Title
Examining Prioritization in Working Memory for Verbal and Visual Stimuli

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Williamson, Emily
Moss, Jarrod

Publication Date
2022

Peer reviewed

Examining Prioritization in Working Memory for Verbal and Visual Stimuli

Emily Williamson (eaw361@msstate.edu)
Jarrod Moss (jarrod.moss@msstate.edu)
Department of Psychology, Mississippi State University
Mississippi State, MS, 39762 USA

Abstract

The effect of prioritization on information in working memory has primarily been examined in tasks containing a single type of stimulus and with one item that is prioritized. However, many theories of working memory posit different types of components for the maintenance of verbal or visuospatial information. This study examined differences between prioritized and nonprioritized items as well as word and image stimuli. Participants completed an association learning task in which working memory demands were varied along with the number of items to be prioritized. Following a short delay, retention was tested. Prioritization effects were identified during both the learning and testing phases of the experiment, and the impact of prioritization was moderated by working memory demands of the task. Significant differences in accuracy between word and image stimuli were only observed in the testing phase, with accuracy for verbal information being worse. While prioritization improved accuracy and response times during learning, it led to decreases in the testing phase.

Keywords: working memory; prioritization; learning

Prioritization is a highly productive component of everyday mental functioning, from keeping us safe (e.g., keeping an eye on the cars around us on the road rather than the passing scenery) to playing videogames (e.g., choosing to work towards a specific achievement over others). When referencing its intentional direction towards sensory stimuli, this prioritization is referred to as selective attention. Gazzaley and Nobre (2011) discussed the relationship between working memory and selective attention, which they define as, "the ability to focus our cognitive resources on information relevant to our goals." However, with respect to information already in working memory, research has shown that it is possible to prioritize some information over other information (Allen & Ueno, 2018; Atkinson et. al., 2020; Heuer & Schubo, 2020). Understanding the mechanisms and limitations of prioritizing information in working memory should inform not only theories of working memory, but also how we use these mechanisms in making decisions, learning, and many other tasks.

Working memory prioritization effects have been repeatedly shown to increase accuracy for prioritized information (Allen & Ueno, 2018; Atkinson et. al., 2020; Heuer & Schubo, 2020), but this is not equivalent to increasing the capacity of working memory. Atkinson et. al. (2020) conducted a series of prioritization experiments with a concurrent articulatory suppression task, repeating a word or phrase unrelated to the experiment, designed to suppress one of the maintenance mechanisms involved in verbal working memory. They asked participants to prioritize the item at a specific serial position in the sequence being read to them, then repeat as much of the sequence as they could recall. These experiments indicated that prioritized item at a specified serial position will be maintained at the cost of nonprioritized items as increasing working memory resources are taken up by the prioritized item. In these studies, a single item in a serial recall task was selected to have a higher priority.

However, multiple items may be prioritized simultaneously. Allen and Ueno (2018) demonstrated this by manipulating the number of prioritized items across four visuospatial experiments. Participants were first shown a screen indicating which of four spatial locations would contain items that were meant to be prioritized. Across these four experiments, which contained up to three prioritized items, researchers consistently found significantly higher recall accuracy for prioritized items when compared to nonprioritized items. This task provides evidence that the act of prioritization may be more flexible than thought previously by allowing multiple items to be marked as prioritized. More prior research has examined prioritization in visuospatial working memory, with the recent study by Atkinson et al. (2020) being one of the few to look at prioritization in verbal stimuli. To the best of our knowledge, there are no studies directly examining differences in prioritization between verbal and visual information.

There are theoretical reasons that the type of stimulus information might matter. Some theories of working memory posit different modules or components that specialize in maintaining different kinds of information. In particular, the multicomponent theory of working memory (Baddeley, 1992) and the time-based resource-sharing (TBRS) theory of working memory (Barrouillet & Camos, 2020) both have components that specialize in maintaining certain types of information. Verbal, or phonological information, according to these theories, can be processed within a space known as the phonological loop by Baddeley or the articulatory loop in TBRS. Both theories also describe two methods that this can be used to maintain verbal information so it can be manipulated or recalled. The first is a process involving repeating the information either internally or aloud using the articulatory loop, and the second is a process involving rapid attentional switches, referred to as refreshing, amongst the information in working memory to prevent loss of information. Previous studies identified the presence of more than one maintenance strategy by intentionally blocking one or the other to measure the remaining effect (Atkinson & Allen, 2020). Maintaining information in verbal working memory is often done with the intention of reproducing that

