Visuospatial Working Memory of Serial Order in Humans and Chimpanzees:
The Possible Influence of Language on Recall

Nadine Charanek

Lancaster University

Abstract

In recent years, most researchers have been interested in unravelling cognition which is perhaps the most challenging undertaking by humans to date; but studying our nonhuman counterparts may have made it easier especially in terms of working memory (WM) functions in relation to language. The present paper aims to investigate the possible influence of language as verbal coding on the recall of serial order in visuospatial working memory (VSWM) in humans compared to chimpanzees in previous research. Such was investigated in 4 experimental masking tasks: 2 verbal Arabic numerals tasks and 2 nonverbal symbols tasks. Both numerals and symbols tasks utilized a dual-task paradigm of an 8-digit sequence as verbal interference (VI) to investigate the influence of language as manifest in verbal coding. For all experiments, there were variations in latency times for which characters were shown before masked. Findings revealed that a) chimpanzees clearly outperform humans in the Arabic numerals task; b) human capacity is limited but is dependent on latency time, and c) language as manifested in verbal coding does have an influence but such is dependent on the mode of presentation of serial order. Finally, limitations in methodology are discussed followed by implications especially those pertaining to bilingual studies.

Keywords: visuospatial, working memory, chimpanzees, serial order, trade-off

1. Introduction

The ability to encode, process, and recall information is considered as an essential cognitive ability, be it in humans or nonhumans. When it comes to working memory (WM), several studies have been carried out that mostly involved the phonological loop or verbal working memory (VWM) – a WM subcomponent. However, fewer studies[11] have focused on the visuospatial WM (VSWM) (Baddeley, 2007) especially how verbal interference (VI) per se might influence the processing of sequentially presented spatial locations with most studies focusing on the visual component of WM (e.g. Larsen & Baddeley, 2003; Logi & Marchetti, 1991; Logie, Saito, Morita, Varma, & Norris, 2016) or the verbal (e.g. Jones, Macken, & Nicholls, 2004; Jones, Hughes, & Macken, 2006; Nelson, Brooks, & Borden, 1973) with less focus on the spatial (e.g. Bouma, 1987; Chuah & Maybery, 1999; Depoorter & Vandierendonck, 2009; Jalbert, Saint-Aubin, & Tremblay, 2008; Dent & Smyth, 2005). So, the question remains then as to how does VI influence spatial recall of serial order?

1.1 Visuospatial working memory of humans

VSWM is a subcomponent of WM and it is thought to be responsible for the manipulation of visual and spatial information ones (for a full review of the WM multicomponent model see Baddeley, 1986, 2007; Cornoldi & Vecchi, 2003; Logie, 1995). When it comes to combining the study of Baddeley's (2007) WM components and serial order, several theories have risen (See Baddeley, 2007 for an extensive review). An example of which is the associative theory which claims that we store information by making associations between the beginning and the

---

[11] Most studies focused on which type of interference (verbal, spatial, or visual) influences recall of visuospatial information the most (e.g. Deyzac, Logie, & Denis, 2006; Postle, D'Esposito, & Corkin, 2005; Zhao, Chen, & West, 2010; Zimmer, Speiser, & Seidler, 2003).

end of a sequence (see Shiffrin & Cook, 1978). It was further explained through the primacy model by Page and Norris (1998) where items recalled sequentially are always associated with the first item but with decreasing strength hence errors in recall arise due to decay, but this may be reactivated through rehearsal. In other words, recall of serial order according to this model is supposed to be enhanced by verbal rehearsal. Furthermore, any potential interference makes these items vulnerable to trace decay. Accordingly, verbal interference is likely to negatively influence recall of serial order especially when it is phonologically similar to the recalled items (Baddeley, 2007, Baddeley, Eysenck, & Anderson, 2015; Logie, 1995).

Based on these serial order theories, it would be interesting to investigate if they also apply for VSWM serial recall as influenced by verbal rehearsal mechanisms and this is because most previous research has focused on tasks involving one WM subsystem, but few studies have explored the possible interface between VWM and VSWM in serial order processing. One way to explore this would be through a dual-task paradigm where subjects perform a primary memory task and a secondary memory task simultaneously. If the two tasks utilize the same limited cognitive resources, then a dual-task methodology will show weaker performance on the primary task than when this task is performed separately. Previous researchers have utilized this framework using verbal interference such as articulatory suppression (e.g. verbal repetition of digit sequence) as the secondary task while engaged in a primary memory task. Of interest for the present study are tasks that involve remembering "where" and remembering "what". For instance, several studies found no significant effects of retaining a visual pattern when encoding and recalling a digit sequence at the same time (Cocchini et al., 2002) nor for spatial patterns in a syntactic category task (Postle, D'esposito, Corkin, 2005). This could be interpreted as the possible separability of operations of the VSWM and the VWM at least for the tasks provided. However, a more positive correlation between the two systems was found

where engagement in a verbal reasoning task concurrently with serial digit recall showed a disruption in performance on both tasks when processing became increasingly difficult (e.g. Baddeley et al., 1984; Posner & Konick, 1966). This finding could signal the presence of rehearsal that was disrupted by digit recall i.e. rehearsal is necessary for serial order[12].

Nonetheless, Nelson et al. (1973) found that in several variations of a word to picture trials, subjects performed better in picture-picture trials where verbal codes might be less present than in a word task. The absence of a verbal code has perhaps improved recall of the serial order. The authors explained this in terms of the nonverbal system being capable of coding sequential information as well as the verbal one. But a cognitive trade-off hypothesis may be a plausible explanation where a supposed negative influence or pressure from language might be impeding recall for serial order.

