understand the relation between general interference between items and capacity limit; whether these are compatible probably depends on the mathematical expression of interference and the test situation.

Signal detection models. Schurgin et al. (2020) advanced a model of performance in color reproduction tasks that treats the number of items in an array similarly to other factors that influence performance. They find a psychophysical function that depends on the discriminability between items in a continuous manner, according to a signal detection model in which the number of items is only one factor that can alter discriminability. Although this model appears to have nothing to do with capacity limits, it could well be that the performance function across the number of items reflects the difficulty of a hub of attention, such as the IPS, in keeping track of multiple items and their relations to one another (Gosseries et al. 2018). Thus, although this contribution is a major one, whether it replaces or complements a capacity approach is still an open question.

Relation of embedded processes to alternative approaches. The embedded processes approach was designed with the relation between attention and memory in the foreground. However, we would suggest that the approach taken here can be used to complement and improve most of the other approaches without abandoning them, similar to how we have fine-tuned the embedded processes approach in this article. This model is in contrast with the highly modular models in which there is sometimes an effort to account for memory and behavior with little, if any, involvement of attention. However, the largely modular approach of Baddeley & Hitch (1974) and Baddeley (2000) straddles these two extremes by placing considerable stock in both modality-specific and attention-based, general processes (for recent work on the latter, see Allen &

Ueno 2018, Hu et al. 2016). The adaptive control of thought shares with embedded processes the important role of activation and attention. The adaptive control approach excels in offering a set of equations to broadly situate attention and other processing within memory in a computational model, which is a very useful endeavor, but it may require quite a long time for the various strands of research integrated here to be taken into consideration within such a model. The approaches involving what happens as a function of time or interference are meant to deal with specific circumstances in memory for lists and arrays (e.g., Barrouillet & Camos 2021, Oberauer & Lin 2017), but these can be tried out within an embedded processes framework. Signal detection models can offer elegant fits to the data but can still be complemented by investigations of what mechanisms underlie limits related to the number of items that can be held in attention, the nature of attention to the items during encoding and retrieval of LTM, and so on. To encourage this endeavor, we have strived to make our thinking accessible without imposing rigid modeling assumptions.

CONCLUSION

No matter one's theoretical view, it seems clear that there is a rich body of convergent and complementary evidence about the relation between attention and memory in the fields of behavior, computational modeling, individual and developmental differences, brain science, and neuropathology. Cross-fertilization between these fields is not an easy matter, but recent work shows some multidisciplinary convergence. This development reflects our optimism about the current directions of this vast field.

SUMMARY POINTS

- The relation between attention and memory is important for both psychological theory and practical issues (e.g., education, job performance, eyewitness testimony).
- The role of attention differs between information that is or is not consciously memorable.
- Both working- and long-term memory include attention-dependent and attention-independent processes.
- Formation of habits, procedures, routines, and gist require less attention than conscious, verbatim memories.
- Memory guides the direction of attention toward more important stimuli through both voluntary executive processes and involuntary orienting processes.
- Brain and behavioral evidence both point to several working memory limits: the capacity of the focus of attention, persistence of activation of information outside of that focus, and interference between active items.
- With childhood development, there is an increase in the ability to form patterns to remember while sparing the focus of attention and to allocate attention proactively; with old age, these abilities decline somewhat.
- Findings from diverse fields including individual differences, development, neuropsychology, neuroimaging, and computational modeling provide convergent information about the attention–memory relation.

FUTURE ISSUES

- When an attended item or event is held in working memory but is not later retrieved from long-term memory, can this situation reflect an absence of long-term storage, or does it always reflect some other reason for retrieval failure?
- Can we identify the types of attentive processing that prevent the decay of unattended representations across several seconds, such as categorization of an item?
- Is there a trade-off between the limit in how many separate items can be attended to during perception and how many can be retained in the focus of attention?
- Although children and older adults both have poorer attention control than young adults, to what extent are attention and memory protected in older adults because of a lifetime of knowledge?
- Do modality differences reflect separate working memory mechanisms, or do they indicate a general attention-related capacity limit combined with effects of feature similarity?
- When does similarity between items to attend to or remember make a big difference and when does it not matter, given that both results have been obtained?
- Is activated long-term memory outside of the focus of attention represented by neural activity or some other mechanism, such as altered synaptic connection weights?
- Is there a special role of attention for associations and order, beyond its role for remembering items?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work received support from NIH grant R01-HD021338 to N.C.

LITERATURE CITED

- Alderson RM, Kasper LJ, Hudec KL, Patros CH. 2013. Attention-deficit/hyperactivity disorder (ADHD) and working memory in adults: a meta-analytic review. *Neuropsychology* 27(3):287–302
- Allen RJ, Hitch GJ, Mate J, Baddeley AD. 2012. Feature binding and attention in working memory: a resolution of previous contradictory findings. *Q. J. Exp. Psychol.* 65(12):2369–83
- Allen RJ, Ueno T. 2018. Multiple high-reward items can be prioritized in working memory but with greater vulnerability to interference. *Atten. Percept. Psychophys.* 80(7):1731–43
- Anderson JR, Bothell D, Byrne MD, Douglass S, Lebiere C, Qin Y. 2004. An integrated theory of the mind. *Psychol. Rev.* 111(4):1036–60
- Antoine S, Ranzini M, van Dijck J-P, Slama H, Bonato M, et al. 2019. Hemispatial neglect and serial order in verbal working memory. *J. Neuropsychol.* 13(2):272–88
- Arrington CN, Kulesc PA, Francis DJ, Fletcher JM, Barnes MA. 2014. The contribution of attentional control and working memory to reading comprehension and decoding. *Sci. Stud. Read.* 18(5):325–46
- Baars BJ, Franklin S. 2003. How conscious experience and working memory interact. *Trends Cogn. Sci.* 7:166–72
- Baddeley AD. 1986. *Working Memory*. Oxford, UK: Clarendon Press