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

information at some later point (e.g., repeating a phone number over and over until you can put it in your phone), so its capacity is often determined by accuracy of serial recall. While working memory possesses two mechanisms for holding onto information, the area that handles visual information only utilizes the attentional refreshing method (Heuer & Schubo, 2016). This difference in mechanisms available for verbal and visual information may impact how different stimuli can be prioritized in working memory.

Examining varying prioritization demands has been done primarily in either visuospatial working memory tasks or via a serial recall task with auditory stimuli. The goal of the current study is to examine prioritization effects with different working memory loads in a learning task with either image or word stimuli. The rationale for using a simple association learning task is to see if similar prioritization effects can be observed when working memory is being used in the service of completing a larger task. The association learning task has been used previously to examine interactions of working memory and reinforcement learning (Collins, 2018; Collins & Frank, 2012). These prior results have been interpreted as supporting a combination of working memory learning and reinforcement learning, with set size being the distinguishing factor. Within the smaller set sizes that were still within the capacity limit of working memory, the results were interpreted as supporting primarily maintenance in working memory with limited contribution of reinforcement learning. The larger set sizes, which exceeded participants' capacity limits, required reinforcement learning in addition to working memory from the onset of the task. The key result was that at a delayed test, associations from the larger set size were retained better than associations learned at a smaller set size where the associations could be maintained primarily within working memory.

Within the task itself, participants are given instructions to earn as many points as possible by guessing which of three key presses is the correct key to earn points for that stimulus. Prioritization is incentivized through a scoring system that makes some stimuli worth twice as many points. The number of stimuli being learned at one time within a block is manipulated in the range of 3 to 6 resulting in varying working memory demands. In the priority condition, half of the items in a block are prioritized. In the control condition, no stimuli are prioritized. Comparison of the prioritized and non-prioritized items in the priority condition will provide evidence of prioritization effects. Non-prioritized stimuli in the priority condition can also be compared to the control condition to see if non-prioritized items were any worse because resources were allocated to prioritized items. The primary question of interest is whether prioritization effects would be observed and if they were moderated by working memory load. Our initial hypothesis was that the impact of prioritization would increase as set size increased; some of this difference between prioritized and non-prioritized items might be because the non-prioritized item performance was lower due to allocation of resources to the prioritized item.

The other main comparison in this study will be differences between word and image stimuli. Because it may be easier to represent words in the articulatory loop, having an additional subsystem might make prioritization different from image stimuli. For this reason, we hypothesized that it may be easier to prioritize items in the word condition than in the image condition.

Method

Design

A 2 x 3 x 2 mixed repeated-measures design is used in this experiment, with major factors being priority condition (priority, control), set size (3, 4, and 6), and stimulus type (word or image). Fifteen different categories of nameable items were first developed, then six different images (to account for the maximum set size) were obtained for each category to serve as the image condition. Once all image stimuli were gathered, a text version of each of these visual items (i.e., image: image of cat, word: "cat") were generated for the word condition. These stimuli were chosen to identify hypothesized differences between the mechanisms involved in maintaining these types of information.

The main task for this experiment is an adapted version of the association learning task meant to measure reinforcement learning as well as possible working memory learning. The present version of this task contained 15 blocks, each containing 3, 4, or 6 different stimuli from the same category (e.g., animals, sports, fruit). The task for each block was to learn which of the three keys ("A," "S," or "D") was the correct key to press for each image. Correct/incorrect feedback was provided after each attempt in the form of points (0, 1, or 2 points). No points were earned for incorrect key presses, so that receiving points indicates the key was the correct key. Each stimulus was presented 13 times in a block with all stimuli for the block being presented in a pseudorandom order. Participants were randomly assigned to one of four conditions formed by crossing the priority condition and stimulus type factors. In the control condition, all stimuli were worth 1 point. In the priority condition, half of the stimuli were worth 2 points such that 1, 2, and 3 items were worth more points in the 3, 4, and 6 set size conditions respectively. Set size was manipulated within participants with 5 blocks of each set size. The order of set sizes across blocks was randomized.