Despite the supposed negative influence of language on recall for serial order, most research on WM viewed language more positively. For instance, in investigating limitations in capacity, researchers varied between verbal and nonverbal stimuli based on the assumption that verbal rehearsal is beneficial for recall. Several studies on information coding revealed that our STM might be limited in how many features we can remember for items (Baddeley, 2007; Cowan, 2010) and retention of such information may be influenced by verbal rehearsal. For example, it has been shown that one can remember quite easily both the colour and shape of a

---

[12] Such importance of the verbal system for visuospatial serial order has been shown in several studies. For instance, in a series of experiments Nelson et al. (1973) have shown that learning the serial order of pictures involves the verbal system where subjects reported using verbal codes to help recall the serial order of the pictures. The authors explained this in light of Paivio's (1971) dual coding hypothesis. According to Paivio, there are 2 independent processing systems: one specialized for imaginal information and the other for coding verbal information. The visual system is reserved for coding simultaneously presented information while the verbal is involved in coding sequential or serial information (Paivio, 2014). As such, according to Paivio's hypothesis, verbal coding is supposed to aid in recall.

letter after it has been briefly presented then masked by an asterisk (e.g. Irwin & Andrews, 1996). Nonetheless, such finding cannot completely rule out the influence of verbal rehearsal because letters can be verbally coded. Hence there was a need to use abstract symbols that are not easily recalled through verbal rehearsal in any similar study and research saw the use of more nonverbal stimuli such as wallpaper patterns (e.g. Broadbent & Broadbent, 1981), faces (e.g. Smyth et al., 2005), and bars (e.g. Luck & Vogel, 1997; Corsi, 1972). The use of these abstract stimuli, however, does not completely rule out the absence of verbal coding (Logie et al., 2000). It seems logical then that occupying the phonological loop is needed to test if an increase in capacity would occur, but findings suggest even in the presence of articulatory suppression and rapid presentation of characters, no changes in performance occur at least for object memory (e.g. Vogel, Woodman, & Luck, 2001). But, will the same pattern emerge for spatial memory in recall of serial order? The present study aims to investigate this by bringing together memory for what (visual STM), memory for where (spatial STM), and a possible how (verbal STM).

The investigation of serial order is not unique to humans, but it further extends to the field of comparative psychology. Numerous studies have been done on humans and nonhumans alike to better understand the underlying mechanisms associated with WM and the encoding of serial order (see Smith & Beran, 2017 for a review). The studies on nonhuman intelligence are particularly significant because they have offered us direct information on what cognitive processes are like in the absence of spoken language and aspects of our own cognition that are dependent on or influenced by language (Fagot & Barbet, 2006; Wasserman & Zentall, 2006).

1.2 A chimp's cognition: extending the frontiers to human intelligence

Comparative studies of cognition that are of concern to us are those pertaining to our supposed ancestors: apes. Stemming from Darwin's Evolution Theory, and based on other research, humans and chimpanzees appear to share some kind of kinship as revealed by their genetic makeup (Seyfarth, Cheney, & Bergman, 2005; Tomasello & Herrmann, 2010; Wasserman & Zentall, 2006). Chimpanzees, for instance, were found capable of finding hidden objects as well as human children (Tomasello & Call (1997) but their spatial retention decays with time (most likely faster than in humans) perhaps because chimpanzees live in the here and now and do not need to recall the past (Matsuzawa, 2012; Tomasello & Herrmann, 2010). But questions remained regarding their faster and more accurate performance than humans, and to further understand such species difference in visuospatial retention, eye-tracking studies have been carried out on children and chimpanzees (See Kano & Tomonaga, 2009, 2011). The difference in the timing of eye movement was attributed to ongoing cognitive processes (e.g. Findlay & Walker, 1999) where humans engage in prolonged fixations in semantically demanding activities such as reading (Clifton et al., 2016; Kano & Tomonaga, 2011) which therefore limits our amount of scanning of the visual field. In other words, chimpanzees may be faster in their scanning than humans because, to the best of our knowledge, they do not have such semantic processing as revealed by studies on both species (Dent & Smyth, 2005; Ginsburg, Archambeau, van Dijck, Chetail, & Gevers, 2017; Inoue & Matsuzawa, 2009; Matsuzawa, 2009; Washburn, Gulledge, James, & Rumbaugh, 2007). Such a lack of semantic processing in chimpanzees might be rather advantageous for WM functions.

1.3 Visuospatial working memory of chimpanzees

As explained thus far, the representation of language in chimpanzees is thought to mirror the representation of language in humans especially in terms of working memory. Several studies have shown that apes tend to engage in internal cognitive processes much like humans when it

comes to reconstructing old and novel patterns (Vauclair, 1994). Such ability was also extended to numerical skills as shown by Boysen and Berntson (1989). To explain how a chimpanzee develops numerical skills, and whether this is representative of the development seen in humans, Boysen & Berntson theorized that chimpanzees possess the capacity to engage in ordering numerical skills and not only subitizing them i.e. they can both recognize numbers and count them ordinally and cardinally[13]. The chimpanzees in their study were able to count novel objects using the Arabic numerals and both their performance and development reflected that observed in human children's numerical skills. In other words, the chimpanzees showed aspects of not only recognition memory but also reproductive one (e.g. Vauclair, 1994). Based on these findings, it was suggested that more research was required on studying the capacity for numerical competence in chimpanzees as it can tell us about the cognitive structure in apes, humans' phylogenetic origins and cognitive evolution, and the ontogeny of number concepts in children (Boysen & Berntson, 1989; Fagot & Barbet, 2006).

1.4 The Cognitive Trade-Off hypothesis

One of the most notable researchers to investigate and expand on both numerical competence and reproductive memory in chimpanzees, and whose theory is the foundational argument in this paper, is primatologist Tetsuro Matsuzawa of Kyoto University in Japan. His decades of research on our supposed nonhuman relatives have put forth revolutionary ideas in terms of WM and language through what he terms the Cognitive Trade-off hypothesis, the theoretical basis for the present study. Put simply, Matsuzawa believes that in the process of evolution,

---

[13] The chimpanzees were trained on counting food items first then these were replaced by the Arabic numerals on glass placards utilizing the numbers 0-5. Then, they were trained on the order of counting from 0-5 and were no longer limited to recognizing the numbers as individual items.

humans had to sacrifice parts of their cognition to make room for even more complex processes and perhaps the most noteworthy is language (Inoue & Matsuzawa, 2007, 2009; Matsuzawa, 2009, 2012, 2013).

Such trade-off notion came from a series of experiments on well-trained chimpanzees at a semi-natural primate habitat at Kyoto University. Matsuzawa and his colleague Inoue (2009) first discovered the higher WM capacity of chimpanzees in a serial order digit span task of 9 Arabic numerals accuracy (See Inoue & Matsuzawa, 2009 for more on these experiments) which later extended into a limited-hold masking task.

In the masking task, the numerals on the screen were immediately covered with white squares when the subject touched the smallest numeral and their positions were to be recalled from memory. Overall, the response time to responding to the first numeral was longer than the others. This led the researchers to believe that perhaps the chimpanzees analyze the visual scene as one whole before touching the screen i.e. they form a cognitive map of space (Tomasello & Call, 1997). Furthermore, young chimpanzees were found to be faster in their judgement than their adult counterparts. Similar findings are generally reported in studies on humans where children perform better than adults on certain memory tests (See Hayes & Heit, 2004; Sloutsky & Fisher, 2004). Matsuzawa and Inoue attributed such findings to their trade-off understanding wherein the process of acquiring linguistic skills children gradually lose their perceptual photographic memory abilities (or eidetic memory). Such may also be why the young chimpanzees perform better than the adult ones (Inoue & Matsuzawa, 2009).