- Baddeley AD. 2000. The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4(11):417–23
- Baddeley AD, Hitch G. 1974. Working memory. In *The Psychology of Learning and Motivation*, Vol. 8, ed. GH Bower, pp. 47–89. New York: Academic
- Balota DA. 1983. Automatic semantic activation and episodic memory encoding. *J. Verbal Learn. Verbal Behav.* 22(1):88–104
- Baltes PB, Staudinger UM, Lindenberger U. 1999. Lifespan psychology: theory and application to intellectual functioning. *Annu. Rev. Psychol.* 50:471–507
- Barrouillet P, Camos V. 2021. The time-based resource-sharing model of working memory. In *Working Memory: State of the Science*, ed. RH Logie, V Camos, N Cowan, pp. 85–115. Oxford, UK: Oxford Univ. Press
- Barrouillet P, Portrat S, Camos V. 2011. On the law relating processing to storage in working memory. *Psychol. Rev.* 118(2):175–92
- Bartholow BD, Fleming KA, Wood PK, Cowan N, Saults JS, et al. 2018. Alcohol effects on response inhibition: variability across tasks and individuals. *Exp. Clin. Psychopharmacol.* 26(3):251–67
- Brady TF, Tenenbaum JB. 2013. A probabilistic model of visual working memory: incorporating higher order regularities into working memory capacity estimates. *Psychol. Rev.* 120(1):85–109
- Brainerd CJ, Reyna VF. 1990. Gist is the grist: fuzzy-trace theory and the new intuitionism. *Dev. Rev.* 10(1):3–47
- Brainerd CJ, Reyna VF. 2015. Fuzzy-trace theory and lifespan cognitive development. *Dev. Rev.* 38(1):89–121
- Broadbent DE. 1958. *Perception and Communication*. New York: Pergamon Press
- Bryan M. 2009. The psycho-phone. *Antique Phonograph News*, July–Aug. <https://web.archive.org/web/20101130103402/http://capsnews.org/apn2009-4.htm>
- Buchsbaum BR, D’Esposito M. 2019. A sensorimotor view of verbal working memory. *Cortex* 112:134–48
- Cai Y, Yu Q, Sheldon AD, Postle BR. 2020. The role of location-context binding in nonspatial visual working memory. *eNeuro* 7(6):ENEURO.0430-0420.2020
- Camos V, Barrouillet P. 2011. Developmental change in working memory strategies: from passive maintenance to active refreshing. *Dev. Psychol.* 47(3):898–904
- Chai WJ, Abd Haid AI, Abdullah JM. 2018. Working memory from the psychological and neurosciences perspectives: a review. *Front. Psychol.* 9:401
- Chambers R, Lo CY, Allen NB. 2008. The impact of intensive mindfulness training on attentional control, cognitive style, and affect. *Cogn. Ther. Res.* 32:303–22
- Chekaf M, Cowan N, Mathy F. 2016. Chunk formation in immediate memory and how it relates to data compression. *Cognition* 155:96–107
- Chen H, Wyble B. 2016. Attribute amnesia reflects a lack of memory consolidation for attended information. *J. Exp. Psychol. Hum. Percept. Perform.* 42(2):225–34
- Christophel TB, Iamschchinina P, Yan C, Allefeld C, Haynes J-D. 2018. Cortical specialization for attended versus unattended working memory. *Nat. Neurosci.* 21(4):494–96
- Chun MM, Turk-Browne NB. 2007. Interactions between attention and memory. *Curr. Opin. Neurobiol.* 17(2):177–84
- Cofresi RU, Watts AL, Martins JS, Wood PK, Sher KJ, et al. 2021. Acute effect of alcohol on working memory updating. *Addiction* 116(11):3029–43
- Colflesh GJH, Conway ARA. 2007. Individual differences in working memory capacity and divided attention in dichotic listening. *Psychon. Bull. Rev.* 14:699–703
- Conway ARA, Cowan N, Bunting MF. 2001. The cocktail party phenomenon revisited: the importance of working memory capacity. *Psychon. Bull. Rev.* 8:331–35
- Conway ARA, Kane MJ, Engle RW. 2003. Working memory capacity and its relation to general intelligence. *Trends Cogn. Sci.* 7(12):547–52
- Cowan N. 1988. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychol. Bull.* 104(2):163–91
- Cowan N. 1995. *Attention and Memory: An Integrated Framework*. New York: Oxford Univ. Press
- Cowan N. 1999. An embedded-processes model of working memory. In *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, ed. A Miyake, P. Shah, pp. 62–101. Cambridge, UK: Cambridge Univ. Press

Provides a gateway to understanding cognitive and brain mechanisms associated with the embedded processes model.