Following the association learning task, participants completed a delay task that lasted 10-20 minutes (M = 16 minutes). The delay task was the automate operation span task (Unsworth et al., 2005), which is a complex span task designed to measure working memory capacity. The data from this task is not analyzed here, and the methodological details are identical to those reported in Unsworth et al. (2005). After the delay task, they completed a testing phase where they were asked to recall the correct key press for each stimulus seen in the association learning task without feedback. Participants were not told in advance that they would be tested on the key press associations. This phase randomly presented each of those associated stimuli from the learning phase, with the instructions that the associated key presses were the same for this phase as the learning phase. Feedback was not provided following each key press, but otherwise the timing of the trials was identical to the learning phase. At the end of the task, they received a cumulative score for the entire testing phase.

Participants

Two hundred and sixteen undergraduate student participants participated in exchange for course credit. They were randomly assigned to one of four conditions (priority image, priority word, control image, and control word) and completed the experiment online from personal computers. Similar to the original Collins (2018) experiment, an a priori learning criterion was set at 75% accuracy across the last three stimulus repetitions in a block in the set size three condition. This criterion was included to exclude participants who were not able to learn a condition in which prior studies have found that mean accuracy is 90% or greater. Additional exclusion criteria included participants which took more than twenty-five minutes to complete the OSPAN task, so those with a longer than average delay between the learning and testing phases would not be analyzed. Finally, participants that scored lower than 80% averaged accuracy for the algebra problems within the OSPAN were excluded on the basis of not performing the task to their full ability.

Based on prior data from the association learning task (without prioritization), half of the difference between performance in the set size three and set size six conditions was used as an estimate of the effect size of prioritization for purposes of a power analysis. Monte Carlo simulations based on this effect size were used to determine that 160 participants should yield a power of .9 to detect the prioritization effect. After applying the learning criterion, 161 participants remained in the analysis.

Procedure

All participants saw the same instructions at the onset of the experiment. They were instructed to place their hand over the three keys they were using to choose associations on the keyboard (A, S, and D). As they proceeded through the instructions at their own pace, they were shown what a set of images looked like for each block of the main experiment. Next they practiced, step by step, associating keys with practice images and received feedback as they would in real trials. The first condition difference came from an extra set of instructions that appeared to those in the priority condition, indicating items with a blue border were worth two points when associated correctly. Once they read through the instructions, they performed a practice block with a set size of three. This was identical to later blocks, except that the stimuli used here was not used in later blocks. A performance criterion was set at 80% of the last five presentations of each stimulus for them to proceed. If they did not meet the criterion, then they reread the instructions and tried again.

Participants in the image condition received images as their stimuli for the entirety of the experiment, and those in the word condition received word stimuli. At the beginning of a block of trials, all the stimuli they would be learning in that block were presented on the screen as shown in Figure 1. They could study these stimuli for up to one minute. For each trial, a 500 ms fixation cross was presented followed by the stimulus for up to 2000 ms, during which participants made their response. Following the response or the time limit, they received feedback in the form of "+0" for incorrect guesses, "+1" for correctly associated control/nonpriority items, and "+2" for prioritized items in the priority condition. Failure to respond in the time limit led to "Too Slow" as feedback. Feedback was presented for 500 ms before cycling to the next stimulus. Each of the stimuli were presented 13 times in a pseudorandom order such that the number of intervening trials between successive presentation of the same stimulus followed a uniform distribution. Following the last trial in a block, the total cumulative points earned in that block were presented for 2500 ms. Each participant completed 5 blocks of each of the three set sizes.

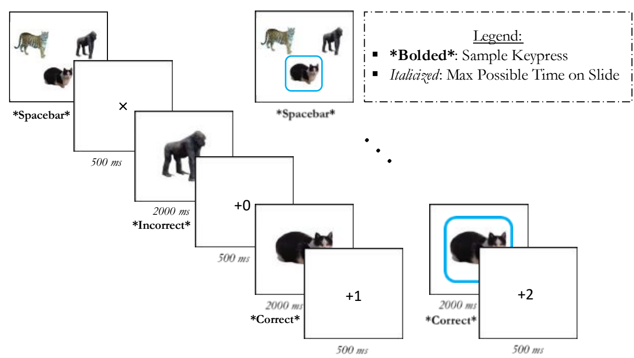

Figure 1: Association learning task, showing two trials and difference in priority condition.