To better understand the working memory functions of chimpanzees, and better yet humans, Inoue and Matsuzawa (2007) extended their masking task to 9 university students. The students overall performed worse (40% accuracy) than the three young chimpanzees (80%

accuracy). To further investigate such phenomenon, the masking task was used. Performance of humans was compared to that of the best mother performer Ai, and the best young performer, her son, Ayumu (then aged 5.5). Ai's accuracy and that of the human adults declined as the hold duration decreased while Ayumu continued to perform with almost perfect accuracy regardless of the timings. It was speculated that age was a contributing factor to the decline of working memory in both species considering that Ai was, in her prime years, the best in such task just as Ayumu is now. Furthermore, differences in performance between the two were thought to be due to Ai being language trained while Ayumu was not thus aligning with Matsuzawa's hypothesis (Roberts & Quillinan, 2014). As such, while age seems to explain the obvious lapse in memory within-species, it is not enough, however, to explain why chimpanzees are a step ahead than humans in their visuospatial working memory, at least in the task at hand making the trade-off explanation more reasonable. Inoue and Matsuzawa speculate that this trade-off may have exchanged our eidetic memory- which may be the underlying mechanism for extraordinary chimpanzee memory- in favour of language (See section 1.4). Furthermore, such eidetic memory has been shown to be relatively better in children and to decline with age (Giray, Altkin, Vaught, & Roodin, 1976; Haber & Haber, 1988) which might explain why Ai and the adult chimpanzees performed worse than the younger ones.

The above studies give us a glance of why the trade-off notion was coined: to better understand one of the main differences between humans and animals – language. By language here I mean making semantic associations (or symbolic representations) which are thought to impose semantic pressure on our cognition causing us to give up our instant eidetic memory for something more complex and perhaps more essential to us: language (Inoue & Matsuzawa, 2007, 2009; Matsuzawa, 2009). But, is language really the only reason why our immediate

memory is slower than that of these magnificent apes? Of course, any person might argue for the obvious reason: biological makeup owing to species difference. For instance, eye movement may be what influences speed of visuospatial scan, however, such argument is not valid for several reasons. First, several studies have shown that the observed variation in gaze timing is influenced by cognitive processes taking place (e.g. Hollingworth, Richard, & Luck, 2008) and that gaze timing increases as the visual scene becomes more demanding such as in reading (Clifton et al., 2016). Second, the current study does not violate the biological turf of the human body as the lowest speed of 210ms is the least frequency for the human saccadic eye to move from central fixation to peripheral (Bartz, 1962) and such eye movement has been shown to be directed by visual short-term memory (Hollingworth, Richard, & Luck, 2008) and not the reverse. Third, chimpanzees have been discovered to make longer fixations per second than humans when shown several patterns; this might be interpreted as different attentional focus for each species, but some have found that our attentional shifts do not influence visuospatial performance (e.g. Johnson, Hollingworth, & Luck, 2008) suggesting that perhaps each species adheres to certain visual scanning strategies independent of biological makeup (Kano & Tomonaga, 2009, 2011; Myowa-Yamakoshi, Scola, & Hirata, 2012).

Another criticism for semantic pressure came from replication studies. For instance, one study showed that training on the masking task can be a factor (Sielberg & Kearns, 2009), but results revealed that despite training humans could not surpass accuracy for 5 numerals, while Ayumu is reported to have 80% accuracy for 8 numerals unaffected by the limited-hold durations (Matsuzawa, 2009). This suggests that, for the time being, humans cannot outperform a chimpanzees' performance even with training (Matsuzawa, 2013). Such attestation is further highlighted in a recent replication of Inoue & Matsuzawa's study by Roberts & Quillinan (2014)

where despite humans performing better than those in the original study, they still performed significantly worse than Ayumu.[14]

So, the question remains then: what is causing such difference in performance on the VSWM masking task between chimpanzees and humans? Based on the reviewed literature, some have claimed that the competitive nature of chimpanzees and their lack of language interference makes them process a visual field much faster than humans who process everything in a social context i.e. we seem to make meaning out of everything we see (Ginsburg et al., 2017; Matsuzawa, 2013). As such, perhaps what is uniquely ours cognitively – language – may be what is affecting our visual scanning strategies and subsequent perceptual analysis (Kano & Tomonaga, 2009, 2011; Matsuzawa, 2013).

1.5 Aim of the study

From the cognitive trade-off perspective then, any disruption in performance because of articulatory suppression would suggest that a binding mechanism between the verbal and the spatial codes is taking place and that the VWM and VSWM, though separable systems (Baddeley, 2007), are dependent on one another. However, if there are no significant differences across performances, then this might be more of a reflection of their separability as maintained by the multicomponent WM model (Baddeley, 2007). The aim of this study is to investigate the feasibility of an interface between the VSWM and the VWM, and the possible

---

[14] Some studies on why chimpanzees outperform humans in pattern recognition and detailed perception (see Martin, Bhui, Bossaerts, Matsuzawa, & Camerer, 2014 for details) revealed that a chimp's memory for patterns is better than a human's and that a cognitive advantage does somehow exist. Such advantage is attributed to chimpanzees' competitive lifestyle where strategic thinking and pattern detection are necessary to detect enemies. And, unlike humans, chimpanzees have been shown to be better at competitive games than cooperative ones (e.g. Hare & Tomasello, 2004; Yamamoto & Tanaka, 2009; Martin et al., 2014; Matsuzawa, 2013; Tomasello & Herrmann, 2010).

influence on the processing of serial order and location which will, in turn, provide an understanding of the WM functions and limitations. This will be investigated through a dual-task performance using articulatory suppression along with a verbal task or a nonverbal one. Such focus is, to the best of my knowledge, unlike previous studies where although studies have used articulatory suppression in the form of digit recall, but such was limited to testing the capacity for feature retention (colour, shape, size…) of visually presented information and not the capacity for the serial location of patterns. Furthermore, no research as far as I know has attempted to study the direct influence of verbal rehearsal on VSWM compared to chimpanzees. Thus, studies on the role of VI on VSWM show a gap in memory research for the relationship between serial memory for spatial locations and verbal rehearsal embedded in comparative psychology.