Explains working memory development in the context of cognitive developmental theories based on attention based skills.

- Cowan N. 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24(1):87–185
- Cowan N. 2014. Working memory underpins cognitive development, learning, and education. *Educ. Psychol. Rev.* 26(2):197–223
- Cowan N. 2016. Working memory maturation: Can we get at the essence of cognitive growth? *Perspect. Psychol. Sci.* 11(2):239–64
- Cowan N. 2017. The many faces of working memory and short-term storage. *Psychon. Bull. Rev.* 24:1158–70
- Cowan N. 2019. Short-term memory based on activated long-term memory: a review in response to Norris 2017. *Psychol. Bull.* 145(8):822–47**
- Cowan N. 2022a. Working memory development: a 50-year assessment of research and underlying theories. *Cognition* 224:105075**
- Cowan N. 2022b. Item-position binding capacity limits and word limits in working memory: a reanalysis of Oberauer 2019. *J. Cogn.* 5(1):3
- Cowan N, AuBuchon AM, Gilchrist AL, Blume CL, Boone AP, Saults JS. 2021. Developmental change in the nature of attention allocation in a dual task. *Dev. Psychol.* 57(1):33–46
- Cowan N, Belletier C, Doherty JM, Jaroslawska AJ, Rhodes S, et al. 2020. How do scientific views change? Notes from an extended adversarial collaboration. *Perspect. Psychol. Sci.* 15(4):1011–25
- Cowan N, Beschin N, Della Sala S. 2004. Verbal recall in amnesiacs under conditions of diminished retroactive interference. *Brain* 127:825–34
- Cowan N, Blume CL, Saults JS. 2013. Attention to attributes and objects in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 39(3):731–47
- Cowan N, Elliott EM. 2022. Deconfounding serial recall: response timing and the overarching role of grouping. *J. Exp. Psychol. Learn. Mem. Cogn.* 49(2):249–68
- Cowan N, Elliott EM, Saults JS, Morey CC, Mattox S, Hismjatullina A, Conway AR. 2005. On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cogn. Psychol.* 51(1):42–100
- Cowan N, Fristoe NM, Elliott EM, Brunner RP, Saults JS. 2006. Scope of attention, control of attention, and intelligence in children and adults. *Mem. Cogn.* 34:1754–68
- Cowan N, Hardman KO. 2021. Immediate recall of grouped serial numbers with or without multiple item repetitions. *Memory* 21:744–61
- Cowan N, Li D, Moffitt A, Becker TM, Martin EA, et al. 2011. A neural region of abstract working memory. *J. Cogn. Neurosci.* 23:2852–63
- Cowan N, Li Y, Glass B, Saults JS. 2018. Development of the ability to combine visual and acoustic information in working memory. *Dev. Sci.* 21(5):e12635
- Cowan N, Lichty W, Grove TR. 1990. Properties of memory for unattended spoken syllables. *J. Exp. Psychol. Learn. Mem. Cogn.* 16(2):258–69
- Cowan N, Morey CC. 2007. How can dual-task working memory retention limits be investigated? *Psychol. Sci.* 18:686–88
- Cowan N, Morey CC, AuBuchon AM, Zwilling CE, Gilchrist AL. 2010. Seven-year-olds allocate attention like adults unless working memory is overloaded. *Dev. Sci.* 13(1):120–33
- Cowan N, Rachev NR. 2018. Merging with the path not taken: Wilhelm Wundt's work as a precursor to the embedded-processes approach to memory, attention, and consciousness. *Conscious. Cogn.* 63:228–38
- Cowan N, Rouder JN, Blume CL, Saults JS. 2012. Models of verbal working memory capacity: What does it take to make them work? *Psychol. Rev.* 119(3):480–99
- Cowan N, Saults JS, Blume CL. 2014. Central and peripheral components of working memory storage. *J. Exp. Psychol. Gen.* 143(5):1806–36
- Craik FIM. 2020. Remembering: an activity of mind and brain. *Annu. Rev. Psychol.* 71:1–24
- Craik FIM, Eftekhari E, Binns MA. 2018. Effects of divided attention at encoding and retrieval: further data. *Mem. Cogn.* 46(8):1263–77
- Daneman M, Carpenter PA. 1980. Individual differences in working memory and reading. *J. Verbal. Learn. Verbal Behav.* 19:450–66
- Darwin CJ, Turvey MT, Crowder RG. 1972. An auditory analogue of the Sperling partial report procedure: evidence for brief auditory storage. *Cogn. Psychol.* 3:255–67

