Review article

Up to the magical number seven: An evolutionary perspective on the capacity of short term memory

Majid Manoochehri [*]

Tiktaalik Institute, Kerman, Iran

ABSTRACT

Working memory and its components are among the most determinant factors in human cognition. However, in spite of their critical importance, many aspects of their evolution remain underinvestigated. The present study is devoted to reviewing the literature of memory studies from an evolutionary, comparative perspective, focusing particularly on short term memory capacity. The findings suggest the limited capacity to be the common attribute of different species of birds and mammals. Moreover, the results imply an increasing trend of capacity from our non-human ancestors to modern humans. The present evidence shows that non-human mammals and birds, regardless of their limitations, are capable of performing memory strategies, although there seem to be some differences between their ability and that of humans in terms of flexibility and efficiency. These findings have several implications relevant to the psychology of memory and cognition, and are likely to explain differences between higher cognitive abilities of humans and non-humans. The adaptive benefits of the limited capacity and the reasons for the growing trend found in the present study are broadly discussed.

1. Introduction

The terms 'short term memory' (STM) and 'working memory' (WM) are sometimes used similarly or interchangeably (see Aben et al., 2012; Norris, 2017). Many elaborate studies, however, have proposed different definitions for them and have provided empirical evidence on their differences (for empirical evidence, see Engle et al., 1999; Kail and Hall, 2001; McCabe et al., 2005; for neural evidence, see Lewis-Peacock et al., 2012). The present study follows the idea that STM refers to the slave components of the WM system, known as phonological loop and visuo-spatial sketchpad, which are assumed to be the temporary, passive storages of material (e.g., Sebastián and Hernández-Gil, 2012; Sebastián and Mediavilla, 2015). This issue has been stipulated by Brydges et al. (2018):

> … STM could be considered a subcomponent of WM, consistent with Baddeley and Hitch's (1974) model of WM where the slave systems (the phonological loop and the visuospatial sketchpad) are STM constructs, and the central executive is associated with the active manipulation and updating of information. (p. 118) (see also Kail and Hall, 2001; McCabe et al., 2005)

Moreover, it should be pointed out that STM has been traditionally studied by *memory span* tests. Memory span refers to the ability of an individual to reproduce immediately, after one presentation, a series of discrete stimuli in their original order (Blankenship, 1938; Dempster, 1981; Toyoshima et al., 2018). In the recent years, the term 'memory span' has also been used to refer to the capacity of WM or the focus of attention. However, in the current study, this term has mostly been employed to denote the capacity of STM.

Of major concern of the present study, however, is that STM is typically measured with traditional simple span tasks (e.g., forward digit span, word span, etc.), whereas WM is estimated with complex span tasks (e.g., reading span, operation span, etc.) (Bailey et al., 2011; Rowe et al., 2019). Numerous studies indicated a remarkable difference between the results of these two test groups. The mean STM score of adults is estimated to be about 6.56 ($\pm$2.39), while the mean WM score of adults is estimated to be about 4.88 ($\pm$2.58) (Gignac, 2015). WM capacity, however, has been estimated or suggested almost always much smaller than STM capacity. Cowan, in a series of studies (e.g., Cowan, 2001, 2010), suggested a capacity of 4 ($\pm$1) for WM. In addition, Gobet and Clarkson (2004), based on a series of experiments with chess players, discussed that Cowan's proposal that WM holds four chunks may be an

overestimate since the real capacity appears to be below three and even close to two chunks.

WM is a determinant factor in human cognition. Many of cognitive processes, if not all of them, are affected by WM's characteristics and functions. This makes it necessary to understand the selection pressures that shaped the present features of WM and the evolutionary forces that are still affecting this system. As noted by Conway and Christiansen (2001), human cognition is merely one specific instance of animal cognition, evolved in the same place and from the same ancestors. Therefore, by exploring the capabilities and limitations of our closer evolutionary relatives, as well as their similarities and differences with us, we can come to understand the underlying nature of our cognitive system. Nevertheless, despite more than a century of research on memory span, only scant studies have reviewed the literature from an evolutionary perspective (e.g., Read, 2008).

The objective of this study is to provide a comprehensive picture of the available evidence as well as to highlight similarities and differences between human and non-human cognition, focusing particularly on the capacity of STM. Given this, there are unanswered questions, which this study aims to explore. What is the pattern of STM evolution and what does that pattern mean? What selective forces have generated characteristics and alterations of STM during the course of evolution? And what are the outcomes of these possible alterations? Needless to say, in order to answer these questions, in addition to capacity, other features and functions of WM, such as memory strategies, must be reviewed and discussed. To this end, I begin first with reviewing experimental studies concerning the capacity of STM.

2. Capacity

In the first place, it has to be mentioned that studies that used simple span tasks to estimate the capacity of the temporary, passive storage of material are of most relevance for this review. The pioneering study of Miller (1956), which suggested the *magical number seven* as the capacity of STM was according to the findings of simple span tasks. In addition, in the recent years there has been a growing interest in using simple span tasks in animal studies. But the findings of these studies have not received adequate attention or in some cases have been misinterpreted.

However, it is illuminating to briefly outline the results of the few available reviews, which studied the capacity of WM form a comparative view. In one of these studies, Read (2008) drew the conclusion that the capacity of WM in chimpanzees (perhaps as the highest or equal highest among primates) is only about two items. Contrary to this, Carruthers (2013) suggested that the WM limits of at least primate species fall within the human range. Both studies agree on a capacity of about three to four items for human WM.

2.1. Humans

Hermann Ebbinghaus (1850–1909) is known as one of the first researchers to study the capacity of STM. He examined the number of trials that it took him to memorize sequences of nonsense syllables (Ebbinghaus, 1885). In reporting his results, he commented that: "The question can be asked: What number of syllables can be correctly recited after only one reading? For me the number is usually seven" (p. 36). In the years after him, several studies of memory span with various methods, tasks, and materials have been done by others (for major reviews, see Blankenship, 1938; Dempster, 1981). Surprisingly, the average memory span of adults was almost always the same. In 1956, Miller, in his classical article, stipulated that: "Everybody knows that there is a finite span of immediate memory and that for a lot of different kinds of test materials this span is about seven items in length" (p. 11). The subsequent findings from other nations were also in accordance with these conclusions, suggesting that a memory span of about seven items is a universal characteristic of human beings. The more recent psychological literature also confirms these results (e.g., Chen et al., 2009; Gignac, 2015;

Ostrosky-Solís and Lozano, 2006; Schwartze et al., 2019). Klingberg (2009), for instance, noted that:

> … if we ask twenty students to remember a series of random digits, most of them will be able to repeat between six and eight of them. If we test their visuospatial memories, some will remember five positions and some eight; whatever the results, the average will often lie remarkably close to Miller's limit of seven. (p. 56)

2.2. Non-humans

In the first place, it should be noticed that owing to the difficulties of studying cognitive functions in non-human animals, designed tasks in most of the cases are easier than classical memory span tests in humans. In order to design an experiment that carefully measures STM span, some points should be taken into consideration. In the classical literature of human memory, the duration of STM has been considered to be 30 s or less if the material is not rehearsed (Atkinson and Shiffrin, 1968; Hatano and Osawa, 1983). Moreover, memory span scores are sensitive to practice effects (Ericsson et al., 1980; Martin and Fernberger, 1929; Rowe, Hasher and Turcotte, 2008), size of items (Baddeley et al., 1975; Campoy, 2008; Chen et al., 2009; Tolan and Tehan, 2005), and types of tests (recollection vs. recognition; see Chubala et al., 2019). Indeed, memory span is about seven items when materials are as large as digits or letters, and the test is a recollection test. Contrary to these primary bases, most studies of non-humans include several hours of practice before the main test or, in some cases, have ignored the importance of time intervals (see also Carruthers, 2013). As the last point, it should be noted that because of the scarcity of experimental studies, the focus of this section is only on mammalian and avian species.

2.2.1. Chimpanzees and other non-human primates

Chimpanzees have been known as one of our nearest primate relatives and also as one of the smartest non-human animals. Therefore, it may be expected to find the largest memory span of non-humans in chimpanzees. A line of studies by Japanese researchers shows that the average performance of chimpanzees in a visuospatial memory span task is slightly lower than humans. In one study, a female adult chimpanzee could remember the correct sequence of five items (Kawai and Matsuzawa, 2000). In another study by Inoue and Matsuzawa (2009), the authors examined five chimpanzees, including three young and two adults. The two adults could remember four items. Also, two of the young chimpanzees (both female) managed to obtain the scores six and seven, and the last young subject (a male named Ayumu) could remember nine items. In order to estimate memory span in these studies, the Arabic numerals (1–9) were presented on the touch-screen monitor. Initially, subjects mastered the skill of touching numerals 1 to 9 in ascending order. Then, in the main task, numerals appeared in random locations of the screen and after touching the first numeral all other numerals were replaced by white squares. The subject had to remember which numeral appeared in which location, and then touch them based on the knowledge of numerical sequence (Inoue and Matsuzawa, 2009). These findings by Japanese researchers, particularly the results of the last young male (i.e., Ayumu), have prompted the notion that chimpanzees have a better or equal memory span than humans. But it has to be mentioned that chimpanzees received many training sessions during the experiment, which has the potential of increasing performance. Cook and Wilson showed that after adequate practice, human subjects (i.e., university students) substantially outperformed chimpanzees (Cook and Wilson, 2010). Furthermore, some researchers commented on these results that the first number was touched while it was visible on the screen, and so it did not need to be memorized (Read, 2008) (for additional critiques, see Carruthers, 2013).