Based on the cognitive trade-off hypothesis, it is predicted that participants will perform worse than chimpanzees, but such performance is better in the symbols task than in the numerals task. Furthermore, participants are likely to perform better in the presence of VI because occupying the phonological loop reduces the chances of using verbal rehearsal for recall and according to Matsuzawa's hypothesis, such rehearsal reduces WM capacity.

2. Methodology

2.1 Participants

Participants were 16 English British monolingual students at Lancaster University in the UK aged 18-26 (mean age M= 21.56). There were (7) males and (9) females. They were recruited via social media platforms and email.

2.2 Materials and Procedure

The experiment was an online game adapted from Roberts and Quillinan (2014). Responses were collected utilizing a touch screen design to simulate the original study's set up, but 8-digit sequences were used as VI and chosen randomly for memorization (See Appendix, Table 1).

Because the present study focuses on language as a possible influence on visuospatial working memory, the adaptation of the experiment from Roberts and Quillinan (2014) comprised two sets of tasks: one task that denotes verbal/semantic coding (numerals) and an abstract symbols task that is not easily verbalizable (Figure 1) (Baddeley, 2007; Logie, 1995).[15]
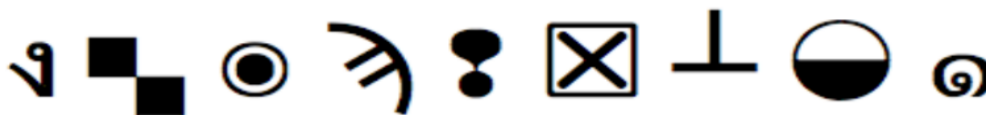
Figure 1. showing the 9 abstract symbols used in the abstract symbol tasks.

Participants had to engage in two sets of tasks: a numerals task and a symbols task each with two variations (one control and one with VI). Each set of tasks completed for each variation included three modes:

a)   An arcade mode that starts with one character then gradually increases the number of characters one by one to all 9. The first character is presented at a longer latency time of about 1300s then the rest are presented at gradually faster speeds until 210ms.

---

[15] These specific symbols are adapted from Roberts & Quillinan (2014). It has been argued that even the use of abstract symbols does not necessarily rule out verbal coding processes and such, as Roberts and Quillinan (2014) note, may be a limitation and a reason why no improvements were seen in the abstract symbols task. To further limit the role of verbal coding then, participants are required to occupy the verbal rehearsal system by continuously repeating a simple sequence (e.g. Luck & Vogel, 1997) i.e. engaging in articulatory suppression to prevent verbal recoding of visuospatial information (Baddeley, 2007; Baddeley, Eysenck, & Anderson, 2015; Dent & Smyth, 2005) and the present study utilized 8-digit sequences. The choice for serially ordered numerals as VI comes from the need to activate the same working memory systems (Baddeley, 2007; Baddeley, Eysenck, & Anderson, 2015) hence limiting any unsolicited influence of other cognitive systems that is not of focus in the present study.

b)   A challenge mode that presents all 9 characters simultaneously and gradually decreases

latency time from 3s to 210ms.

c)   A chimp mode that starts with 5 characters up to 9. All characters are presented for

210ms.

Participants were instructed to recall the order and location of 9 characters presented in

4 tasks: numerals (Task 1), numerals with VI (Task 2), symbols (Task 3), and symbols with VI

(Task 4). After a certain time, the characters were masked by white squares which subjects had

to tap on to indicate the recalled order and location of the numerals (See Figure 2).
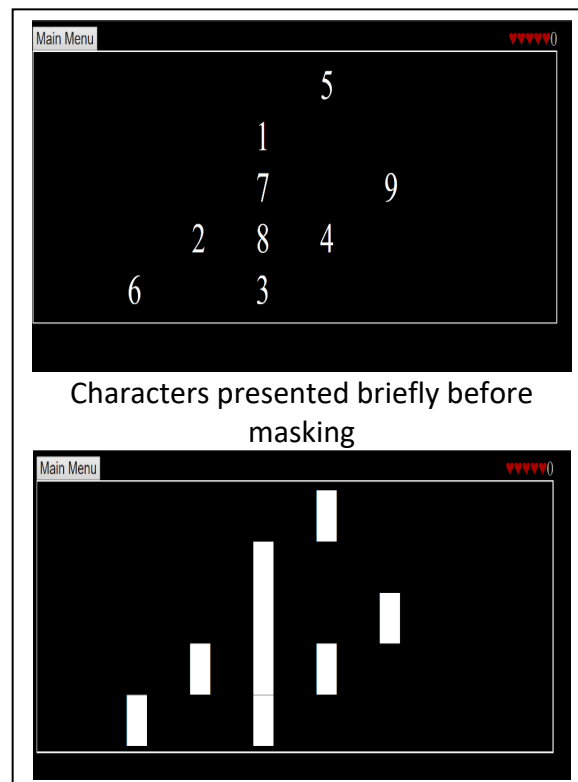
Figure 2. showing a preview of the limited-hold masking task as viewed by participants before masking and after.

All the tasks followed the same sequence and procedure except in tasks 3 and 4 where VI was

introduced. In these two tasks, participants had to memorize and recite the order of an 8-digit

sequence while performing each task. The sequence was changed at the end of each round (i.e.

after all lives were lost and the game was over). Participants had to recall the sequence after the end of the round. Failure to recall the sequence at the end suggests that verbal coding may have occurred for the characters being shown.

2.3 Data Analysis

A quantitative analysis was done per subject with calculations of recall accuracy for all 4 conditions: 2 controls, 2 variables.[16] Within each condition, detailed accuracy percentage was calculated per mode: arcade, challenge, and chimp. All these conditions with each mode total to 12 possible variations. Percentage accuracy was calculated as follows:

$$\% \text{ Accuracy } (n) = \frac{correct\ number\ of\ clicks\ per\ round}{maximum\ number\ of\ clicks\ per\ round} \times 100$$

This sole percentage accuracy is part of an aggregated mean accuracy calculated for each attempted numerical capacity (n) by counting the number of trials for that numeral and the average accuracy on each of these trials as follows:

$$\% \text{ Mean Accuracy } = \frac{\Sigma\ \%\ Accuracy\ (n)}{number\ of\ trials\ attempted\ (n)}$$

This % Mean Accuracy is taken as a reflection of WM capacity in this study where total percentage accuracy calculations were made for each of the 12 condition variations by calculating the total average for all 16 participants. The conditions are divided per task type (numerals or symbols), mode type (arcade, challenge, or chimp), and variable type (VI or no VI).