- Davelaar EJ, Goshen-Gottstein Y, Ashkenazi A, Haarman HJ, Usher M. 2005. The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychol. Rev.* 112:3–42
- Davies DR, Parasuraman R. 1982. *The Psychology of Vigilance*. New York: Academic
- Demetriou A, Spanoudis G, Shayer M, Van der Ven S, Brydges CR, et al. 2014. Relations between speed, working memory, and intelligence from preschool to adulthood: structural equation modeling of 14 studies. *Intelligence* 46:107–21
- Dew ITZ, Cabeza R. 2011. The porous boundaries between explicit and implicit memory: behavioral and neural evidence. *Ann. N.Y. Acad. Sci.* 1224:174–90
- Dewar M, Della Sala S, Beschin N, Cowan N. 2010. Profound retroactive interference in anterograde amnesia: What interferes? *Neuropsychology* 24:357–67
- Draheim C, Pak R, Draheim AA, Engle RW. 2022. The role of attention control in complex real-world tasks. *Psychon. Bull. Rev.* 15:1–55
- Ebbinghaus H. 1885 (1913). *Memory: A Contribution to Experimental Psychology*, transl. HA Ruger, CE Bussenius. New York: Teach. Coll. Columbia Univ.
- Eich E. 1984. Memory for unattended events: remembering with and without awareness. *Mem. Cogn.* 12:105–11
- Ekman M, Fiebach CJ, Melzer C, Tittgemeyer M, Derrfuss J. 2016. Different roles of direct and indirect frontoparietal pathways for individual working memory capacity. *J. Neurosci.* 36(10):2894–903
- Elliott EM et al. 2021. Multi-lab direct replication of Flavell, Beach and Chinsky 1966: spontaneous verbal rehearsal in a memory task as a function of age. *Adv. Methods Pract. Psychol. Sci.* 4(2):1–20
- Elliott EM, Cowan N. 2001. Habituation to auditory distractors in a cross-modal, color-word interference task. *J. Exp. Psychol. Learn. Mem. Cogn.* 27:654–67
- Engle RW. 2002. Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11:19–23
- Finke K, Myers N, Bublak P, Sorg C. 2013. A biased competition account of attention and memory in Alzheimer's disease. *Philos. Trans. R. Soc. B* 368(1628):20130062
- Forsberg A, Blume C, Cowan N. 2021b. The development of metacognitive accuracy in working memory across childhood. *Dev. Psychol.* 57:1297–317
- Forsberg A, Guitard D, Adams EJ, Pattanakul D, Cowan N. 2022a. Children's long-term retention is directly constrained by their working memory capacity limitations. *Dev. Sci.* 25(2):e13164
- Forsberg A, Guitard D, Cowan N. 2021a. Working memory limits severely constrain long-term retention. *Psychon. Bull. Rev.* 28:537–47
- Forsberg A, Guitard D, Greene NR, Naveh-Benjamin M, Cowan N. 2022b. The proportion of working memory items recoverable from long-term memory remains fixed despite adult aging. *Psychol. Aging* 37(7):777–86
- Friedman NP, Miyake A, Corley RP, Young SE, DeFries JC, Hewitt JK. 2006. Not all executive functions are related to intelligence. *Psychol. Sci.* 17:172–79
- Fukuda K, Vogel EK. 2019. Visual short-term memory capacity predicts the “bandwidth” of visual-long term memory encoding. *Mem. Cogn.* 47:1481–97
- Gagnon S, Foster J, Turcotte J, Jongenelis S. 2004. Involvement of the hippocampus in implicit learning of supra-span sequences: the case of SJ. *Cogn. Neuropsychol.* 21:867–82
- Gernsback H. 1911 (2014). *Ralph 124C 41+*. Mansfield, CT: Martino Fine Books
- Gershkoff-Stowe L. 2001. The course of children's naming errors in early word learning. *J. Cogn. Dev.* 2:131–55
- Gilchrist AL, Cowan N. 2011. Can the focus of attention accommodate multiple separate items? *J. Exp. Psychol. Learn. Mem. Cogn.* 37(6):1484–502
- Gosseries O, Yu Q, LaRocque JJ, Starrett MJ, Rose NS, et al. 2018. Parietal-occipital interactions underlying control- and representation-related processes in working memory for nonspatial visual features. *J. Neurosci.* 38:4357–66
- Gray S, Green S, Alt M, Hogan T, Kuo T, et al. 2017. The structure of working memory in young school-age children and its relation to intelligence. *J. Mem. Lang.* 92:183–201
- Gray S, Levy R, Alt M, Hogan TP, Cowan N. 2022. Working memory predicts new word learning over and above existing vocabulary and nonverbal IQ. *J. Speech Lang. Hear. Res.* 65:1044–69

Demonstrates that dividing attention in young adults cannot fully simulate older adults' associative memory deficit.

Demonstrates that attention to different features of lists alters readiness for item versus order tests.

Maps three different types of neural activity during working memory onto the embedded-processes model components.