The memory span of other non-human primates appears to be close to that of chimpanzees. Studies in baboons revealed a capacity of about four

to five items. Fagot and De Lillo (2011) studied two male baboons (*P. papio*) by using a similar task to that described above. Their results showed that one baboon had a span of four and the other of five items. The authors also tested human subjects (i.e., undergraduate students) to compare the performance of humans and baboons. As expected, human subjects showed a quantitatively longer memory span. They stipulated that even after extensive training, the immediate serial spatial span of baboons does not match that of humans tested in similar conditions. Another study of immediate serial recall by Botvinick et al. (2009) showed a similar memory span to that of baboons for rhesus monkeys (*Macaca mulatta*). Moreover, Buschman et al. (2011) reported data indicating a capacity between three and four objects for visual STM of rhesus macaques.

2.2.2. Non-primate mammals and birds

In comparison to primates, non-primate mammals and birds do not appear to have larger memory spans. Herman (2010) studied the memory span of bottlenose dolphins (*Tursiops truncatus*) for sounds. A female dolphin listened to a list of novel sounds, each of 2 s duration and separated by 0.5 s silent intervals. After finishing the list, the probe sound was presented. The dolphin had to determine whether the probe sound was a member of the list or not. According to this experiment, Herman suggested a memory span of about four to five items for dolphins. In the case of pigeons, Terrace (1993) has discussed that "The amount of time it takes a pigeon to learn a four-item list (3–4 months) suggests that four items may approach the limit of the pigeon's memory span" (p. 164). Moreover, Balakhonov and Rose (2017), according to a series of experiments, proposed that crows (*Corvus corone*) and monkeys (*Maccaca mulatta* and *Maccaca fascicularis*) show a remarkably similar capacity of about four items in visual STM. Also, researchers who studied memory span of rats by using object recognition tasks proposed a capacity of about four items for this species (Sugita et al., 2015; Toyoshima et al., 2018). It has to be said that in these studies with rats, the time length of different phases of the experiments is longer than the standard duration (e.g., 5 min for sample phase, 5 min for a delay period, etc.). However, due to the overall lack of evidence, these findings have been presented here to provide support for other data.

The obvious implication of these data is that the limited capacity is the common attribute of different species of birds and mammals (for a similar conclusion, see Wright and Elmore, 2016). Moreover, these data imply that STM of humans has the largest capacity among mammalian and avian species. This latter conclusion is in accordance with studies that examined different species with the same STM task. Fagot and De Lillo (2011), for instance, measured baboons and humans (i.e., undergraduate students) by the same memory span test. As expected, human subjects showed a quantitatively longer memory span. Besides, Wright (1989) indicated that in a serial probe recognition task with visual stimuli, humans outperform monkeys and monkeys outperform pigeons. He noted that pigeons did acquire the serial probe recognition task, but they required more training and shorter list lengths than monkeys. In another study, Wright and Elmore (2016) indicated that monkeys and pigeons were considerably less accurate than humans in a visual STM task with the same array sizes (2, 4, and 6 items). Taken together, these findings imply an increasing trend of capacity from our non-human ancestors to modern humans. This is in accordance with Coolidge and Wynn (2005) archeological argument, which proposed that an enhancement of capacities occurred in the relatively recent human past, most likely after the first appearance of anatomically modern humans. Once we accept the idea of an increase in STM capacity, the question arises of what is the cause of this phenomenon. In later sections of this review, some possibilities are discussed.

3. Invariability

In addition to its limited capacity, memory span has other unique characteristics, one of which is invariability. Wechsler (1939) argued

that approximately 90% of the adult population appear to recall somewhere between five and eight digits (for supporting evidence, see GréGoire & Van Der Linden, 1997; Karakaş et al., 2002; Manoochehri, 2020a). In accordance with the invariability hypothesis of Wechsler, surprisingly, there is almost no report of a true large memory span. It should be noted that there are, of course, some reports of extreme performances, such as a memory span of about 80 items obtained by Subject S.F., who managed to obtain this score after more than 230 h of practice in the laboratory (Ericsson et al., 1980; for a review of similar cases, see Parker et al., 2006). But neither of these cases can be considered as a true large memory span. As discussed by Parker et al. (2006), these high scores are because of using advanced strategies, rather than the true capacity of STM. Subject S.F., for instance, was only able to gain high scores in forward digit span test, but neither in letter span test nor when he was tested by uncodable sequences of digits (Ericsson et al., 1980). In addition, generally speaking, it is more difficult to find a score higher than 8 than a score lower than 5 in normal groups of young adults (GréGoire and Van Der Linden, 1997; Karakaş et al., 2002). Consistent with the invariability hypothesis, the memory span scores do not show remarkable sex differences. Indeed, many previous studies failed to observe any difference (e.g., GréGoire and Van Der Linden, 1997; Monaco et al., 2013) or only observed small differences (e.g., Choi et al., 2014; Lynn and Irwing, 2008; Sebastián and Mediavilla, 2015).

Up to here, we discussed the invariability of memory span scores in human subjects. Interestingly, the results of the current review convey the idea that, on a larger scale, there is also invariability of performance across mammalian and avian species. This idea can be further supported by the findings of a meta-analysis review by Lind et al. (2015). They indicated that there is a remarkable similarity among different species in delayed matching-to-sample tasks. They expressed that even primates do not appear to stand out from other mammals (or other species in general) in either zero-delay performance or performance half-life. Given this discussion, the question arises, what is the reason behind the invariability of STM capacity?

4. Resistance to the Flynn effect

Memory span scores are also resistant to the Flynn effect. A substantial amount of empirical evidence suggests that some cognitive ability test scores have increased significantly since about 1930. This phenomenon is called the Flynn effect. The effects have been found to be more substantial on measures of fluid intelligence, such as Raven's Progressive Matrices (Flynn, 2009; Gignac, 2015). As discussed by Gignac (2015), if the Flynn effect is occurring, it would appear to be a phenomenon that is completely independent of memory span. He stipulated that there were no meaningful changes in memory span scores (including both STM and WM span scores) from 1923 to 2008.

The important point is that while some factors were and are powerful enough to affect other cognitive ability test scores, memory span scores are not affected by them. Why is that so? One possibility is that the present size of STM is a vital characteristic for cognitive systems. Therefore, it resists against the Flynn effect or, as discussed in the previous section, it shows a small variability in humans and even in a variety of phylogenetically distant species. By the way, if we accept that the present size of STM is a vital feature of human and animal cognition, the next key question, naturally, is why? Until today, few articles have investigated this issue. The last section of this article is devoted to reviewing these studies.

5. Memory strategies

Humans use a variety of strategies (e.g., rehearsal, chunking, visualization, etc.) in daily life. This is particularly more apparent in the face of higher cognitive functions, such as language or complex problem solving. Most of the people have a good mastery over a range of strategies and use them to enhance cognitive performances. This is the case not only in

general, but also in unexpected and time-limited cognitive tasks. Because of the prominent role of strategies in WM functions, they have been discussed in memory studies, since the early works (e.g., Blankenship, 1938). Owing to the central and critical role of rehearsal and chunking in the WM system, the focus of this review is only on these two major strategies. Furthermore, much of the discussion is based on two primary questions: can non-humans perform a certain strategy? And what are the differences between human and non-human performance?

5.1. Rehearsal

We are not the only species who can use strategies. A growing body of animal studies provides evidence of performing memory strategies by non-humans. Rehearsal is a relatively simple and underlying strategy, which is likely to be found in different animal species. The future planning behavior is interpreted by some researchers, such as Carruthers, as the hallmark of this strategy in non-humans. This is because it seems impossible to perform a true future planning behavior without mentally rehearsing information (Carruthers, 2013). One instance of future planning, discussed by Carruthers, is the stone-collecting behavior of an alpha male chimpanzee (Osvath, 2009; Osvath and Karvonen, 2012). The animal collects and stores piles of stones early in the morning to throw at zoo visitors later in the day. As argued by Carruthers, at the times when he collected and concealed his stashes, he was in a calm state, in the absence of the stimuli (i.e., human visitors) that would provoke his rage later. A similar instance of future planning in chimpanzees is the case of using tools by wild chimpanzees of Congo to extract termite prey (Sanz et al., 2004). Carruthers (2013) explained that chimpanzees never arrived at the site of a subterranean termite nest without bringing a stick unless one had previously been left at the site. And this was true even though the nearest appropriate tree was tens of meters away in the forest, from which point the nest site could not be seen (for more evidence of future planning in primates, see Mulcahy and Call, 2006; Naqshbandi and Roberts, 2006; Osvath and Osvath, 2008). In the case of birds, Taylor et al. (2010) indicated that crows are able to settle a novel three-stage metatool problem, which demands complex cognition and future planning. Planning for the future has also been shown in western scrub-jays (Raby et al., 2007; for a review of future cognition in non-humans, see Roberts, 2012). These findings indicate that rehearsal does exist in non-human animals, and at least some species of primates and birds can use this strategy. Yet, despite the similarities, it could be said that no information processing system among animal species relies on rehearsing information as much as ours, because we spend a vast amount of time to review the past events or to plan for the future, and this is of key importance for our higher cognitive functions.