3. Results

---

[16] The control conditions refer to the numerals task and the symbols task without VI. The variable conditions refer to the numerals task and the symbols task with VI.

A summary of the obtained percentage mean accuracies depending on mode type and task type is shown in Figure 3 along with I&M's in Figure 4. It appears that three main patterns emerge from our results. The first pattern which has been shown throughout the 4 experiments concerns performance being dependent on how many numerals are presented and for how long i.e. performance appeared to be a ramification of numerals capacity and latency time (both of which are dependent on mode type). More specifically, the arcade mode was consistently the easiest for participants as revealed by their relatively higher accuracies than the other two modes whereas the challenge mode was the most difficult with participants scoring even lower than the chimp mode.
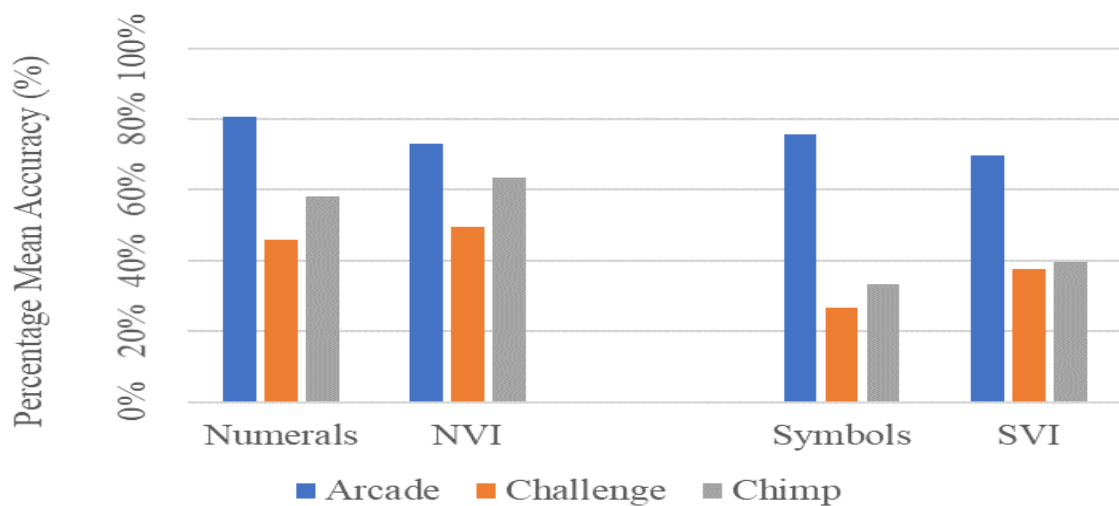
Figure 3. *Summary of percentage mean accuracies according to mode type and task type.*
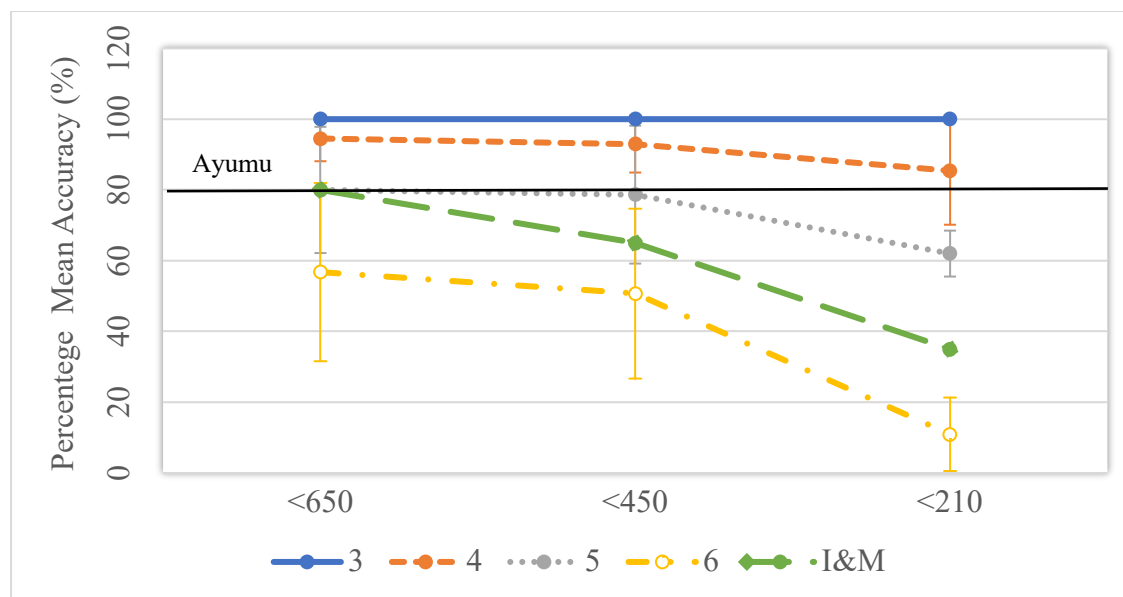
Figure 4. Percentage Mean Accuracy for 6 numerals compared to Ayumu and the humans from the original I&M study at 3 different latency times.

The second pattern concerns the influence of VI on accuracy and instead of having a positive influence on performance for all modes as was predicted, a negative influence has been observed depending on mode type in each experiment and is highlighted again in Figure 3. This negative influence was limited to the arcade mode whereas the challenge and chimp modes have improved with the administering of VI. The third pattern observed concerns the numerals task being easier for participants than the symbols task and such is reflected by the obvious higher accuracies for numerals than symbols across all modes (See Figure 3).