They then completed the operation span task. Following its completion, they completed the testing phase. Participants were not told in advance that they would be tested on the key press associations. This phase presented each of the stimuli they learned associations for during the learning phase in a random order, and they received instructions that the associated key presses are the same for this phase as the learning phase. Feedback is not provided following each key press, but otherwise the timing of the trials was identical to the learning phase. At the end of the testing phase, they received a cumulative score for the testing phase.

Analyses

Generalized linear mixed effects models were developed in R to analyze this data with degrees of freedom being approximated by Satterwhaite's method as implemented in the lmerTest package. The dependent measures examined were accuracy during the learning phase, correct response time during the learning phase, and accuracy during the testing phase. For each dependent measure, one set of analyses examined the difference between prioritized and non-prioritized items in the Priority condition. The purpose of this set of analyses was to examine whether prioritization had any impact. In these analyses, priority status (within priority condition, prioritized vs. non-prioritized items), stimulus type (image vs. word), and set size as a continuous variable were included as fixed effects. For random effects, there were random intercepts of participant and stimulus with random slopes for all factors that were manipulated within-participant or within-items (i.e., priority status and set size for participants, and priority status for items).

A second set of analyses compared between-subject effects using the non-prioritized items in the priority condition and performance in the control condition. The purpose of these analyses was to examine whether prioritizing some items lead to a decrease in performance on non-prioritized items. In these analyses, priority condition (priority vs. control), stimulus type (image vs. word), and set size as a continuous variable were included as fixed effects. For random effects, there were random intercepts of participant and stimulus with random slopes for all factors that were manipulated within-participant or within-items (i.e., set size for participants, and priority condition for items). For all analyses, all fixed effects were allowed to interact but nonsignificant interactions were dropped in order to provide easier interpretation of lower order effects.

Results

Learning Phase

Both accuracy during the learning phase and correct response time were examined. Based on prior results with this association learning task, the smaller set size conditions are likely to be easier than the larger set size condition (Collins & Frank, 2012). As such, there may be a ceiling effect in the set size 3 and 4 conditions limiting the effect of prioritization. It may be that learning phase response times are a better indicator of prioritization in these smaller set sizes.

Prioritized vs. Non-prioritized Items Mean accuracy for the learning phase is shown in Figure 2. For the general linear mixed effects model contrasting prioritized and non-prioritized items in the priority condition, set size interacted with priority status, $z = 4.18$, $p < .001$. Contrasts showed that this interaction is due to priority status improving accuracy in the set size six condition, $z = 3.03$, $p = .003$, but not in the smaller set sizes, $p > .05$ for both set size three and four. Increasing set size also decreased accuracy, $z = -16.02$, $p < .001$). There were no significant effects of stimulus condition or priority status, and no other significant interaction.

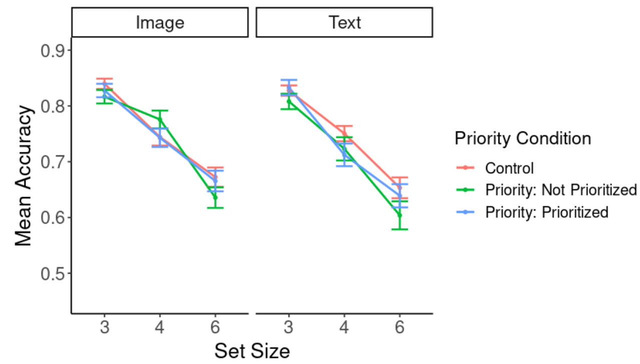

Figure 2: Accuracy in learning phase across set size, separated by stimulus condition. Note that priority condition was manipulated across participants and priority status within the priority condition was manipulated within participants.