Elaborate studies of memory reveal that there are more than one type of rehearsal (Thalmann et al., 2019b). Besides, needless to say, our assessment of the ability of an animal to use rehearsal strategy depends on our definition of this mental process. For instance, the basis of this discussion and its conclusion up to this point was Carruthers's definition, which considered rehearsal as the off-line rehearsals of action schemata that can be used to populate and sustain some of the contents of WM (i.e., elaborative rehearsal). In contrast to this, if we define rehearsal as a continuous refreshing process that sustains representations of WM (i.e., maintenance rehearsal) (Jonides et al., 2008; Thalmann et al., 2019b), it will be difficult to imagine a species among birds and mammals that does not use this strategy. It is because many activities of non-human animals seem to be impossible to accomplish without a mental process that sustains representations of WM. Foraging situations are perhaps among the best examples. For instance, while a prey species is approached or attacked by a predator and tries to avoid capture, there can be moments that the prey does not receive any kind of sensory input from the predator, and this does not cause the prey to stop evasive behaviors. Needless to say, a similar discussion can be held for predators. Indeed, empirical evidence suggests that a wide range of animal species are capable of maintenance rehearsal in such situations (e.g., Cross and Jackson, 2017).

Lastly, it is probably the unique characteristic of humans that they deliberately repeat information in order to memorize it, and by doing so they purposely take advantage of rehearsal strategy (Atkinson and Shiffrin, 1968; Shen, 2004; Thalmann et al., 2019b). From this last aspect, humans are able to use this strategy more consciously and more intentionally.

Taken together, in spite of the similarities, some of these data give the impression that non-humans are not as flexible and efficient as humans in performing rehearsal strategy. As a further case in point, Wright (1989) indicated that at least some of the primates have some difficulties taking advantage of rehearsal in time-limited laboratory tasks of WM. By using a recognition memory task in the laboratory, he indicated that monkeys do not show the typical interstimulus interval (blank time between items) results found with humans. In human subjects, increasing the interstimulus interval improves performance.

5.2. Chunking

Chunking is the process of grouping information into units and giving a label to them so that sets of information can be efficiently represented and used as integrated units (Huang and Awh, 2018; Miller, 1956). The prominent role of chunking in memory and other cognitive functions, such as language, visual perception, and motor skills, has been widely discussed by numerous studies (e.g., Chekaf et al., 2016; Fonollosa et al., 2015; Huntley et al., 2011; Jones, 2012; McCauley and Christiansen, 2015; McCauley et al., 2017; Solopchuk et al., 2016). In addition, chunking has been suggested to have a key role in the problem-solving process (Leighton and Sternberg, 2003), though chunking itself can be seen as an incipient form of problem solving. Chunks directly represent the results of previously-encountered subproblems. This means that fewer subgoals and, therefore, fewer steps are necessary to solve a problem (Bayazitoglu et al., 1993). Chunking has also been proposed to be a form of the brain's categorization or organization (Capaldi, 2003; Fonollosa et al., 2015).

An advantage of chunking is that it is then unnecessary to keep all items of the chunk in the capacity-limited WM, but just some index to each chunk (Chen and Cowan, 2009). Indeed, chunking reduces the load on WM and provides an opportunity for the information processing system to rely on LTM (Thalmann, Souza and Oberauer, 2019a). It is necessary to add that some studies proposed that despite its advantages, chunking has some costs for information processing systems, such as decreasing the speed (Bayazitoglu et al., 1993; Fonollosa et al., 2015; Norris et al., 2019). In an evolutionary context, this means that an information processing system may evolve to prefer simple solutions that are faster to achieve but are not very efficient or may evolve to prefer more elaborate solutions that need more time to achieve but are more profitable and effective. Needless to say, chunking reduces the time needed for processing information when the complexity of problems is the same (Jones, 2012). Taken together, in spite of its simple nature, chunking appears to be one of the most important functions in our information processing system.

Clearly, the WM system has a central role in forming and learning new chunks (Chen and Cowan, 2005). Also, given the present discussion, it may be helpful to distinguish two types of chunking: simple chunking and hierarchical chunking. While simple chunking refers to forming a single chunk by using elementary items, hierarchical chunking refers to the process in which already existing chunks or their indexes are grouped to form new chunks and these in turn shape super chunks and so forth (for hierarchical chunking, see Conway and Christiansen, 2001; Fonollosa et al., 2015; Rabinovich et al., 2014). Language is a very good example of hierarchical chunking. Moreover, hierarchical chunking has been considered by some researchers as the most complex form of sequential learning (Conway and Christiansen, 2001).

Similar to what was discussed for maintenance rehearsal, it is difficult to imagine how non-humans could accomplish some of their activities without chunking information. For instance, at least in some mammalian

and avian species, identifying members of their own group, mate, offspring, prey, and predators may be partly by relying on chunking. Even encounters with predators are not always the same, as at any given time some predators are in hunting mode while others are not actively hunting prey (Dugatkin, 2013). It is not very unlikely that in such situations chunking plays a role in processing information.

In spite of its critical importance, there are little comparative studies on chunking. There is, however, sufficient evidence to indicate that non-humans can apply chunking. I shall perhaps begin with primates. De Lillo et al. (1997) tested the search abilities of capuchin monkeys. Their results showed that the search efficiency of monkeys was higher in a search space suitable to organization in spatial chunks. In addition, Scarf et al. (2018) reported data indicating that both humans and monkeys chunk sequences of visual items (see also Terrace, 2002). One of their findings, which should be highlighted, is that monkeys, similar to humans, spontaneously chunk sequences that are composed of uniform items. This is important because many similar studies used sequences that their elements share some visual or spatial characteristics. They also observed that both species regularly paused at one or multiple points in sequences. Needless to say, pausing is well-known as the hallmark of chunking (Fonollosa et al., 2015). Moreover, while sequences that were used for humans were one item longer (eight vs. seven), monkeys needed several more trials to acquire the lists (e.g., about 40 times more to acquire the first list), which also means a slow process for forming new chunks. A more recent study by Völter et al. (2019) provided some evidence of chunking in WM tasks in chimpanzees. They discussed that their best performing subject seemed to engage in a chunking strategy in that he tended to end his search with the outer stimuli. According to their findings, they declared that differences in using strategies seem to be a promising candidate for a dividing line between humans and chimpanzees.

Primates are not the only ones who can apply chunking. Terrace indicated that pigeons are able to perform chunking strategy in serial learning tasks and by doing so they even increase their performance (Terrace, 1987, 1993). He demonstrated that pigeons learn five-element sequences of colors and achromatic geometric forms, segregated into distinct groupings, approximately twice as fast as homogeneous color sequences or heterogeneous, non-clustered sequences, which could not be chunked.

Another line of studies suggests that rats are also able to perform chunking (e.g., Capaldi et al., 1986; Dallal and Meck, 1990; Fountain and Annau, 1984). For example, Macuda and Roberts (1995) indicated that rats use chunking in a spatial memory task (i.e., radial maze) when it is possible and this improves their performance. In a review of literature, Capaldi (2003) described three different sorts of chunks of varying degrees of complexity (i.e., *trial*, *series*, and *list*) and maintained that rates are able to use all of them. The first and second types of chunking described by him (i.e., trial, series) are respectively what have been called simple and hierarchical chunking in the present review. To explain the second one, he noted that:

> The next highest chunk is called a series chunk, which consists of the animal's combining trial chunks into a higher level chunk. For example, a rat trained under four nonrewarded trials followed by a rewarded trial responds as follows: It begins by running slowly to the initial nonrewarded trial, the progressively increases its running speed over the successive nonreward trials, until by the terminal nonrewarded trial, the animal runs about as fast as it is able. Such responding indicates that the rat is treating the five trials, four nonrewarded followed by a reward, as a single organized whole or a chunk. (p. 406).

This indicates that non-humans can perform hierarchical chunking. But hierarchies, after all, can be different in complexity and the number of layers. Up to date, it is not clear to what extent non-humans can perform this process. But considering the fact that some higher cognitive functions that need complex hierarchical chunking do not exist in non-humans (e.g., language, see Manoochehri, 2020b; Progovac, 2019) or only exist in highly restricted forms (e.g., music, see Manoochehri, 2020c), it is tempting to propose that there is a significant inefficiency in performing complex hierarchical chunking in non-human animals (see also Conway and Christiansen, 2001). In this regard, Conway and Christiansen (2001) suggested that limitations in complex hierarchical chunking may help explain why non-human primates lack human-like language. This can be further supported by the evidence indicating the fundamental role of chunking in language processing (McCauley and Christiansen, 2015; McCauley et al., 2017). It is also tempting to postulate that more restricted slave components of the WM system in non-humans may not be able to efficiently support complex hierarchical chunking (for supportive neural evidence, see Aboitiz et al., 2010). It remains for future research to investigate hierarchical chunking in non-humans further. In summary, however, the results found here permit the conclusion that with some level of inefficiency, non-humans can perform chunking, which is similar to that found for rehearsal.