4. General Discussion

4.1 The limits of human visuospatial memory compared to chimps

Part of this study aimed to investigate how participants performed compared to the best chimpanzee performer Ayumu and I&M's participants (Inoue & Matsuzawa, 2007, 2009). Results in experiment 1 revealed that although participants performed better than I&M's, they could not perform up to Ayumu's level at 210ms. Performance up to 5 numerals was not

affected (unlike in I&M's) at 650ms or 450ms until latency time decreased to 210ms, which is unlike Ayumu who maintained near-perfect accuracy across all latencies and for all 9 numerals. The participants' accuracy, however, decreased significantly when 6 numerals were presented, and performance was worse at all 3 latency times and was no longer at Ayumu's level. This suggests that as humans, we are limited in our hold capacity to 5 numerals and that we need more time to process these numerals optimally, preferably more than 210ms. This finding is in line with previous research where capacity has been shown to be limited to 3-5 objects regardless of whether each object has different features like colour, shape, and size (Baddeley, 2007; Cowan, 2010; Irwin & Andrews, 1996; Luck & Vogel, 1997; Vogel et al., 2001). Furthermore, previous studies have shown that chimpanzees engage in a preplanning-at-a-glance strategy for coding visuospatial information while humans seem to take more time to do so due to the involvement of semantic processing through verbal coding (Clifton et al., 2016; Dent & Smyth, 2005; Ginsburg et al., 2017; Matsuzawa, 2009) and as previously revealed by eye-tracking research (Kano & Tomonaga, 2009, 2011). Indeed, humans tend to make use of linguistic coding as mnemonic support as we have particularly seen for the nonverbal symbols (Washburn, Gulledge, James, & Rumbaugh, 2007) and while this helps in retaining information for longer periods due to forged semantic associations, it limits the amount of information retained.

As such, our WM capacity appears to be limited especially for visuospatial information due to semantic processing, and according to Cowan (2010), this means that "superior beings from another planet can accomplish feats that we cannot because they have a larger working memory limit, similar to our digital computers …" (p. 56) and this is perhaps because these beings, like Ayumu, do not possess language as mandated by the cognitive trade-off hypothesis giving them an edge in making rapid perceptual judgements of the visual and spatial.

4.2 Language as a negative influence on VSWM capacity

Tasks 2, 3, & 4 have investigated the supposed negative influence of language as seen in previous studies (Inuoe & Matsuzawa, 2007; Matsuzawa, 2011, 2012, 2013) with tasks 3&4 implementing a VI task. Taken together with the previously mentioned WM capacity research, if a language is removed, be it by using a nonverbal stimulus or engaging in articulatory suppression, more storage capacity would be freed to process and retain information. An increase in capacity has been reflected in the present study through an increase in accuracy percentages in the presence of VI (though only for specific modes) but not when a nonverbal task was implemented alone.

Capacity seemed to be limited to 3-5 objects in experiment 2, much like what was found in experiment 1. While it is obvious that for experiment 1 the overt verbal rehearsal of the semantic Arabic numerals might be what adversely influenced accuracy if viewed from the trade-off notion, the expectation was that accuracy would improve with a nonverbal symbols task. However, contrary to the initial prediction of improved performance, results showed a much more negative performance. This would mean that the trade-off notion is rejected. But, based on the strategies utilized to complete the symbols task, it appears that the trade-off notion is made even stronger. Indeed, almost all participants engaged in overt verbal coding of the abstract symbols which conforms with the previous research on chimpanzees and humans (Roberts & Quillinan, 2014) where the use of abstract symbols does not necessarily inhibit the use of verbal rehearsal (Baddeley, 2007; Logie et al., 2000) and such has been found to be the main strategy for nonverbal visuospatial memory where linguistic coding is used as a main retention strategy by humans and language-trained chimpanzees as opposed to other primates (see Washburn, Gulledge, James, & Rumbaugh, 2007). For instance, one participant attempted to name the symbols such as "v" for ᐯ and "squares" for ▰. Another participant used "two

circles" for ◉ and "bow" for ﮄ. These examples highlight how as humans we tend to interact semantically with visual information. In other words, it seems as though we are incapable of processing foreign information unless it is semantically coded (Bouma, 1987; Ginsburg et al., 2017). Such coding of the symbols then may have resulted in a negative effect on performance as making meaning out of the visual field seems to occupy too much time as is seen in several visual scanning activities and contrary to chimpanzees' rapid all-at-once scanning strategies (Clifton et al., 2016; Kano & Tomonaga, 2009, 2011; Myowa-Yamakoshi, Scola, & Hirata, 2012). Furthermore, familiarity with the Arabic numerals, as opposed to the abstract symbols, cannot be neglected as humans and language-trained chimpanzees have been shown to perform better when the symbols are linguistically meaningful as opposed to completely new ones that needed to be semantically coded before processing and retention (Washburn, Gulledge, James, & Rumbaugh, 2007).

Regardless of how information was presented, and unlike previous studies on articulatory suppression and VI, an overall increase in accuracy was observed in the presence of VI which confirms the trade-off notion and further explains why WM capacity might be limited at least in the present study. Furthermore, some participants have reported that the presence of VI has allowed them to engage in the hold task unconsciously which may suggest that shifting serial focus to a secondary verbal task has suspended conscious rehearsal of the primary task (which is the premise of the notion of articulatory suppression e.g. Baddeley, 2007; Baddeley, Eysenck, & Anderson, 2015; Cornoldi & Vecchi, 2003; Logie, 1995) but has actually caused accuracy to increase (and not remain unchanged despite attentional shift as in Johnson, Hollingworth, & Luck, 2008). Participants further mentioned that the VI has provided a rhythm for them to tap on character locations be it in the symbols task or numerals task and indeed when participants made errors in the to-be-recalled sequence, performance broke down. However, I argue that it

is not a rhythm that the VI has provided but as mentioned it has caused a temporary suspension of verbal rehearsal of the digits and indeed this is revealed by the mistakes made in reciting the VI. For instance, participants often started a VI sequence correctly but halfway through a trial, they started citing the digits presented on the screen causing performance to break down when attempting to shift back to the VI sequence again. This could mean that engaging in overt articulatory suppression in a secondary task can improve the visuospatial capacity of the primary task i.e. when language is suppressed, visuospatial memory improves which means for a brief moment in time we may have gained the visuospatial cognitive abilities of chimpanzees or eidetic memory due to the absence of language (Matsuzawa, 2009, 2011, 2012, 2013) but this seems to be dependent on the pattern of serial order presentation.

4.3 Patterns of serial order presentation and influence on VSWM capacity

The present study revealed an unexpected pattern regarding mode of presentation of characters. The arcade mode, in which characters are presented sequentially with decreasing latency time, has been found to be the easiest mode but negatively influenced by VI. On the other hand, the chimp mode (which presents 5 characters simultaneously at 210ms) and challenge mode (which presents all 9 characters simultaneously with decreasing latency time) have both been positively influenced by VI and harder than the arcade mode, with the challenge mode being the hardest. These results suggest that the pattern of presentation of serial order could be a factor influencing VSWM capacity.

