Proceduralization and Working Memory in Association Learning

Philip Newlin (philip.newlin@psllab.org)
Department of Psychology, 225 Lee Blvd
Mississippi State, MS 39762 USA

Jarrod Moss (jarrod.moss@msstate.edu)
Department of Psychology, 225 Lee Blvd
Mississippi State, MS 39762 USA

Abstract

Humans are highly variable in their ability to learn and execute complex tasks; however, there are conflicting theories on skill acquisition. This study compared two different explanations for how association learning interacts with other cognitive processes: a) reinforcement learning and working memory are separate, competing processes operating simultaneously on association learning; and, b) associations are proceduralized into production rules and reinforcement learning acts on those rules. Participants completed a simple association learning task followed by a delayed test under two conditions designed to contrast these theories. The results are consistent with a proceduralization account in which reinforcement learning and working memory are not competitive interfering systems, but there remain important questions about how these two accounts may be best integrated.

Keywords: reinforcement learning; working memory; procedural learning; skill acquisition

Imagine you just encountered a light switch for the first time. With some trial and error, you surmise that by positioning your fingers above or below the switch and elevating or lowering your hand, you can reposition the switch at its vertex. After the switch has clicked into place, you notice the level of luminance in the room change. You try to recreate your new discovery and are repeatedly met with predictable results. The next time you encounter a light switch, you could operate it by remembering what you did last time. Upon further interactions with the light switch this process is further proceduralized, up to the point where you no longer think about the positioning of your hand, the actuation of your joints, the mechanism of the pivoting switch, and the range of outcomes that may follow. With enough experience, it is more efficient to use a proceduralized rule: if the goal is to produce light and contained in the room is a light switch, then flip the light switch.

There are at least two theoretical frameworks that describe the processes involved in learning associations such as that between a switch and the correct response to activate the switch. One perspective is that these simple associations are learned via reinforcement learning. Another perspective is that this is a rudimentary skill acquired in the same manner as other skills. The goal of the study reported here was to examine predictions arising from these two theoretical frameworks. Both frameworks have substantial empirical evidence supporting them. Therefore, the long-term goal of this research is to examine ways in which these two frameworks might be integrated into a more comprehensive theoretical framework.

In the skill acquisition framework, skill is theorized to transition from a declarative representation to a procedural representation (Anderson, 1982). The declarative representation can come from reading a set of instructions or via a trial-and-error process like that described for the light switch. As the declarative representation is retrieved in performing the skill, it becomes proceduralized into a procedural memory. Within the ACT-R architecture (Anderson et al., 2004), declarative knowledge is represented as a chunk and procedural knowledge as a production rule that can be executed to perform the skill. A production compilation mechanism explains how declarative knowledge is compiled into new production rules (Taatgen & Anderson, 2002). Once a new production has been compiled, it offers an additional method for doing the task that does not rely on retrieving a declarative memory. The compiled procedure and the procedure to retrieve the declarative representation (i.e., a retrieval-based strategy) are now alternative competing procedures. As these procedures get reinforced upon successful execution of the skill, the faster and less error prone compiled production gains a much higher estimate of utility and gets selected much more often than the retrieval strategy. This skill acquisition mechanism has been used to explain skills such as learning the past tense (Taatgen & Anderson, 2002). In this account of skill learning, reinforcement learning is used to learn which procedure is most likely to yield a quick and accurate response, and working memory holds any retrieved declarative chunks while they are used by a production.

An alternative framework arising from the study of reinforcement learning proposes that reinforcement learning competes with a more explicit working-memory based system of learning (Collins, 2018; Collins et al., 2017; Collins et al., 2017; Collins & Frank, 2012). A recent study explored a potential interaction between reinforcement learning and working memory (Collins, 2018). From this perspective, working memory is a short-term, capacity limited form of memory used to briefly hold task-relevant information, and reinforcement learning is a slower, capacity unlimited mechanism whereby actions increase or decrease in their likelihood of being repeated as a function of their consequences. These two systems are said to both simultaneously operate on learning tasks such that items being held in working memory do not receive updates to their value via reinforcement learning (Collins, 2018).

Original Study

The original study used a simple associative learning task, wherein participants learned which of three keys was the correct key to press for each image in a set of images (Collins, 2018). For example, if an image of a triangle was shown, one might guess that the "a" key was correct. Correct/incorrect feedback would then be provided, and the next image shown. Over many presentations of the triangle with feedback, one would eventually learn the correct key press. During the learning phase of this task, participants were asked to learn a block of images and the correct key that corresponded with each image. Each block consisted of a set of either 3 or 6 images that were categorically related (e.g., colors, shapes). Each image in a block was presented 12-14 times in a pseudorandom order, and there were 12 blocks of images. Half of the blocks had 3 images in the set (small set size) and half had a set size of 6 images (large set size). Different set sizes (3 and 6) were used in the learning phase to introduce different levels of working memory load. The assumption was that for the small set size, but not for the large set size, it was more likely that all of the stimulus-response associations could be held in working memory.