6. The primacy and recency effect

Human subjects show the primacy and recency effect in STM tasks (Düzel et al., 1996; Farrand et al., 2001; Jahnke, 1965). The primacy effect refers to good memory for first-presented items of lists and the recency effect refers to good memory for later-presented items (Martín et al., 2013; Wright, 1989). The primacy effect is a hallmark of LTM and the recency effect is a hallmark of WM (Düzel et al., 1996; Wright, 1989). This cognitive phenomenon has also been documented in different species of non-humans, including, for instance, monkeys (Botvinick et al., 2009; Castro and Larsen, 1992; Scarf et al., 2018; Wright, 2007), rats (Reed et al., 1991), rabbits (Wagner and Pfautz, 1978), dolphins (Herman, 2010), and pigeons (Wright, 2007). Among these studies, of particular interest for the present purposes are the studies of Wright (1989, 2007), because they examined the performance of different species, including humans, in the same task. Wright and colleagues showed that similar to humans, rhesus and capuchin monkeys, and pigeons exhibit both the primacy and recency effect in a visual probe recognition experiment. They indicated that for all four species at very short probe delays there was no primacy effect, only a strong recency effect (a condition to the advantage of WM and the disadvantage of LTM). At intermediate probe delays, the serial position functions were U-shaped, showing that there were primacy as well as recency effects. At the longest delay, there was no recency effect, only a primary effect (a condition to the advantage of LTM and the disadvantage of WM). These findings indicate the existence of a similar memory structure in these species. The reason for this conclusion is that if similar procedures trigger similar improvements or impairments of memory performance in different species, then one can infer that memory in these species relies on similar mechanisms (Bobrowicz, 2019). Wright and colleagues themselves repeatedly interpreted these findings as proof of similar underlying processes of memory in mammalian and avian species (e.g., Wright, 1989; Wright and Elmore, 2016). Similarities in the memory structures, in turn, guarantee the possibility of numerous comparative discussions, such as that of capacity in this article.

A significant difference among the species examined by Wright and colleagues was in the time course. The effective probe delays for humans were longer than monkeys, and for monkeys were longer than pigeons. For instance, in humans at a delay of 100 s, the function showed a strong primacy effect and no recency effect, which was longer than that of monkeys (30 s) and pigeons (10 s). These differences in time course indicate that the duration of STM in humans is longer than monkeys and in monkeys is longer than pigeons. This may also imply the notion that the duration of STM in humans is among the longer ones among mammalian and avian species (for more supportive evidence, see Fassihi et al., 2014; Smeele et al., 2019). However, more evidence is needed to draw a firm conclusion. Supportive evidence of a time-limited STM

storage for many species has been provided by Lind et al. (2015). Differences in the duration of STM might be at least partly because of differences in the memory span size.

7. Hypotheses concerning the capacity

Why memory span has a limited capacity or why there is an increasing trend of capacity towards humans? In the first place, I will argue the potential reasons for the limited capacity. In order to provide a more explicit discussion, the relevant studies are divided into two groups: those that based their discussion on a capacity about seven items or a temporary, passive storage (i.e., STM) and those that based their discussion on a capacity about three to four items or the focus of attention (i.e., WM).

7.1. Hypotheses of the limited capacity

7.1.1. STM hypotheses of the limited capacity

To begin with, some previous studies have suggested that "short-term memory limitations do not have a rational explanation" (Anderson, 1990, pp. 91/92) or larger capacities are biologically expensive or impossible. For instance, it has been postulated that greater STM size may have required additional tissue, which increases body mass and energetic expenditure, and therefore it is impossible with the biological characteristics of humans (e.g., Dukas, 1999). Other researchers rejected both of these assumptions (Todd et al., 2005). Moreover, the second assumption (i.e., assuming larger capacities as biologically expensive/impossible options) does not seem reasonable considering the diversity of extraordinary physiological and behavioral characteristics of different animal species. Also, if any of these suggestions is correct, we should perhaps be able to find various capacities of STM in different animals, which the present study does not indicate it.

One of the studies concerning the capacity of STM has been conducted by MacGregor (1987). Using a mathematical model, he highlighted the importance of efficient retrieval for STM. According to him, the limited capacity of STM could be the consequence of an efficiency of design. He argued that chunking facilitates retrieval when there are seven or five items in an unorganized memory. In a memory system evolved for efficiency, there is an upper effective limit to STM and a capacity beyond this limit would not be required.

In another study, Saaty and Ozdemir (2003) argued that in making preference judgments on pairs of elements in a group, the number of elements in the group should be no more than seven. The mind is sufficiently sensitive to improve large inconsistencies but not small ones and the most inconsistent judgment is easily determined. When the number of elements is seven or less, the inconsistency measurement is relatively large with respect to the number of elements involved. As the number of elements being compared is increased, the measure of inconsistency decreases slowly. Therefore, in order to serve both consistency and redundancy, it is best to keep the number of elements seven or less. When the number of elements increases past seven, the resulting increase in inconsistency is too small for the mind to single out the element that causes the greatest inconsistency to scrutinize and correct its relation to the other elements.

In a series of studies, Kareev has proposed that capacity limitation maximizes the chances for the early detection of strong and useful relations (Kareev, 1995; 2000; Kareev et al., 1997; for a controversial discussion of this hypothesis see Anderson et al., 2005; Juslin and Olsson, 2005; Kareev, 2005). From his standpoint, a STM capacity of size seven, which characterizes human adults, is of particular value in detecting imperfect correlations between features in the environment. The limited capacity may serve as an amplifier, strengthening signals which may otherwise be too weak to be noticed. He argued that, because correlations underlie all learning, their early detection is of great importance for the functioning and well-being of organisms. Therefore, the cognitive system might have evolved so as to increase the chances for early detection of strong correlations. In addition to the theoretical contribution, Kareev and colleagues in an experimental study found that people with smaller STMs are more likely to perceive a correlation than people with larger STMs (Kareev et al., 1997).

Some of the suggestions for the reason behind the limited capacity can be found in the studies of decision-making cognition. Here, it has been shown that people tend to rely on relatively *small samples* from payoff distributions (Hertwig and Pleskac, 2010). The size of these samples is often considered related to the capacity of STM (Hahn, 2014; Hertwig et al., 2004; Hertwig and Pleskac, 2010). In this context, a capacity-limited STM has been proposed as a possible cause (Hahn, 2014; Hertwig et al., 2004; Hertwig and Pleskac, 2010; Todd et al., 2005) or a requirement (Plonsky et al., 2015) for relying on small samples. More relevant to the present discussion, Todd et al. (2005) suggested that the benefits of using small samples or the costs of using too much information resulted in selective pressures that have produced particular patterns of forgetting in LTM and limits of capacity in STM (see also Hahn, 2014). So, what are these costs and benefits? Limited information use can lead simple heuristics to make more robust generalizations in new environments (Todd et al., 2005). Small samples amplify the difference between the expected earnings associated with the payoff distributions, thus making the options more distinct and choice easier (Hertwig and Pleskac, 2010). Relying on small samples has also been suggested to result in saving time and energy (Plonsky et al., 2015; Todd et al., 2005). Even if we assume that there is no cost (energy or time) for gathering information, by considering too much information, we are likely to add noise to our decision process, and consequently make worse decisions (Martignon and Hoffrage, 2002; Todd et al., 2005). Among these, the one which is perhaps associated with strong selective forces is saving time. There are different occasions that timely decisions play a vital role in the life of animals. But perhaps of most importance is the case of hunting situations. The encounters between prey and predators were an integral part of the daily life of our ancestors through deep evolutionary time. It is also clear that the penalties for any kind of inefficiency in such encounters are immediate and fatal, which thus results in intense selection for particular cognitive abilities and predation avoidance mechanisms (see Mathis and Unger, 2012; Rosier and Langkilde, 2011; Whitford et al., 2019). For instance, any prey that is attacked by several predators and cannot quickly decide which one to avoid at first or which way and which method to choose for escaping or perhaps defending will be eliminated at once. A similar discussion can be developed for predators (see Lemasson et al., 2009).

Another line of studies has stressed the importance of the limited capacity for foraging activities (e.g., Bélisle and Cresswell, 1997; Real, 1991; 1992; Thuijsman et al., 1995). According to it, the limited capacity may result in an overall optimization of food search behaviors. Similarly, Murray et al. (2017) have contended that the memory systems of anthropoids have been primarily evolved to reduce foraging errors. Foraging activities, however, do not appear to be the underlying reason for the capacity-limited STM. This is because, if foraging were the fundamental reason, then there would be remarkable sex differences in memory span, similar to that observed, for instance, in spatial abilities (Ecuyer-Dab and Robert, 2007; Voyer, Postma, Brake and Imperato-McGinley, 2007). According to the division of labor in ancestral hunter-gatherer societies, men were predominantly hunters and women were gatherers (Ecuyer-Dab and Robert, 2007; Marlowe, 2007), and it is likely that each one of these activities demands a different memory span. Namely, because a hunter has to focus on prey and ignore distracting information, while a successful gatherer can, or should, simultaneously consider many stationary targets (e.g., seeds, fruits, etc.). Contrary to this, many studies of sex differences in memory span show no significant difference (GrÉGoire and Van Der Linden, 1997; Monaco et al., 2013; Orsini et al., 1986; Peña-Casanova et al., 2009). Foraging activities, if they were the underlying reason, could also result in remarkable differences among different species. The present study, however, does not indicate such differences. Therefore, although the limited capacity may

have provided benefits for foraging activities, it seems reasonable to propose that foraging, after all, is not the main and direct reason for the limited memory space.