- Greene NR, Chism S, Naveh-Benjamin M. 2022. Levels of specificity in episodic memory: insights from response accuracy and subjective confidence ratings in older adults and in younger adults under full or divided attention. *J. Exp. Psychol. Gen.* 151(4):804–19
- Greene NR, Martin BA, Naveh-Benjamin M. 2021. The effects of divided attention at encoding and at retrieval on multidimensional source memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 47(11):1870–87
- Greene NR, Naveh-Benjamin M. 2020. A specificity principle of memory: evidence from aging and associative memory. *Psychol. Sci.* 31(3):316–31
- Greene NR, Naveh-Benjamin M. 2022a. Adult age differences in specific and gist associative episodic memory across short- and long-term retention intervals. *Psychol. Aging* 37(6):681–97
- Greene NR, Naveh-Benjamin M. 2022b. Effects of divided attention at encoding on specific and gist representations in working and long-term memory. *J. Mem. Lang.* 126:104340
- Greene NR, Naveh-Benjamin M. 2022c. The effects of divided attention at encoding on specific and gist-based associative episodic memory. *Mem. Cogn.* 50(1):59–76
- Greene NR, Naveh-Benjamin M, Cowan N. 2020. Adult age differences in working memory capacity: spared central storage but deficits in ability to maximize peripheral storage. *Psychol. Aging* 35(6):866–80
- Guérard K, Saint-Aubin J, Boucher P, Tremblay S. 2011. The role of awareness in anticipation and recall performance in the Hebb repetition paradigm: implications for sequence learning. *Mem. Cogn.* 39:1012–22
- Guitard D, Cowan N. 2023a. Attention allocation between item and order information in short-term memory. *Q. J. Exp. Psychol.* 76(6):1391–409
- Guitard D, Cowan N. 2023b. The tradeoff between item and order information in short-term memory does not depend on encoding time. *J. Exp. Psychol. Hum. Percept. Perform.* 49(1):51–70**
- Guitard D, Saint-Aubin J, Cowan N. 2021. Asymmetrical interference between item and order information in short-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 47:243–63
- Guitard D, Saint-Aubin J, Cowan N. 2022. Tradeoffs between item and order information in short-term memory. *J. Mem. Lang.* 122:104300
- Halford GS, Cowan N, Andrews G. 2007. Separating cognitive capacity from knowledge: a new hypothesis. *Trends Cogn. Sci.* 11:236–42
- Halpern DF. 1998. Teaching critical thinking for transfer across domains: dispositions, skills, structure training, and metacognitive monitoring. *Am. Psychol.* 53:449–55
- Hardman KO, Cowan N. 2015. Remembering complex objects in visual working memory: Do capacity limits restrict objects or features? *J. Exp. Psychol. Learn. Mem. Cogn.* 41(2):325–47
- Hebb DO. 1961. Distinctive features of learning in the higher animal. In *Brain Mechanisms and Learning: A Symposium*, ed. JF Delafresnaye, pp. 37–51. Oxford, UK: Blackwell
- Heitz RP, Engle RW. 2007. Focusing the spotlight: individual differences in visual attention control. *J. Exp. Psychol. Gen.* 136(2):217–40
- Henderson JM, Weeks PA Jr., Hollingworth A. 1999. The effects of semantic consistency on eye movements during complex scene viewing. *J. Exp. Psychol. Hum. Percept. Perform.* 25(1):210–28
- Hitch GJ, Hu Y, Allen RJ, Baddeley AD. 2018. Competition for the focus of attention in visual working memory: perceptual recency versus executive control. *Ann. N.Y. Acad. Sci.* 1424:64–75
- Hu Y, Allen RJ, Baddeley AD, Hitch GJ. 2016. Executive control of stimulus-driven and goal-directed attention in visual working memory. *Atten. Percept. Psychophys.* 78:2164–75
- Jacoby LL, Woloshyn V, Kelley C. 1989. Becoming famous without being recognized: unconscious influences of memory produced by dividing attention. *J. Exp. Psychol. Gen.* 118:115–25
- James W. 1892 (1985). *Psychology: The Briefer Course*. Cambridge, MA: Harvard Univ. Press
- Jaroslawska AJ, Gathercole SE, Logie MR, Holmes J. 2016. Following instructions in a virtual school: Does working memory play a role? *Mem. Cogn.* 44:580–89
- Jenkins JJ. 1979. Four points to remember: a tetrahedral model of memory experiments. In *Levels of Processing in Human Memory*, ed. LS Cermack, FIM Craik, pp. 429–46. Hillsdale, NJ: Erlbaum
- Jiang Q, Cowan N. 2020. Incidental learning of list membership is affected by serial position in the list. *Memory* 28:669–76
- Kamiński J, Rutishauser U. 2020. Between persistently active and activity-silent frameworks: novel vistas on the cellular basis of working memory. *Ann. N.Y. Acad. Sci.* 1464:64–75**