Mean correct response time for the learning phase is shown in Figure 3. For the linear mixed effects model contrasting prioritized and non-prioritized items in the Priority condition, prioritized items were responded to faster than non-prioritized items, $t(150.85) = -3.65$, $p < .001$. Larger set sizes also took longer to respond to, $t(78.50) = 9.32$, $p < .001$. There were no significant effects of stimulus condition or significant interactions.

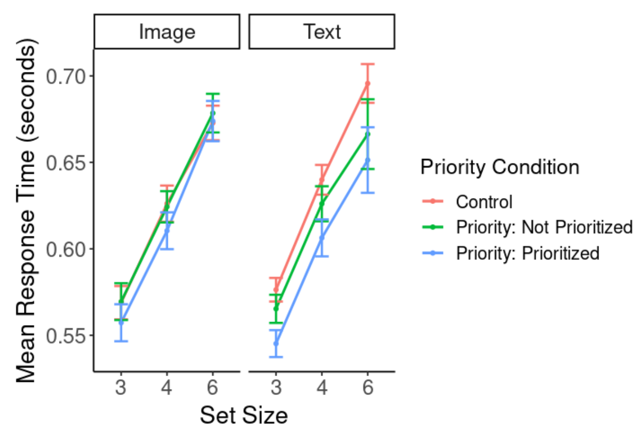

Figure 3: Mean response time in the learning phase across set size, separated by stimulus condition.

Non-prioritized Items vs. Control Condition Non-prioritized items were compared to items in the control condition to examine whether any benefits of prioritization came at the cost of worse performance for the items that were not prioritized in the priority condition. For accuracy, higher set sizes had lower accuracy, $z = -21.95$, $p < .001$, but there was no evidence that priority condition had an effect, $z = -0.89$, $p = .38$. There were also no significant effects of stimulus condition. Similarly for response times, set size led to an increase in response time, $t(152.90) = 15.54$, $p < .001$, but there was no evidence of an effect of priority condition, $t(179.46) = -0.93$, $p = .36$ or significant interactions.

Testing Phase

Prioritized vs. Non-prioritized Items Mean accuracy for the testing phase is shown in Figure 4. For the general linear mixed effects model contrasting prioritized and non-prioritized items in the priority condition, set size interacted with priority status, $z = 4.06$, $p < .001$. Contrasts showed that this interaction is due to priority status leading to decreases in accuracy for set sizes of three and four (Set 3: $z = -4.72$, $p < .001$; Set 4: $z = -3.99$, $p < .001$), but no significant difference in accuracy for set size six. Set size also interacted with stimulus condition, but only had a significant effect for image accuracy, $z = 2.83$, $p = .005$. This interaction was due to set size affecting image stimuli accuracy, $z = -3.61$, $p < .001$, but not word stimuli accuracy. Increasing set size led to lower overall accuracy, $z = -2.36$, $p = .02$, and there was overall lower accuracy for prioritized items, $z = -3.13$, $p = .002$. The word condition also had lower accuracy than the image condition, $z = -4.35$, $p < .001$. There were no significant interactions between priority status and stimulus condition.

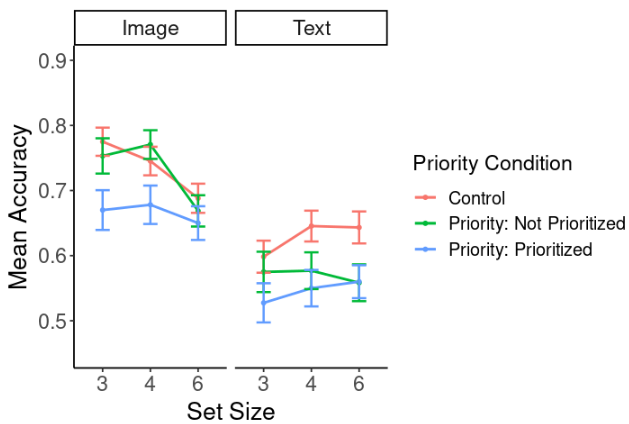

Figure 4: Accuracy in testing phase across set size, separated by stimulus condition.