Among the hypotheses reviewed here, Kareev's suggestion (i.e., early detection of useful relations) is among the ones that have received relatively more attention. Also, his assumption seems reasonable in a comparative context and appears consistent with the findings of the present review. But of more importance is the fact that a memory system that has the ability of early detection of useful relations is likely to cause higher performance in associative learning and also saving time in decision making. In the case of learning, Kareev himself noted that: "Because correlations underlie all learning, their early detection and, subsequently, accurate assessment are of great importance for the functioning and well-being of organisms" (Kareev, 2000, p. 398). Leaning, certainly, is one of the first and main challenges of any cognitive system. Besides, there are broad similarities in basic forms of learning in different species (Dugatkin, 2013). It is also certain that through deep evolutionary time there has been intense selection for individuals with higher performance in learning. In this regard, Dugatkin (2013) stated that: "The ability to learn should be under strong selection pressure, such that individuals that learn appropriate cues that are useful in their particular environment should be strongly favored by natural selection" (p. 141). In summary, these considerations motivate the idea that associative learning and saving time in decision making are most likely the underlying reasons for the emergence and maintenance of limited capacity.

7.1.2. WM hypotheses of the limited capacity

There are, on the other side, some other studies of the limited capacity that based their analyses on a capacity about three to four chunks or the focus of attention (i.e., WM). Some of them will be briefly reviewed here. Sweller (2003), for instance, proposed that no more than two or three elements can be handled in WM, because any more elements would result in more potential combinations than could be tested realistically. According to him, as the number of elements in WM increases, the number of permutations rapidly becomes very large (e.g., $5! = 120$). With random choice, the greater the number of alternatives from which to choose while problem solving, the less likelihood that an appropriate choice will be made.

Many other possibilities have been discussed by Cowan (Cowan, 2001, 2005, 2010). For instance, based on the notion that it is biologically impossible for the brain to have a larger capacity, he declared that the representation of a larger number of items could fail because together they take too long to be activated in turn (Cowan, 2010). Another discussion by Cowan is that the WM capacity limit is the necessary price of avoiding too much interference (Cowan, 2005). According to him, activation of the memory system would go out of control if WM capacity was not limited to about four items at once. A relatively small central WM may allow all concurrently active concepts to become associated with one another without causing confusion or distraction (Cowan, 2010). Oberauer and Kliegl (2006) similarly stated that:

> The capacity of working memory is limited by mutual interference between the items held available simultaneously. Interference arises from interactions between features of item representations, which lead to partially degraded memory traces. The degradation of representations in turn leads to slower processing and to retrieval errors. In addition, other items in working memory compete with the target item for recall, and that competition becomes larger as more items are held in working memory and as they are more similar to each other. (p. 624).

7.2. The increasing trend of capacity

Archaeological evidence of an enhancement in the WM system has been presented by Coolidge and Wynn (2005) (see also Coolidge et al.,

2013; Wynn and Coolidge, 2006). The core idea of their hypothesis is a genetic mutation that affected neural networks approximately 60,000 to 130,000 years ago and increased the capacity of general WM and phonological storage. In the case of phonological storage, which is of more interest in the present review, they stipulated that: "A relatively simple mutation that increased the length of phonological storage would ultimately affect general working-memory capacity and language" (Coolidge and Wynn, 2005, p. 14). They proposed that the enhancement of WM capacities was the final piece in the evolution of human executive reasoning ability, language, and culture. From their point of view, the larger capacity is a necessary precondition for symbolic thought, which selective pressures contributed to the growth of it. They noted that an increase in WM capacities of pre-modern *H. sapiens* would have allowed greater articulatory rehearsal, consequently allowing for automatic long-term storage, and the beginnings of introspection, self-reflection, and consciousness. In line with Coolidge and colleagues' hypothesis, Aboitiz et al. (2010) proposed that during the course of human evolution, a development in the phonological loop occurred. They maintained that this development produced a significant increase in STM capacity and subsequently resulted in the evolution of language.

Many researchers, at least in the field of archeology, tend to agree with the idea of enhanced WM (e.g., Aboitiz et al., 2010; Haidle, 2010; Lombard and Wadley, 2016; Nowell, 2010; Putt, 2016, for a review of criticisms, see Welshon, 2010), though there seems to be a disagreement on its time. Almost all, however, suggest a time in the Pleistocene about or after the appearance of the genus *Homo* (Aboitiz et al., 2010; Coolidge and Wynn, 2005; Haidle, 2010; Putt, 2016). Some also suggest a gradual development (Haidle, 2010).

Once we accept the idea of the enhanced WM, important questions arise as to the cause and the process of this phenomenon. The enhancement of WM has been argued as a prerequisite for the evolution of some complex cognitive abilities of humans, such as language (Aboitiz et al., 2010; Coolidge and Wynn, 2005) and tool use (Haidle, 2010; Lombard and Wadley, 2016). For instance, Aboitiz et al. (2010) pointed out the existence of selective benefits for individuals with larger phonological capacities, which, in their view, were linguistically more apt. From their standpoint:

> The development of the phonological loop produced a significant increase in short-term memory capacity for voluntary vocalizations, which facilitated learning of complex utterances that allowed the establishment of stronger social bonds and facilitated the communication of increasingly complex messages, eventually entailing external meaning and generating a syntactically ordered language. (p. 55).

In the case of tool use, Haidle (2010) argued that a basic trait of all object behaviors is the increased distance between problems and solutions. Given this, more complex object behaviors possess longer distances. According to her, during the process of tool use the immediate desire (e.g., getting the kernel of the nut) must be set aside and replaced by one or several intermediate objectives, such as finding or producing an appropriate tool. Thus, thinking must depart from the immediate problem and shift to abstract conceptualizations of potential solutions, which results in sequences of physical actions with objects appropriate to achieve a solution in the near future (see also Lombard and Wadley, 2016). Given her discussion, it is clear how individuals or populations with an enhanced WM system, which provides the possibility of maintaining and manipulating more information, could take advantage of their superiority to excel others in tool-use performance and, consequently, to win competitions.

Arguably, if we assume the enhancement of WM as a gradual process, which has started long before our common ancestor with chimpanzees (as it was found by the present study), neither tool use nor language can be considered as the primary reason for it. But complex problem solving, because of its commonness, can be nominated as the primary cause,

which has then been supported by tool use and language (see also Putt, 2016). This assumption can be aided by the evidence indicating the critical role of an elaborate WM system in problem solving tasks (Logie et al., 1994; Zheng et al., 2011).

After all, there are many obscure aspects regarding the evolution of WM. Needless to say, deep disagreements in the related fields and issues, such as the process and timeline of language evolution (Progovac, 2019), make the puzzle more difficult to solve. It goes beyond the limits of the present article to pursue this further, but perhaps a possible way to settle this problem is looking for advantages and disadvantages of high and low STM/WM capacities (Engle, 2010). Such findings from experimental psychology in conjunction with archaeological and comparative evidence can shed light on the evolution of the WM system.

8. Conclusion

The first and obvious implication of the present findings is that the limited capacity is the common attribute of different species of birds and mammals. The present results also indicate an increasing trend of capacity from our non-human ancestors to modern humans. Among the potential explanations of the limited capacity, *associative learning* and *saving time in decision making*, particularly because of the strong selective forces that associate with them and their vital importance for different species, seem to be more likely. On the other hand, the enhancement of the WM system appears to be a prerequisite for the evolution of some higher cognitive abilities of modern humans, such as *language, tool use,* and *complex problem solving.*

A question yet to be answered is whether the current size of STM/WM in humans is the end of the line or not. The current size has been considered by some to be the end point (e.g., MacGregor, 1987). As opposed to it, Cowan declared that it is possible to imagine that larger capacities would have been preferable or doable, but still did not happen. Therefore, our current capacity just reflects our place in the middle of an ongoing evolutionary process, not an end point. If this is the case, one might expect the present capacity to expand in the future, assuming that it offers a sufficient survival advantage (Cowan, 2005). However, the current review suggests that considering the resistance of memory span scores to the Flynn effect, it is difficult to expect substantial changes in small periods of time.

All in all, many of us, instead of the wild nature, are living in artificial and unnatural environments. Are these unnatural environments along with their overwhelming and escalating complex problem solving tasks imperceptibly pushing us towards a WM system with even a larger capacity and, if so, what is the price for that? Here, the most important point to stress is that evolution does not drive towards perfection. However, thanks to our elaborate information processing system and consciousness, we now have the ability to purposefully plan for the future of our own evolution.