- Kane MJ, Brown LH, McVay JC, Silvia PJ, Myin-Germeys I, Kwapil TR. 2007. For whom the mind wanders, and when: an experience-sampling study of working memory and executive control in daily life. *Psychol. Sci.* 18(7):614–21
- Kane MJ, Engle RW. 2003. Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *J. Exp. Psychol. Gen.* 132(1):47–70
- Kane MJ, Hambrick DZ, Tuholski SW, Wilhelm O, Payne TW, Engle RE. 2004. The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *J. Exp. Psychol. Gen.* 133:189–217
- Keane MM, Cruz ME, Verfaelli M. 2015. Attention and implicit memory: Priming-induced benefits and costs have distinct attentional requirements. *Mem. Cogn.* 43(2):216–25
- Keppel G, Underwood BJ. 1962. Proactive inhibition in short-term retention of single items. *J. Verbal Learn. Verbal Behav.* 1:153–61
- Koutstaal W. 2003. Older adults encode—but do not always use—perceptual details: intentional versus unintentional effects of detail on memory judgments. *Psychol. Sci.* 14(2):189–93
- Koutstaal W, Schacter DL. 1997. Gist-based false recognition of pictures in older and younger adults. *J. Mem. Lang.* 37(4):555–83
- Kowialiewski B, Lemaire B, Portrat S. 2022. Between-item similarity frees up working memory resources through compression: a domain-general property. *J. Exp. Psychol. Gen.* 151(11):2641–65
- Lemaire B, Pageot A, Plancher G, Portrat S. 2018. What is the time course of working memory attentional refreshing? *Psychon. Bull. Rev.* 25(1):370–85
- Lemaire B, Portrat S. 2018. A computational model of working memory integrating time-based decay and interference. *Front. Psychol.* 9:416
- Lepsien J, Thornton I, Nobre AC. 2011. Modulation of working-memory maintenance by directed attention. *Neuropsychologia* 49:1569–77
- Levett LM, Haigh CB, Perez G. 2021. Toward a broader framework of eyewitness identification behavior. *J. Appl. Res. Mem. Cogn.* 10(3):341–45
- Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. 2012. Neural evidence for a distinction between short-term memory and the focus of attention. *J. Cogn. Neurosci.* 24:61–79
- Li D, Christ SE, Cowan N. 2014. Domain-general and domain-specific functional networks in working memory. *NeuroImage* 102:646–56
- Li Y, Cowan N. 2021. Attention effects in working memory that are asymmetric across sensory modalities. *Mem. Cogn.* 49(5):1050–65
- Logie RH. 2016. Retiring the central executive. *Q. J. Exp. Psychol.* 69:2093–109
- MacKay DG. 2019. The earthquake that reshaped the intellectual landscape of memory, mind and brain: case HM. In *Cases of Amnesia: Contributions to Understanding Memory and the Brain*, ed. SE MacPherson, S Della Sala, pp. 16–39. London: Routledge**
- Madore KP, Wagner AD. 2022. Readiness to remember: predicting variability in episodic memory. *Trends Cogn. Sci.* 26(8):707–23
- Majerus S, Cowan N, Péters F, Van Calster L, Phillips C, Schrouff J. 2016. Cross-modal decoding of neural patterns associated with working memory: evidence for attention-based accounts of working memory. *Cereb. Cortex* 26:166–79
- Majerus S, Péters F, Bouffier M, Cowan N, Phillips C. 2018. The dorsal attention network reflects both encoding load and top-down control during working memory. *J. Cogn. Neurosci.* 30:144–59
- Marcel AJ. 1983. Conscious and unconscious perception: experiments on visual masking and word recognition. *Cogn. Psychol.* 15:197–237
- Matthews G, Wohleber RW, Lin J. 2020. Stress, skilled performance, and expertise: overload and beyond. In *The Oxford Handbook of Expertise*, ed. P Ward, JM Schraagen, J Gore, E Roth, pp. 490–524. Oxford, UK: Oxford Univ. Press
- McGhee JD, Cowan N, Beschin N, Mosconi C, Della Sala S. 2020. Wakeful rest benefits before and after encoding in anterograde amnesia. *Neuropsychology* 34(5):524–34
- McKone E, Dennis C. 2000. Short-term implicit memory: visual, auditory, and cross-modality priming. *Psychon. Bull. Rev.* 7:341–46

Provides new insight into the well-known case of amnesia, shedding light on the attention–memory relation.