The fact that our participants consistently scored better in the arcade mode implies that as humans we process information better when it is presented in chunks and not all at once. Such finding is consistent with previous encoding theories of visuospatial information pertaining to the primacy model (Baddeley, 2007; Logie et al., 2016). According to the primacy

model, participants were able to recall items sequentially by associating them with the first item but with decreasing strength hence errors in recall have risen due to decay. Such decay appears to be counteracted through rehearsal where performance was better in the absence of VI, but its presence as predicted by the primacy model has made the characters (numerals and symbols alike) vulnerable to trace decay. This decay is more pronounced when VI is phonologically similar to the recalled items (Baddeley, 2007, Baddeley, Eysenck, & Anderson, 2015; Logie, 1995) which were observed in a slightly more negative influence for the numerals task than the symbols task where VI- which is also a numerical sequence- decreased accuracy by about 3% more in the numerals task (See Figure 3). Such phonological similarity effect has been noted before (Baddeley, 2007, Baddeley, Eysenck, & Anderson, 2015) and being a language component further asserts the negative influence of language on recall.

Another unexpected finding was that the challenge mode, and not the chimp mode, was consistently the hardest for participants. These low scores observed in the challenge mode – which presents all 9 characters simultaneously as opposed to 5 numerals only like in the chimp mode- suggests limitations in capacity for a number of objects and not only their features which is something that previous studies did not fully investigate (Baddeley, 2007; Cowan, 2010). Much like our capacity for 3-5 features (Baddeley, 2007; Cowan, 2010; Luck & Vogel, 1997; Vogel et al., 2001), spatial capacity for object locations seems to be limited to 3-5 characters at least when characters are presented all at once. This capacity seems to be slightly better when characters are presented sequentially suggesting the utilization of different strategies to encode and recall visuospatial information (Kano & Tomonaga, 2009, 2011; Matsuzawa, 2013; Myowa-Yamakoshi, Scola, & Hirata, 2012). For instance, some participants reported the use of hand gestures to draw a line pattern between characters during the encoding process and try to replicate this pattern in the recall process while at the same time verbally coding the

characters i.e. they attempted to make semantic associations through both a visual code and a phonological code and both of these codes have been previously shown to support immediate visuospatial recall of serial order (Dent & Smyth, 2005; Logie et al., 2000, 2016; Washburn, Gulledge, James, & Rumbaugh, 2007). In other words, the two WM subcomponents do not appear to be as separable as is posed by most WM models but on the contrary, there is crosstalk happening between them whether in humans or language-trained chimpanzees (Baddeley, 2007; Washburn, Gulledge, James, & Rumbaugh, 2007).

5. Conclusion

What this study has shown, relative to the small sample size, is that language as verbal coding, can influence VSWM capacity in different ways. However, what these findings reveal is that the long argument surrounding the separability of WM subcomponents, VSWM and VWM, might be neglecting their interaction in operations as simple as recall of serial order. Furthermore, while we might think that language makes us superior, we must be reminded that other creatures out there may do things we cannot because they do not have language, a skill we fully labelled our own. In addition, while language might be disadvantageous for certain tasks, it is rather invaluable for others especially when we depend on our linguistic capabilities to create new things, share ideas, and imagine our future. Such language advantages are long-held by bilingualism and investigating the controversy of the bilingual advantage relative to the trade-off notion studied in this paper may reveal other interesting patterns.

References

Baddeley, A. (1986). *Working memory*. Oxford, UK: Oxford University Press.

Baddeley, A. (2007). *Working memory, thought, and action* (Vol. 45). Oxford, UK: Oxford University Press.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2015). *Memory*. East Sussex: Psychology Press.

Bartz, A. (1962). Eye movement latency, duration, and response time as a function of angular displacement. Journal of Experimental Psychology, 64(3), 318.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (Pan troglodytes). *Journal of Comparative Psychology, 103*(1), 23.

Bouma, A. (1987). Effects of concurrent spatial and verbal memory loads on serial position functions of laterally presented letter strings. *Brain and Cognition, 6*(3), 295-320.

Clifton Jr, C., Ferreira, F., Henderson, J. M., Inhoff, A. W., Liversedge, S. P., Reichle, E. D., & Schotter, E. R. (2016). Eye movements in reading and information processing: Keith Rayner's 40-year legacy. *Journal of Memory and Language*, *86*, 1-19.

Chuah, Y. L., & Maybery, M. T. (1999). Verbal and spatial short-term memory: Common sources of developmental change? *Journal of Experimental Child Psychology*, *73*(1), 7-44.

Cocchini, G., Logie, R. H., Della Sala, S., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition*, *30*(7), 1086-1095.

Cornoldi, C., & Vecchi, T. (2003). *Visuo-spatial working memory and individual differences*. London: Psychology Press.

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International, 34*, 891B.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current directions in psychological science*, *19*(1), 51-57.

Dent, K., & Lancaster University. Theses. Psychology. (2003). *Representation and Capacity in Visual-spatial Short-term Memory*.

Dent, K., & Smyth, M. M. (2005). Verbal coding and the storage of form-position associations in visual–spatial short-term memory. *Acta Psychologica*, *120*(2), 113-140.

Depoorter, A., & Vandierendonck, A. (2009). Evidence for modality-independent order coding in working memory. *The Quarterly Journal of Experimental Psychology*, *62*(3), 531-549.

Deyzac, E., Logie, R. H., & Denis, M. (2006). Visuospatial working memory and the processing of spatial descriptions. *British Journal of Psychology*, *97*(2), 217-243.

Fagot, J., & Barbet, I. (2006). Grouping and segmentation of visual objects in baboons (Papio Papio) and humans (Homo sapiens). In *Comparative cognition: Experimental explorations of animal intelligence* (pp. 15-28). Oxford University Press, USA.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, *22*(4), 661-674.

Ginsburg, V., Archambeau, K., van Dijck, J. P., Chetail, F., & Gevers, W. (2017). Coding of serial order in verbal, visual and spatial working memory. *Journal of Experimental Psychology: General*, *146*(5), 632.

Giray, E. F., Altkin, W. M., Vaught, G. M., & Roodin, P. A. (1976). The incidence of eidetic imagery as a function of age. *Child development*, 1207-1210.

Haber, R., & Haber, L. (1988). Eidetic imagery as a cognitive skill. In L. Obler & D. Fein (Eds.), *The exceptional brain: The neuropsychology of talent and special skills* (pp. 218-241). New York: Guildford Press.

Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571-581.

Hayes, B. K., & Heit, E. (2004). Why learning and development can lead to poorer recognition memory. *Trends in cognitive sciences*, *8*(8), 337-339.

Herrmann, E., Call, J., Hernández-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317*, 1360–1366.