The evidence reviewed in this paper shows that many species of birds and mammals are capable of performing memory strategies, although there seem to be some differences between humans and non-humans in terms of flexibility and efficiency. An enhancement in the capacity of the WM system might be the reason, or part of the reason, for the emergence of superior memory strategies in humans.

Striking similarities in the primacy and recency effect in conjunction with other evidence, such as similarities in the size of STM and performing memory strategies, suggest a similar memory structure in different species of birds and mammals. This is in accordance with Wright's inference that there is a qualitative similarity in memory processing across mammals and birds. The present findings have several implications relevant to the psychology of memory and cognition. For instance, the differences found in the ability to perform memory strategies and the size of STM, may provide an explanation for some of the differences between cognitive abilities of humans and non-humans.

Declarations

Author contribution statement

Funding statement

Data availability statement

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

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Aben, B., Stapert, S., Blokland, A., 2012. About the distinction between working memory and short-term memory. Front. Psychol. 3 (AUG).

Aboitiz, F., Aboitiz, S., García, R.R., 2010. The phonological loop: a key innovation in human evolution. Curr. Anthropol. 51 (S1), S55–S65.

Anderson, J.R., 1990. The Adaptive Character of Thought. Psychology Press.

Anderson, R.B., Doherty, M.E., Berg, N.D., Friedrich, J.C., 2005. Sample size and the detection of correlation - a signal detection account: comment on Kareev (2000) and Juslin and Olsson (2005). Psychol. Rev. 112 (1), 268–279.

Atkinson, R.C., Shiffrin, R.M., 1968. Human memory: a proposed system and its control processes. In: Psychology of Learning and Motivation, 2. Elsevier, pp. 89–195.

Baddeley, A.D., Hitch, G., 1974. Working memory. Psychol. Learn. Motiv. 8, 47–89. Academic press.

Baddeley, A.D., Thomson, N., Buchanan, M., 1975. Word length and the structure of short-term memory. J. Verb. Learn. Verb. Behav. 14 (6), 575–589.

Bailey, H., Dunlosky, J., Kane, M.J., 2011. Contribution of strategy use to performance on complex and simple span tasks. Mem. Cognit. 39 (3), 447–461.

Balakhonov, D., Rose, J., 2017. Crows rival monkeys in cognitive capacity. Sci. Rep. 7 (1), 1–8.

Bayazitoglu, A., Johnson, T.R., Smith, J.W., 1993. Limitations of the unique-attribute representation for a learning system. In: Proceedings of 9th IEEE Conference on Artificial Intelligence for Applications. IEEE, pp. 219–225.

Bélisle, C., Cresswell, J., 1997. The effects of a limited memory capacity on foraging behavior. Theor. Popul. Biol. 52 (1), 78–90.

Blankenship, A.B., 1938. Memory span: a review of the literature. Psychol. Bull. 35 (1), 1.

Bobrowicz, K., 2019. Memory for Problem Solving: Comparative Studies in Attention, Working and Long-Term Memory. Lund University.

Botvinick, M.M., Wang, J., Cowan, E., Roy, S., Bastianen, C., Mayo, J.P., Houk, J.C., 2009. An analysis of immediate serial recall performance in a macaque. Anim. Cognit. 12 (5), 671–678.

Brydges, C.R., Gignac, G.E., Ecker, U.K.H., 2018. Working memory capacity, short-term memory capacity, and the continued influence effect: a latent-variable analysis. Intelligence 69, 117–122.

Buschman, T.J., Siegel, M., Roy, J.E., Miller, E.K., 2011. Neural substrates of cognitive capacity limitations. Proc. Natl. Acad. Sci. Unit. States Am. 108 (27), 11252–11255.

Campoy, G., 2008. The effect of word length in short-term memory: is rehearsal necessary? Q. J. Exp. Psychol. 61 (5), 724–734.

Capaldi, E.J., 2003. Animal memory and cognition. In: Handbook of Psychology. Wiley Online Library, pp. 399–422.

Capaldi, E.J., Nawrocki, T.M., Miller, D.J., Verry, D.R., 1986. Grouping, chunking, memory, and learning. Q. J. Exp. Psychol. Sect. B 38 (1b), 53–80.

Carruthers, P., 2013. Evolution of working memory. Proc. Natl. Acad. Sci. Unit. States Am. 110 (Supplement 2), 10371–10378.

Castro, C.A., Larsen, T., 1992. Primacy and recency effects in nonhuman primates. J. Exp. Psychol. Anim. Behav. Process. 18 (4), 335.

Chekaf, M., Cowan, N., Mathy, F., 2016. Chunk formation in immediate memory and how it relates to data compression. Cognition 155, 96–107.

Chen, Z., Cowan, N., 2005. Chunk limits and length limits in immediate recall: a reconciliation. J. Exp. Psychol. Learn. Mem. Cognit. 31 (6), 1235.

Chen, Z., Cowan, N., 2009. Core verbal working-memory capacity: the limit in words retained without covert articulation. Q. J. Exp. Psychol. 62 (7), 1420–1429.

Chen, Z.Y., Cowell, P.E., Varley, R., Wang, Y.C., 2009. A cross-language study of verbal

and visuospatial working memory span. J. Clin. Exp. Neuropsychol. 31 (4), 385–391.

Choi, H.J., Lee, D.Y., Seo, E.H., Jo, M.K., Sohn, B.K., Choe, Y.M., Woo, J.I., 2014. A normative study of the digit span in an educationally diverse elderly population. Psychiat. Invest. 11 (1), 39–43.

Chubala, C.M., Neath, I., Surprenant, A.M., 2019. A comparison of immediate serial recall and immediate serial recognition. Can. J. Exp. Psychol. 73 (1), 5.

Conway, C.M., Christiansen, M.H., 2001. Sequential learning in non-human primates. Trends Cognit. Sci.

Cook, P., Wilson, M., 2010. Do young chimpanzees have extraordinary working memory? Psychonomic Bull. Rev. 17 (4), 599–600.

Coolidge, F.L., Wynn, T., 2005. Working memory, its executive functions, and the emergence of modern thinking. Camb. Archaeol. J. 15 (1), 5–26.

Coolidge, F.L., Wynn, T., Overmann, K.A., 2013. The evolution of working memory. In: Working Memory. Psychology Press, pp. 51–74.

Cowan, N., 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav. Brain Sci. 24 (1), 87–114.

Cowan, N., 2005. Working memory capacity. In: Working Memory Capacity.

Cowan, N., 2010. The magical mystery four: how is working memory capacity limited, and why? Curr. Dir. Psychol. Sci. 19 (1), 51–57.

Cross, F.R., Jackson, R.R., 2017. Representation of different exact numbers of prey by a spider-eating predator. Interf. Focus 7 (3).

Dallal, N.L., Meck, W.H., 1990. Hierarchical structures: chunking by food type facilitates spatial memory. J. Exp. Psychol. Anim. Behav. Process. 16 (1), 69.

De Lillo, C., Visalberghi, E., Aversano, M., 1997. The organization of exhaustive searches in a patchy space by capuchin monkeys (Cebus apella). J. Comp. Psychol. 111 (1), 82.

Dempster, F.N., 1981. Memory span: sources of individual and developmental differences. Psychol. Bull. 89 (1), 63–100.

Dugatkin, L.A., 2013. Principles of Animal Behavior, Third international student edition. WW Norton & Company.

Dukas, R., 1999. Costs of memory: ideas and predictions. J. Theor. Biol. 197 (1), 41–50.

Düzel, E., Hufnagel, A., Helmstaedter, C., Elger, C., 1996. Verbal working memory components can be selectively influenced by transcranial magnetic stimulation in patients with left temporal lobe epilepsy. Neuropsychologia 34 (8), 775–783.

Ebbinghaus, H., 1885. Memory: a contribution to experimental psychology. Classics in the History of Psychology.

Ecuyer-Dab, I., Robert, M., 2007. The female advantage in object location memory according to the foraging hypothesis: a critical analysis. Hum. Nat. 18 (4), 365–385.

Engle, R.W., 2010. Role of working-memory capacity in cognitive control. Curr. Anthropol. 51 (SUPPL. 1).

Engle, R.W., Laughlin, J.E., Tuholski, S.W., Conway, A.R.A., 1999. Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. J. Exp. Psychol. Gen. 128 (3), 309–331.

Ericsson, K.A., Chase, W.G., Faloon, S., 1980. Acquisition of a memory skill. Science 208 (4448), 1181–1182.

Fagot, J., De Lillo, C., 2011. A comparative study of working memory: immediate serial spatial recall in baboons (Papio papio) and humans. Neuropsychologia 49 (14), 3870–3880.

Farrand, P., Parmentier, F.B.R., Jones, D.M., 2001. Temporal–spatial memory: retrieval of spatial information does not reduce recency. Acta Psychol. 106 (3), 285–301.

Fassihi, A., Akrami, A., Esmaeili, V., Diamond, M.E., 2014. Tactile perception and working memory in rats and humans. Proc. Natl. Acad. Sci. U. S. A 111 (6), 2331–2336.

Flynn, J.R., 2009. Are we getting smarter?: rising IQ in the twenty-first century. Are we Getting Smarter?: Rising IQ in the Twenty-First Century. Cambridge University Press.