- Mizrak E, Oberauer K. 2021. What is time good for in working memory? *Psychol. Sci.* 32(8):1325–37
- Morey C. 2018. The case against specialized visual-spatial short-term memory. *Psychol. Bull.* 144:849–83
- Morey CC, Cowan N, Morey RD, Rouders JN. 2011. Flexible attention allocation to visual and auditory working memory tasks: Manipulating reward induces a tradeoff. *Atten. Percept. Psychophys.* 73:458–72
- Morey CC, Mall JT. 2012. Cross-domain costs during concurrent verbal and spatial serial memory tasks are asymmetric. *Q. J. Exp. Psychol.* 65:1777–97
- Morey CC, Morey RD, van der Reijden M, Holweg M. 2013. Asymmetric cross-domain interference between two working memory tasks: implications for models of working memory. *J. Mem. Lang.* 69:324–48
- Morey CC, Rhodes S, Cowan N. 2020. Co-existing, contradictory working memory models are ready for progressive refinement: reply to Logie. *Cortex* 123:200–2
- Moscovitch M, Umiltà C. 1990. Modularity and neuropsychology: modules and central processes in attention and memory. In *Modular Deficits in Alzheimer-Type Dementia*, ed. MF Schwartz, pp. 1–59. Cambridge, MA: MIT Press
- Naveh-Benjamin M, Craik FIM, Guez J, Dori H. 1998. Effects of divided attention on encoding and retrieval processes in human memory: further support for an asymmetry. *J. Exp. Psychol. Learn. Mem. Cogn.* 24:1091–104
- Naveh-Benjamin M, Guez J, Marom M. 2003. The effects of divided attention at encoding on item and associative memory. *Mem. Cogn.* 31:1021–35
- Naveh-Benjamin M, Kilb A, Maddox GB, Thomas J, Fine HC, et al. 2014. Older adults don't notice their names: a new twist to a classic attention task. *J. Exp. Psychol. Learn. Mem. Cogn.* 40(6):1540–50
- Naveh-Benjamin M, Old SR. 2008. Aging and memory. In *Learning and Memory: A Comprehensive Reference*, ed. JH Byrne, H Eichenbaum, R Menzel, HL Roediger, D Sweatt, pp. 787–808. Oxford, UK: Elsevier
- Neely JH. 1977. Semantic priming and retrieval from lexical memory: roles of inhibitionless spreading activation and limited-capacity attention. *J. Exp. Psychol. Gen.* 106:226–54
- Norman DA. 1968. Toward a theory of memory and attention. *Psychol. Rev.* 75(6):522–36
- Oberauer K. 2019. Working memory and attention—a conceptual analysis and review. *J. Cogn.* 2(1):36
- Oberauer K, Eichenberger S. 2013. Visual working memory declines when more features must be remembered for each object. *Mem. Cogn.* 41:1212–27
- Oberauer K, Farrell S, Jarrold C, Pasiecznik K, Greaves M. 2012. Interference between maintenance and processing in working memory: the effect of item-distractor similarity in complex span. *J. Exp. Psychol. Learn. Mem. Cogn.* 38(3):665–85
- Oberauer K, Lewandowsky S. 2008. Forgetting in immediate serial recall: decay, temporal distinctiveness, or interference? *Psychol. Rev.* 115:544–76
- Oberauer K, Lewandowsky S. 2019. Addressing the theory crisis in psychology. *Psychon. Bull. Rev.* 26(5):1596–618
- Oberauer K, Lin HY. 2017. An interference model of visual working memory. *Psychol. Rev.* 124:21–59
- Palva JM, Monto S, Kulashekhar S, Palva S. 2010. Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *PNAS* 107:7580–85
- Panichello MF, Buschman TJ. 2021. Shared mechanisms underlie the control of working memory and attention. *Nature* 592(7855):601–5
- Parton A, Malhotra P, Husain M. 2004. Hemispatial neglect. *J. Neurol. Neurosurg. Psychiatry* 75(1):13–21
- Passolunghi MC, Siegel LS. 2001. Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *J. Exp. Child Psychol.* 80(1):44–57
- Petersen A, Vangkilde S. 2022. Decomposing the attentional blink. *J. Exp. Psychol. Hum. Percept. Perform.* 48(8):812–23
- Popov V, Reder LM. 2020. Frequency effects on memory: a resource-limited theory. *Psychol. Rev.* 127(1):1–46
- Postle BR, Oberauer K. 2022. Working memory. In *The Oxford Handbook of Human Memory*, ed. MJ Kahana, AD Wagner. Oxford, UK: Oxford Univ. Press. In press. <https://doi.org/10.31234/osf.io/963kf>**
- Rabinowitz JC, Craik FIM, Ackerman BP. 1982. A processing resource account of age differences in recall. *Can. J. Psychol.* 36(2):325–44
- Raye CL, Johnson MK, Mitchell KJ, Greene EJ, Johnson MR. 2007. Refreshing: a minimal executive function. *Cortex* 43:135–45

Important review covering basic research on attention and working memory from brain and behavioral standpoints.