Non-prioritized Items vs. Control Condition Non-prioritized items were compared to items in the control condition to examine whether any benefits of prioritization came at the cost of worse performance for the items that were not prioritized in the priority condition. Overall accuracy for nonprioritized items decreased as set size increased, $z = -1.64$, $p < .001$. Accuracy was also significantly worse in the word condition than the image condition, $z = -5.26$, $p < .001$. There was also significant set size by stimulus type interaction, $z = 5.20$, $p < .001$, with accuracy decreasing for items in the image condition with increasing set size, $z = -6.78$, $p < .001$, but not the word condition. No evidence of a difference between nonprioritized items in the priority condition and items in the control condition was identified. There were also no significant main effects of priority status or set size, and no interaction between stimulus condition and priority status.

Discussion

A prioritization effect was observed in both the learning and testing phases, and effects of stimulus condition were found in the testing phase. Our initial hypothesis was that there would be a larger prioritization effect as set size increased. While accuracy decreased as set size increased in the learning phase, an effect of prioritization was present only in set size six during this phase. This result supports the original hypothesis. The second main hypothesis examined whether maintenance for word stimuli would lead to higher accuracy than visual stimuli because the availability of components like the articulatory loop could be used in addition to mechanisms like attentional refreshing in the episodic buffer. However, there were no effects or interactions with stimulus type during the learning phase. The only significant effect of stimulus type was observed during testing, where accuracy for word stimuli was lower than image stimuli. Finding better memory for images as opposed to words is not an original finding, as prior studies have observed higher free recall and recognition memory for pictures (Paivio & Csapo, 1973). However, it was interesting that memory for key presses associated with images were impacted more by increases in set size than were the key presses for words. In any case, one limitation of the current design is that the image stimuli were designed to be easily verbalizable. It is possible that participants recoded the images, and could use the articulatory loop to help with maintaining these items.

The combination of these accuracy and response time results may be due to ceiling effects on accuracy for the smaller set sizes. By the end of the learning phase, mean accuracy was over 90% for set size three and approaching 90% for set size four. Because the smaller set sizes were under the capacity of the working memory maintenance mechanisms involved, prioritization may not have led to lower accuracy for non-prioritized items, as all items could be maintained in some manner in working memory. While accuracy for these lower set sizes was not influenced by prioritization, the significant prioritization effect in response times indicates that prioritization played a role in performance even at smaller set sizes. Within current theories of working memory, there are several possible ways that prioritized word items could be maintained differently than non-prioritized items. For example, in the time-based resource sharing theory (Barrouillet & Camos, 2020), prioritized information could be refreshed in the episodic buffer using attentional refreshing while non-prioritized items were maintained in the articulatory loop or simply retrieved from declarative long-term memory as needed. Another possibility within this same theory is that high priority items are refreshed at a higher rate than non-prioritized items. Similar possible explanations can be constructed for other theories of working memory.

One additional result that constrains these possible explanations is the decrease in accuracy for prioritized items during the testing phase. Whatever mechanism is posited for prioritized items must also lead to lower availability in long-term declarative memory for prioritized items. This decrease in testing accuracy for prioritized items bears some similarity to priori results with this association learning task where learning was better for set size three items than for set size six items, but testing accuracy was lower for set size three than for set size six items (Collins, 2018). Some results have indicated that this accuracy difference may be due to the proceduralization of the responses for set size three items (Newlin & Moss, 2020). Highly practiced items may lead to the formation of procedural memories from the original declarative memories (Anderson, 1982; Taatgen & Anderson, 2002). These procedural memories would be context specific and disrupted by the differences between learning and testing context. Supporting this idea, Newlin and Moss found that when the testing condition matched the learning condition by being blocked by stimulus category instead of randomly presenting items across categories, then testing performance matched learning performance more closely for small set sizes. It may be possible that a similar effect is being observed here where prioritized items become proceduralized, so they did not need to be retrieved from memory any more after learning. However, the testing phase does not match the learning phase, so that procedural knowledge is not available during test. This explanation of performance in this task can be contrasted with that of Collins (2018), which hypothesized that maintaining associations in working memory interfered with reinforcement learning. The current testing results could also be explained with that theory.