Fonollosa, J., Neftci, E., Rabinovich, M., 2015. Learning of chunking sequences in cognition and behavior. PLoS Comput. Biol. 11 (11), e1004592.

Fountain, S.B., Annau, Z., 1984. Chunking, sorting, and rule-learning from serial patterns of brain-stimulation reward by rats. Anim. Learn. Behav. 12 (3), 265–274.

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

Gobet, F., Clarkson, G., 2004. Chunks in expert memory: evidence for the magical number four…or is it two? Memory 12 (6), 732–747.

GrÉGoire, J., Van Der Linden, M., 1997. Effect of age on forward and backward digit spans. Aging Neuropsychol. Cognit. 4 (2), 140–149.

Hahn, U., 2014. Experiential limitation in judgment and decision. Topic. Cogn. Sci. 6 (2), 229–244.

Haidle, M.N., 2010. Working-memory capacity and the evolution of modern cognitive potential: implications from animal and early human tool use. Curr. Anthropol. 51 (S1), S149–S166.

Hatano, G., Osawa, K., 1983. Digit memory of grand experts in abacus-derived mental calculation. Cognition 15 (1–3), 95–110.

Herman, L.M., 2010. What laboratory research has told us about dolphin cognition. Int. J. Comp. Psychol. 23 (3).

Hertwig, R., Barron, G., Weber, E.U., Erev, I., 2004. Decisions from experience and the effect of rare events in risky choice. Psychol. Sci. 15 (8), 534–539.

Hertwig, R., Pleskac, T.J., 2010. Decisions from experience: why small samples? Cognition 115 (2), 225–237.

Huang, L., Awh, E., 2018. Chunking in working memory via content-free labels. Sci. Rep. 8 (1), 1–10.

Huntley, J., Bor, D., Hampshire, A., Owen, A., Howard, R., 2011. Working memory task performance and chunking in early Alzheimer's disease. Br. J. Psychiatr. 198 (5), 398–403.

Inoue, S., Matsuzawa, T., 2009. Acquisition and memory of sequence order in young and adult chimpanzees (Pan troglodytes). Anim. Cognit. 12 (1), 59–69.

Jahnke, J.C., 1965. Primacy and recency effects in serial-position curves of immediate recall. J. Exp. Psychol. 70 (1), 130.

Jones, G., 2012. Why chunking should be considered as an explanation for developmental change before short-term memory capacity and processing speed. Front. Psychol. 3, 167.

Jonides, J., Lewis, R.L., Nee, D.E., Lustig, C.A., Berman, M.G., Moore, K.S., 2008. The mind and brain of short-term memory. Annu. Rev. Psychol. 59, 193–224.

Juslin, P., Olsson, H., 2005. Capacity limitations and the detection of correlations: comment on Kareev (2000). Psychol. Rev. 112 (1), 256–267.

Kail, R., Hall, L.K., 2001. Distinguishing short-term memory from working memory. Mem. Cognit. 29 (1), 1–9.

Karakaş, S., Yalin, A., Irak, M., Erzengin, Ö.U., 2002. Digit span changes from puberty to old age under different levels of education. Dev. Neuropsychol. 22 (2), 423–453.

Kareev, Y., 1995. Through a narrow window: working memory capacity and the detection of covariation. Cognition 56 (3), 263–269.

Kareev, Y., 2000. Seven (indeed, plus or minus two) and the detection of correlations. Psychol. Rev. 107 (2), 397–402.

Kareev, Y., 2005. And yet the small-sample effect does hold: reply to Juslin and Olsson (2005) and Anderson, doherty, Berg, and Friedrich (2005). Psychol. Rev. 112 (1), 280–285.

Kareev, Y., Lieberman, I., Lev, M., 1997. Through a narrow window: sample size and the perception of correlation. J. Exp. Psychol. Gen. 126 (3), 278–287.

Kawai, N., Matsuzawa, T., 2000. Numerical memory span in a chimpanzee. Nature 403 (6765), 39–40.

Klingberg, T., 2009. The Overflowing Brain: Information Overload and the Limits of Working Memory. Oxford University Press.

Leighton, J.P., Sternberg, R.J., 2003. Reasoning and problem solving. In: Handbook of Psychology. Wiley Online Library.

Lemasson, B.H., Anderson, J.J., Goodwin, R.A., 2009. Collective motion in animal groups from a neurobiological perspective: the adaptive benefits of dynamic sensory loads and selective attention. J. Theor. Biol. 261 (4), 501–510.

Lewis-Peacock, J.A., Drysdale, A.T., Oberauer, K., Postle, B.R., 2012. Neural evidence for a distinction between short-term memory and the focus of attention. J. Cognit. Neurosci. 24 (1), 61–79.

Lind, J., Enquist, M., Ghirlanda, S., 2015. Animal memory: a review of delayed matching-to-sample data. Behav. Process. 117, 52–58.

Logie, R.H., Gilhooly, K.J., Wynn, V., 1994. Counting on working memory in arithmetic problem solving. Mem. Cognit. 22 (4), 395–410.

Lombard, M., Wadley, L., 2016. Hunting technologies during the Howiesons Poort at Sibudu Cave: what they reveal about human cognition in KwaZulu-Natal, South Africa, between~ 65 and 62 ka. In: Multidisciplinary Approaches to the Study of Stone Age Weaponry. Springer, pp. 273–286.

Lynn, R., Irwing, P., 2008. Sex differences in mental arithmetic, digit span, and g defined as working memory capacity. Intelligence 36 (3), 226–235.

MacGregor, J.N., 1987. Short-term memory capacity: limitation or optimization? Psychol. Rev. 94 (1), 107–108.

Macuda, T., Roberts, W.A., 1995. Further evidence for hierarchical chunking in rat spatial memory. J. Exp. Psychol. Anim. Behav. Process. 21 (1), 20.

Manoochehri, M., 2020a. Sex differences in verbal forward digit span: a brief communication. Cogn. Brain Behav. Interdis. J.

Manoochehri, M., 2020b. The emergence of language: a review of ljiljana progovac (2019). In: A Critical Introduction to Language Evolution: Current Controversies and Future Prospects. Evolutionary Psychological Science.

Manoochehri, M., 2020c. The Evolutionary Starting Point; Musicality or Language. A Review of Henkjan Honing (2019); the Evolving Animal Orchestra: in Search of what Makes Us Musical. Evolutionary Psychological Science.

Marlowe, F.W., 2007. Hunting and gathering: the human sexual division of foraging labor. Cross Cult. Res. 41 (2), 170–195.

Martignon, L., Hoffrage, U., 2002. Fast, frugal, and fit: simple heuristics for paired comparison. Theor. Decis. 52 (1), 29–71.

Martín, M.E., Sasson, Y., Crivelli, L., Roldán Gerschovich, E., Campos, J.A., Calcagno, M.L., Allegri, R.F., 2013. Relevance of the serial position effect in the differential diagnosis of mild cognitive impairment, Alzheimer-type dementia, and normal ageing. Neurologia 28 (4), 219–225.

Martin, P.R., Fernberger, S.W., 1929. Improvement in memory span. Am. J. Psychol. 41 (1), 91–94.

Mathis, A., Unger, S., 2012. Learning to avoid dangerous habitat types by aquatic salamanders, Eurycea tynerensis. Ethology 118 (1), 57–62.

McCabe, D.P., Robertson, C.L., Smith, A.D., 2005. Age differences in stroop interference in working memory. J. Clin. Exp. Neuropsychol. 27 (5), 633–644.

McCauley, S.M., Christiansen, M.H., 2015. Individual differences in chunking ability predict on-line sentence processing. CogSci.

McCauley, S.M., Isbilen, E.S., Christiansen, M.H., 2017. Chunking ability shapes sentence processing at multiple levels of abstraction. In: Proceedings of the 39th Annual Conference of the Cognitive Science Society, pp. 2681–2688.

Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. 63 (2), 81–97.

Monaco, M., Costa, A., Caltagirone, C., Carlesimo, G.A., 2013. Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. Neurol. Sci. 34 (5), 749–754.

Mulcahy, N.J., Call, J., 2006. Apes save tools for future use. Science 312 (5776), 1038–1040.

Murray, E.A., Wise, S.P., Graham, K.S., 2017. The Evolution of Memory Systems: Ancestors, Anatomy, and Adaptations. Oxford University Press.

Naqshbandi, M., Roberts, W.A., 2006. Anticipation of future events in squirrel monkeys (Saimiri sciureus) and rats (Rattus norvegicus): tests of the Bischof-Kohler hypothesis. J. Comp. Psychol. 120 (4), 345.

Norris, D., 2017. Short-term memory and long-term memory are still different. Psychol. Bull. 143 (9), 992–1009.

Norris, D., Kalm, K., Hall, J., 2019. Chunking and redintegration in verbal short-term memory. J. Exp. Psychol. Learn. Mem. Cognit.

Nowell, A., 2010. Working memory and the speed of life. Curr. Anthropol. 51 (S1), S121–S133.

Oberauer, K., Kliegl, R., 2006. A formal model of capacity limits in working memory. J. Mem. Lang. 55 (4), 601–626.