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General*, *137*(1), 163.

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, *17*(23), R1004-R1005.

Inoue, S., & Matsuzawa, T. (2009). Acquisition and memory of sequence order in young and adult chimpanzees (Pan troglodytes). *Animal cognition*, *12*(1), 59-69.

Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. *Attention and performance XVI: Information integration in perception and communication*, *16*, 125-155.

Jalbert, A., Saint-Aubin, J., & Tremblay, S. (2008). Short Article: Visual Similarity in Short-Term Recall for Where and When. *Quarterly journal of experimental psychology*, *61*(3), 353-360.

Johnson, J. S., Hollingworth, A., & Luck, S. J. (2008). The role of attention in the maintenance of feature bindings in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 41.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 656.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and language*, *54*(2), 265-281.

Kano, F., & Tomonaga, M. (2009). How chimpanzees look at pictures: a comparative eye-tracking study. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1664), 1949-1955.

Kano, F., & Tomonaga, M. (2011). Species difference in the timing of gaze movement between chimpanzees and humans. *Animal cognition*, *14*(6), 879-892.

Larsen, J. D., & Baddeley, A. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: Do they have a common source? *The Quarterly Journal of Experimental Psychology Section A*, *56*(8), 1249-1268.

Logie, R. (1995). *Visuo-spatial working memory.* Hove: Lawrence Erlbaum.

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *The Quarterly Journal of Experimental Psychology Section A*, *53*(3), 626-646.

Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive? In *Advances in psychology* (Vol. 80, pp. 105-115). North-Holland.

Logie, R. H., Saito, S., Morita, A., Varma, S., & Norris, D. (2016). Recalling visual serial order for verbal sequences. *Memory & Cognition*, *44*(4), 590-607.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279.

Martin, C. F., Bhui, R., Bossaerts, P., Matsuzawa, T., & Camerer, C. (2014). Chimpanzee choice rates in competitive games match equilibrium game theory predictions. *Scientific reports*, *4*, 5182.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature, 315*, 57–59.

Matsuzawa, T. (2009). Symbolic representation of number in chimpanzees. *Current opinion in neurobiology*, *19*(1), 92-98.

Matsuzawa, T. (2012). Outgroup: The study of chimpanzees to know the human mind. In *Evolution of Language, The-Proceedings of The 9th International Conference (Evolang9)* (pp. 227-234). World Scientific.

Matsuzawa, T. (2013). Evolution of the brain and social behavior in chimpanzees. *Current Opinion in Neurobiology*, *23*(3), 443-449.

Myowa-Yamakoshi, M., Scola, C., & Hirata, S. (2012). Humans and chimpanzees attend differently to goal-directed actions. *Nature communications*, *3*, 693.

Nelson, D. L., Brooks, D. H., & Borden, R. C. (1973). Sequential memory for pictures and the role of the verbal system. *Journal of Experimental Psychology*, *101*(2), 242.

Paivio, A. (1971). Imagery and verbal processes. New York, NY: Holt, Rinehart & Winston.

Paivio, A. (2014). Intelligence, dual coding theory, and the brain. *Intelligence*, *47*, 141-158.

Posner, M. I., & Konick, A. F. (1966). Short-term retention of visual and kinesthetic information. *Organizational Behavior and Human Performance*, *1*(1), 71-86.

Postle, B. R., D'Esposito, M., & Corkin, S. (2005). Effects of verbal and nonverbal interference on spatial and object visual working memory. *Memory & cognition*, *33*(2), 203-212.

Roberts, S. G., & Quillinan, J. (2014). The Chimp Challenge: Working memory in chimps and humans. In *The Past, Present and Future of Language Evolution Research: Student volume of the 9th International Conference on the Evolution of Language* (pp. 31-39). EvoLang9 Organising Committee.

Seyfarth, R. M., Cheney, D. L., & Bergman, T. J. (2005). Primate social cognition and the origins of language. *Trends in cognitive sciences*, *9*(6), 264-266.

Sloutsky, V. M., & Fisher, A. V. (2004). When development and learning decrease memory: Evidence against category-based induction in children. *Psychological science*, *15*(8), 553-558.

Smith, T. R., & Beran, M. J. (2017). Arabic Numerals. *Encyclopedia of Animal Cognition and Behavior*, 1-3.

Smyth, M. M., Hay, D. C., Hitch, G. J., & Horton, N. J. (2005). Serial position memory in the visual-spatial domain: Reconstructing sequences of unfamiliar faces. *The Quarterly Journal of Experimental Psychology Section A*, *58*(5), 909-930.

Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford University Press, USA.

Tomasello, M., & Herrmann, E. (2010). Ape and human cognition: What's the difference? *Current Directions in Psychological Science*, *19*(1), 3-8.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92.

Washburn, D. A., Gulledge, J. P., James, F., & Rumbaugh, D. M. (2007). A species difference in visuospatial working memory: does language link "what" with "where"? *International Journal of Comparative Psychology*, *20*(1).

Wasserman, E. A., & Zentall, T. R. (Eds.). (2006). Comparative cognition: a natural science approach to the study of animal intelligence. In *Comparative cognition: Experimental explorations of animal intelligence* (pp. 3-11). Oxford University Press, USA.

Vauclair, J. (1994). In *Language and intelligence in monkeys and apes: Comparative developmental perspectives* (pp. 313-329). Cambridge: Cambridge University Press.

Yamamoto, S., & Tanaka, M. (2009). Do chimpanzees (Pan troglodytes) spontaneously take turns in a reciprocal cooperation task? *Journal of Comparative Psychology*, *123*(3), 242.

Zhao, X., Chen, A., & West, R. (2010). The influence of working memory load on the Simon effect. *Psychonomic bulletin & review*, *17*(5), 687-692.

Zimmer, H. D., Speiser, H. R., & Seidler, B. (2003). Spatio-temporal working memory and short-term object location tasks use different memory mechanisms. Acta Psychologica, 114, 41–65.

Appendix

Table 1

8-digit sequences used as verbal interference

21 43 58 68	22 31 90 78	21 33 89 58
26 14 77 38	29 38 72 28	23 29 00 78
25 82 37 98	29 04 88 98	21 44 98 28
20 69 31 48	29 74 93 08	29 03 67 08
29 94 71 88	20 66 23 58	27 03 55 78

*Note: sequences were chosen randomly by the researcher and varied from one mode to the next. Each sequence had to be recalled once all trials ended for each mode.
Mode refers to arcade, challenge, or chimp mode.