While smaller set sizes may have been proceduralized or maintained differently in working memory, set size six contained too many items to continue this strategy. The capacity of these maintenance mechanisms is low, so prioritizing the items that were worth more points became the favorable strategy. Items that were not prioritized were not maintained as successfully, leading to the observed difference in accuracy. Additionally, lower accuracy for prioritized items during the testing phase indicates those items were not maintained as well across the delay. These results, as well as the nature of prioritization being a focused, deliberate process leads us to posit that prioritization is more aligned with working memory learning rather than reinforcement learning. Future work should further test this hypothesis, as well as explore designs that test alternate working memory mechanisms that could also account for the prioritization effects in this study.

The incentive to prioritize items may also be a limitation in this study. Points in this task were not associated with any monetary or credit incentive for participants, and it is possible that such an incentive would yield increased prioritization effects. Another manipulation that could be examined in the future is a concurrent articulatory suppression task that inhibited verbal maintenance mechanisms. Such a manipulation would begin to examine some of the explanations proposed for the results.

In conclusion, these results show that it is possible to obtain prioritization effects in working memory with multiple prioritized items across different types of stimuli. The association learning task used in this study is different from the immediate serial recall used in studies of verbal working memory as well as from visuospatial tasks used in studies of visual working memory. Prioritization in working memory can therefore have an effect when performing a task in which there can be strategic variation across participants. Tasks similar to this one may allow further hypotheses to be tested regarding interactions of working memory components with other cognitive systems that operate during learning, such as the potential proceduralization or skill acquisition effects discussed here.

Acknowledgments

This research was supported by Office of Naval Research grant N00014-21-1-2617.

References

Allen, R. J., & Ueno, T. (2018). Multiple high-reward items can be prioritized in working memory but with greater vulnerability to interference. *Attention, Perception, & Psychophysics,80*,1731–1743. https://doi.org/10.3758/s13414-018-1543-6

Anderson, J. R. (1982). Acquisition of Cognitive Skill. *Psychological Review 1982, Vol. 89*, No. 4, 369-406.

Atkinson, A. L., Allen, R. J., Baddeley, A. D., Hitch, G. J., & Waterman, A. H. (2020). Can Valuable Information Be Prioritized in Verbal Working Memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*, 747-764.

Atkinson, A. L., Berry, E. D., Waterman, A. H., Baddeley, A. D., Hitch,G. J., & Allen, R. J. (2018). Are there multiple ways to direct attention in working memory? *Annals of the New York Academy of Sciences,1424*, 115–126. https://doi.org/10.1111/nyas.13634

Baddeley, A. (1992). Working Memory. *Science, 255*, 556-559. https://doi.org/10.1126/science.1736359.

Barrouillet, P., & Camos, V. (2020). The Time-Based Resource-Sharing Model of Working Memory. In P. Barrouillet & V. Camos, Working Memory (pp. 85–115). Oxford University Press. https://doi.org/10.1093/oso/9780198842286.003.0004

Camos, V., & Barrouillet, P. N. (2014). Attentional and nonattentional systems in the maintenance of verbal information in working memory: The executive and phonological loops. *Frontiers in Human Neuroscience,8*, 900. https://doi.org/10.3389/fnhum.2014.00900

Collins, A. G. E. (2018). The Tortoise and the Hare: Interactions between Reinforcement Learning and Working Memory. *Journal of Cognitive Neuroscience*, 1-11.

Collins, A. G., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience 35*, 1024-1035. https://doi.org/10.1111/j.1460-9568.2011.07980.x

Heuer, A., Schubo, A. (2016). The Focus of Attention in Visual Working Memory: Protection of Focused Representations and Its Individual Variation. *PLOS One: Cognitive Science*, *11*(4): e0154228. https://doi.org/10.1371/journal.pone.0154228.

Newlin, P., Moss, J. (2020). Proceduralization and working memory in association learning. *Proceedings of the 42nd Annual Meeting of the Cognitive Science Sociaty* (pp. 191-197). Toronto: Cognitive Science Society.

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*(2), 176-206. https://doi.org/10.1016/0010-0285(73)90032-7.

Taatgen, N. A., & Anderson, J. R. (2002). Why do children learn to say "Broke"? A model of learning the past tense without feedback. *Cognition 86*, 123–155.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498.