Orsini, A., Chiacchio, L., Cinque, M., Cocchiaro, C., Schiappa, O., Grossi, D., 1986. Effects of age, education and sex on two tests of immediate memory: a study of normal subjects from 20 to 99 years of age. Percept. Mot. Skills 63 (2 II), 727–732.

Ostrosky-Solís, F., Lozano, A., 2006. Digit span: effect of education and culture. Int. J. Psychol. 41 (5), 333–341.

Osvath, M., 2009. Spontaneous planning for future stone throwing by a male chimpanzee. Curr. Biol.

Osvath, M., Karvonen, E., 2012. Spontaneous innovation for future deception in a male Chimpanzee. PloS One 7 (5).

Osvath, M., Osvath, H., 2008. Chimpanzee (Pan troglodytes) and orangutan (Pongo abelii) forethought: self-control and pre-experience in the face of future tool use. Anim. Cognit. 11 (4), 661–674.

Parker, E.S., Cahill, L., McGaugh, J.L., 2006. A case of unusual autobiographical remembering. Neurocase 12 (1), 35–49.

Peña-Casanova, J., Quiñones-Úbeda, S., Quintana-Aparicio, M., Aguilar, M., Badenes, D., Molinuevo, J.L., Blesa, R., 2009. Spanish multicenter normative studies (NEURONORMA project): norms for verbal Span, visuospatial Span, letter and number sequencing, trail making test, and symbol digit modalities test. Arch. Clin. Neuropsychol. 24 (4), 321–341.

Plonsky, O., Teodorescu, K., Erev, I., 2015. Reliance on small samples, the wavy recency effect, and similarity-based learning. Psychol. Rev. 122 (4), 621.

Progovac, L., 2019. A Critical Introduction to Language Evolution: Current Controversies and Future Prospects. Springer.

Putt, S.S., 2016. Human Brain Activity during Stone Tool Production: Tracing the Evolution of Cognition and Language. University of Iowa.

Rabinovich, M.I., Varona, P., Tristan, I., Afraimovich, V.S., 2014. Chunking dynamics: heteroclinics in mind. Front. Comput. Neurosci. 8, 22.

Raby, C.R., Alexis, D.M., Dickinson, A., Clayton, N.S., 2007. Planning for the future by western scrub-jays. Nature 445 (7130), 919.

Read, D.W., 2008. Working memory: a cognitive limit to non-human primate recursive thinking prior to hominid evolution. Evol. Psychol. 6 (4), 147470490800600.

Real, L.A., 1991. Animal choice behavior and the evolution of cognitive architecture. Science 253 (5023), 980–986.

Real, L.A., 1992. Information processing and the evolutionary ecology of cognitive architecture. Am. Nat. 140, S108–S145.

Reed, P., Chih-Ta, T., Aggleton, J.P., Rawlins, J.N.P., 1991. Primacy, recency, and the von Restorff effect in rats' nonspatial recognition memory. J. Exp. Psychol. Anim. Behav. Process. 17 (1), 36–44.

Roberts, W.A., 2012. Evidence for future cognition in animals. Learn. Motiv. 43 (4), 169–180.

Rosier, R.L., Langkilde, T., 2011. Behavior under risk: how animals avoid becoming dinner. Nature Edu. Know. 2 (8).

Rowe, A., Titterington, J., Holmes, J., Henry, L., Taggart, L., 2019. Interventions targeting working memory in 4–11 year olds within their everyday contexts: a systematic review. Dev. Rev. 52, 1–23.

Rowe, G., Hasher, L., Turcotte, J., 2008. Age differences in visuospatial working memory. Psychol. Aging 23 (1), 79–84.

Saaty, T.L., Ozdemir, M.S., 2003. Why the magic number seven plus or minus two. Math. Comput. Model. 38 (3–4), 233–244.

Sanz, C., Morgan, D., Gulick, S., 2004. New insights into chimpanzees, tools, and termites from the Congo basin. Am. Nat. 164 (5), 567–581.

Scarf, D., Smith, C.D., Jaswal, V.K., Magnuson, J.S., Terrace, H., 2018. Chunky Monkey? The spontaneous temporal chunking of simultaneous chains by Humans (Homo Sapiens) and rhesus monkeys (Macaca mulatta). Stud. Rhesus Monkey Behav.

Schwartze, M., Brown, R.M., Biau, E., Kotz, S.A., 2019. Timing the "magical number seven": presentation rate and regularity affect verbal working memory performance. Int. J. Psychol.

Sebastián, M.V., Hernández-Gil, L., 2012. Developmental pattern of digit span in Spanish population. Psicothema 24 (2), 183–187. Retrieved from. http://www.ncbi.nlm.nih.gov/pubmed/22420342.

Sebastián, M.V., Mediavilla, R., 2015. Verbal forward digit span in Spanish population. Psicothema 27 (2), 93–98.

Shen, H.H., 2004. Level of cognitive processing: effects on character learning among non-native learners of Chinese as a foreign language. Lang. Educ. 18 (2), 167–182.

Smeele, S.Q., Hansen, K.A., Ortiz, S.T., Johansson, F., Kristensen, J.H., Larsson, J., Wahlberg, M., 2019. Memory for own behaviour in pinnipeds. Anim. Cognit. 22 (6), 947–958.

Solopchuk, O., Alamia, A., Olivier, E., Zénon, A., 2016. Chunking improves symbolic sequence processing and relies on working memory gating mechanisms. Learn. Mem. 23 (3), 108–112.

Sugita, M., Yamada, K., Iguchi, N., Ichitani, Y., 2015. Hippocampal NMDA receptors are involved in rats' spontaneous object recognition only under high memory load condition. Brain Res. 1624, 370–379.

Sweller, J., 2003. Evolution of human cognitive architecture. Psychol. Learn. Motiv. 43, 215–266.

Taylor, A.H., Elliffe, D., Hunt, G.R., Gray, R.D., 2010. Complex cognition and behavioural innovation in New Caledonian crows. Proc. Biol. Sci. 277 (1694), 2637–2643.

Terrace, H.S., 1987. Chunking by a pigeon in a serial learning task. Nature 325 (6100), 149.

Terrace, H.S., 1993. The phylogeny and ontogeny of serial memory: list Learning by Pigeons and Monkeys. Psychol. Sci. 4 (3), 162–169.

Terrace, H.S., 2002. The comparative psychology of chunking. In: Animal Cognition and Sequential Behavior. Springer, pp. 23–55.

Thalmann, M., Souza, A.S., Oberauer, K., 2019a. How does chunking help working memory? J. Exp. Psychol. Learn. Mem. Cognit. 45 (1), 37.

Thalmann, M., Souza, A.S., Oberauer, K., 2019b. Revisiting the attentional demands of rehearsal in working-memory tasks. J. Mem. Lang. 105, 1–18.

Thuijsman, F., Peleg, B., Amitai, M., Shmida, A., 1995. Automata, matching and foraging behavior of bees. J. Theor. Biol. 175 (3), 305–316.

Todd, P.M., Hertwig, R., Hoffrage, U., 2005. Evolutionary cognitive psychology. In: The Handbook of Evolutionary Psychology. Wiley Online Library, pp. 776–802.

Tolan, G.A., Tehan, G., 2005. Is spoken duration a sufficient explanation of the word length effect? Memory 13 (3–4), 372–379.

Toyoshima, M., Yamada, K., Sugita, M., Ichitani, Y., 2018. Social enrichment improves social recognition memory in male rats. Anim. Cognit. 21 (3), 345–351.

Völter, C.J., Mundry, R., Call, J., Seed, A.M., 2019. Chimpanzees flexibly update working memory contents and show susceptibility to distraction in the self-ordered search task. Proc. R. Soc. B 286 (1907), 20190715.

Voyer, D., Postma, A., Brake, B., Imperato-McGinley, J., 2007. Gender differences in object location memory: a meta-analysis. Psychonomic Bull. Rev. 14 (1), 23–38.

Wagner, A.R., Pfautz, P.L., 1978. A bowed serial-position function in habituation of sequential stimuli. Anim. Learn. Behav. 6 (4), 395–400.

Wechsler, D., 1939. The Measurement and Appraisal of Adult Intelligence. TheWilliams &Wilkins Company, Baltimore, MD.

Welshon, R., 2010. Working memory, neuroanatomy, and archaeology. Curr. Anthropol. 51 (S1), S191–S199.

Whitford, M.D., Freymiller, G.A., Higham, T.E., Clark, R.W., 2019. Determinants of predation success: how to survive an attack from a rattlesnake. Funct. Ecol. 33 (6), 1099–1109.

Wright, A.A., 1989. Memory processing by pigeons, monkeys, and people. Psychol. Learn. Motiv. 24 (C), 25–70.

Wright, A.A., 2007. An experimental analysis of memory processing. J. Exp. Anal. Behav. 88 (3), 405–433.

Wright, A.A., Elmore, L.C., 2016. Pigeon visual short-term memory directly compared to primates. Behav. Process. 123, 84–89.

Wynn, T., Coolidge, F.L., 2006. The effect of enhanced working memory on language. J. Hum. Evol. 50 (2), 230–231.

Zheng, X., Swanson, H.L., Marcoulides, G.A., 2011. Working memory components as predictors of children's mathematical word problem solving. J. Exp. Child Psychol. 110 (4), 481–498.