- Reinhart RM, Woodman GF. 2014. High stakes trigger the use of multiple memories to enhance the control of attention. *Cereb. Cortex* 24:2022–35
- Rhodes S, Cowan N. 2018. Attention in working memory: Attention is needed but it yearns to be free. *Ann. N.Y. Acad. Sci.* 1424(1):52–63
- Ricker TJ, Hardman KO. 2017. The nature of short-term consolidation in visual working memory. *J. Exp. Psychol. Gen.* 146(11):1551–73
- Ricker TJ, Sandry J, Vergauwe E, Cowan N. 2020. Do familiar memory items decay? *J. Exp. Psychol. Learn. Mem. Cogn.* 46(1):60–76
- Ricker TJ, Vergauwe E. 2022. Boundary conditions for observing cognitive load effects in visual working memory. *Mem. Cogn.* 50:1169–85
- Röer JP, Bell R, Buchner A. 2015. Specific foreknowledge reduces auditory distraction by irrelevant speech. *J. Exp. Psychol. Hum. Percept. Perform.* 41:692–702
- Röer JP, Bell R, Körner U, Buchner A. 2019. A semantic mismatch effect on serial recall: evidence for interlexical processing of irrelevant speech. *J. Exp. Psychol. Learn. Mem. Cogn.* 45:515–25
- Röer JP, Cowan N. 2021. A preregistered replication and extension of the cocktail party phenomenon: One's name captures attention, unexpected words do not. *J. Exp. Psychol. Learn. Mem. Cogn.* 47:234–42
- Rose NS. 2020. The dynamic processing model of working memory. *Curr. Dir. Psychol. Sci.* 29:378–87**
- Ruch S, Henke K. 2020. Learning during sleep: a dream comes true? *Trends Cogn. Sci.* 24:170–72
- Salthouse TA. 2010. Selective review of cognitive aging. *J. Int. Neuropsychol. Soc.* 16(5):754–60
- Salthouse TA. 2021. Individual differences in working memory and aging. In *Current Issues in Memory: Memory Research in the Public Interest*, ed. J Rummel, pp. 299–318. London: Routledge
- Saults J, Cowan N, Sher KJ, Moreno MV. 2007. Differential effects of alcohol on working memory: distinguishing multiple processes. *Exp. Clin. Psychopharmacol.* 15:576–87
- Schacter DL. 1990. Perceptual representation systems and implicit memory: toward a resolution of the multiple memory systems debate. *Ann. N.Y. Acad. Sci.* 608:543–71
- Schurgin MW, Wixted JT, Brady TF. 2020. Psychophysical scaling reveals a unified theory of visual memory strength. *Nat. Hum. Behav.* 4:1156–72
- Scoville WB, Milner B. 1957. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* 2:11–21
- Shallice T, Papagno C. 2019. Impairments of auditory-verbal short-term memory: Do selective deficits of the input phonological buffer exist? *Cortex* 112:107–21
- Shiffrin RM, Schneider W. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol. Rev.* 84:127–90
- Silveri MC, Reali G, Jenner C, Puopolo M. 2007. Attention and memory in the preclinical stage of dementia. *J. Geriatr. Psychiatry Neurol.* 20(2):67–75
- Skodzik T, Holling H, Pedersen A. 2017. Long-term memory performance in adult ADHD: a meta-analysis. *J. Atten. Disord.* 21(4):267–83
- Sokolov EN. 1963. *Perception and the Conditioned Reflex*. New York: Pergamon Press
- Souza AS, Oberauer K. 2016. In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Atten. Percept. Psychophys.* 78:1839–60
- Sperling G. 1960. The information available in brief visual presentations. *Psychol. Monogr.* 74(11):1–29
- Theeuwes J, Bogaerts L, van Moorselaar D. 2022. What to expect where and when: how statistical learning drives visual selection. *Trends Cogn. Sci.* 26(10):860–72
- Treisman M, Rostron AB. 1972. Brief auditory storage: a modification of Sperling's paradigm. *Acta Psychol.* 36:161–70
- Uittenhove K, Chaabi L, Camos V, Barrouillet P. 2019. Is working memory storage intrinsically domain-specific? *J. Exp. Psychol. Gen.* 148(11):2027–57**
- Ünal ZE, Forsberg A, Geary DC, Cowan N. 2022. The role of domain-general attention and domain-specific processing in working memory in algebraic performance: an experimental approach. *J. Exp. Psychol. Learn. Mem. Cogn.* 48:348–74**
- Unsworth N, Miller AL. 2021. Individual differences in the intensity and consistency of attention. *Curr. Dir. Psychol. Sci.* 30(5):391–400

Provides a generally accessible review of neuroimaging research on the relation between attention and memory.

Demonstrates, using recall as opposed to recognition, when working memory storage is general across domains.

Illustrates that individual differences in attention control are important for one type of practical cognition.

- Unsworth N, Robison MK, Miller AL. 2022. On the relation between working memory capacity and the antisaccade task. *J. Exp. Psychol. Learn. Mem. Cogn.* 48(10):1420–47
- Unsworth N, Schrock JC, Engle RW. 2004. Working memory capacity and the antisaccade task: individual differences in voluntary saccade control. *J. Exp. Psychol. Learn. Mem. Cogn.* 30(6):1302–21
- van Ede F, Nobre AC. 2023. Turning attention inside out: how working memory serves behavior. *Annu. Rev. Psychol.* 74:137–65
- Van Gerven PW, Hurks PP, Bovend'Eerdt TJ, Adam JJ. 2016. Switch hands! Mapping proactive and reactive cognitive control across the life span. *Dev. Psychol.* 52:960–72
- Vandierendonck A. 2016. A working memory system with distributed executive control. *Perspect. Psychol. Sci.* 11:74–100
- Vergauwe E, Barrouillet P, Camos V. 2010. Do mental processes share a domain general resource? *Psychol. Sci.* 21:384–90
- Vergauwe E, von Bastian CC, Kostova R, Morey CC. 2022. Storage and processing in working memory: A single, domain-general resource explains multitasking. *J. Exp. Psychol. Gen.* 151(2):285–301
- Watkins MJ. 1984. Models as toothbrushes. *Behav. Brain Sci.* 7(1):86
- White TG. 1982. Naming practices, typicality, and underextension in child language. *J. Exp. Child Psychol.* 33:324–46
- Wolfe JM. 2021. Guided Search 6.0: an updated model of visual search. *Psychon. Bull. Rev.* 28:1060–92
- Wood NL, Stadler MA, Cowan N. 1997. Is there implicit memory without attention? A reexamination of task demands in Eich's 1984 procedure. *Mem. Cogn.* 25:772–79
- Xie W, Campbell S, Zhang W. 2020. Working memory capacity predicts individual differences in social-distancing compliance during the COVID-19 pandemic in the United States. *Psychol. Cogn. Sci.* 117(30):17667–74
- Yue Q, Martin RC, Hamilton AC, Rose NS. 2018. Non-perceptual regions in the left inferior parietal lobe support phonological short-term memory: evidence for a buffer account? *Cereb. Cortex* 29:1398–1413
- Zaretskaya N, Thielscher A, Logothetis NK, Bartels A. 2010. Disrupting parietal function prolongs dominance durations in binocular rivalry. *Curr. Biol.* 20:2106–11