After a delay following the learning phase, there was a surprise test phase, in which participants were presented with the images in random order and asked to select the correct key associated with each image. During learning, the larger set size takes longer to learn and responses to those items take longer. Counterintuitively, during the test phase, responses for images from the larger sets were recalled more accurately (and more slowly) than those from the smaller sets (Collins, 2018). The interpretation of the results was that associations presented in smaller set sizes are more often held in working memory, which operates more quickly leading to better learning performance. Working memory is capacity limited so all of the associations for the large set size could not be held there. Associations held in working memory were not updated via reinforcement learning. Because working memory representations are not held onto over a long delay, only the representations learned via reinforcement learning are available at test. Associations learned in larger sets are updated via reinforcement learning more frequently and were available at test. Therefore, performance on large set size items was higher at test and experienced less forgetting than the low set size. The conclusion, supported by a comparative model fits of models of reinforcement learning and working memory, was that a model with interacting working memory and reinforcement learning systems was the best explanation of the results (Collins, 2018).

Present Study

The theory of skill acquisition and production compilation offers another possible interpretation of these results. When an image is first encountered, each of the three keys has the same relative probability of being correct. Once the correct response for an image has been produced, then an alternative strategy for responding is available (i.e., recall what key was correct last time). Once this response has been recalled a number of times, a compiled production is available that eliminates declarative retrieval: if the goal is to respond to shapes and you see a triangle, then press the "a" key. The utility of the newly compiled rule increases via reinforcement and eventually surpasses the declarative retrieval strategy as it consistently yields the same response more quickly than retrieval (Anderson et al., 2004; Taatgen & Anderson, 2002).

During learning, as shown in Figures 1a and 1b, the average delay between items shown in the smaller set size is smaller than the delay in the large set size. This delay has implications for the rate of learning and the rate of compiling a task-specific production rule, accounting for the differential rate of learning for the two set sizes. The chances are higher that the correct response for a large set size image cannot be recalled because the delay between trials of the same image is longer.

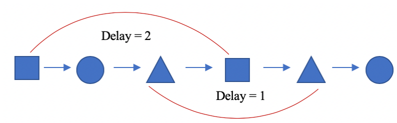

Figure 1a: Stimuli in smaller set sizes have a smaller interval between presentations. Response mappings can be more easily retrieved from declarative memory because of the shorter delay, allowing more opportunities for the response mapping to be compiled into a production rule.

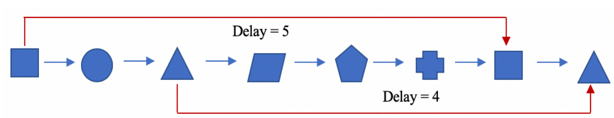

Figure 1b: Stimuli in larger set size blocks have a longer interval between presentations, making it more difficult to retrieve the previous response from declarative memory when the stimulus is next encountered. Consequently, this set size takes longer to learn and there are fewer opportunities for a compiled production to be produced and compete with the retrieval strategy.

During the test phase, a critical aspect of the original study is that stimuli were presented at random, rather than in blocks from the same category as they were presented in the learning phase. If the compiled productions from learning contain information about the category, then they will fail to apply in the test where the category information is not a central part of the task. For example, this rule references the category and would not apply in the test phase: if the goal is to respond to *shapes* and you see a triangle, then press the "a" key. In this case, during the test phase, only the retrieval strategy is available. Retrieval time and success will be related to the frequency and recency that these stimuli were retrieved from declarative memory. In the case of the smaller set size, frequency and recency of retrieval will be on average lower because a compiled production took over and eliminated declarative retrieval. Therefore, the lower accuracy for the small set size during test results from failure to retrieve because the task-specific rule eliminated the need to retrieve that information during learning.

The crux of this alternative explanation for the findings of Collins (2018) is that when task-specific production rules can be compiled quickly and reinforced enough times, then they eliminate the need to retrieve the correct response from declarative memory. Since recency and frequency are both predictive of future retrieval success, reducing recency leads to greater difficulty retrieving the correct response at test when the proceduralized rule is not applicable. These alternative theoretical explanations have important implications for how skill acquisition, working memory, and reinforcement learning interact.

The current study compares these two theories by replicating the method of the prior research (Collins, 2018) and manipulating across participants the presentation of the items at test. Randomization in the test phase presumes that the associations were learned individually, rather than in blocks as they were presented in the learning phase. A between-subjects manipulation where half the participants see the original design with the stimuli randomized at test and the other half of the participants encounter stimuli blocked according to category in the same manner as the learning phase should reveal if response mappings are learned procedurally or retrieved from working memory.

By reinstating the same method at test as during learning, any proceduralized rules are still applicable and will generate faster and more accurate responses than the retrieval strategy. With this small modification to the experiment, it is expected that 1) response accuracy will be the same or greater for the small set size than for the large set size; and, 2) the response time for the small set size will be faster than the large set size. These hypothesized results for the blocked testing condition are the opposite of what is expected in the randomized test condition (e.g., a replication of the original results).

Method

Participants

A sample of 176 adult participants was recruited from undergraduate courses at a large public university in exchange for course credit. Based on an effect size estimated from the results reported in Collins (2018), it is expected that 70 participants in each condition (for a total of 140) should yield .8 power. As in the original study, an asymptotic performance criterion of 75% correct was set for response accuracy in the learning phase of the associative learning task to ensure that participants had learned response mappings prior to test. Asymptotic performance was measured as the average response accuracy on the last 6 trials for each image. 30 participants were eliminated based on this criterion, resulting in a sample size of 146.

Tasks

Participants completed an associative learning task identical to that described by Collins (2018). The learning and test phases of the task were separated by performance of the automated operation span task (Unsworth et al., 2005). The original study used a visual N-back task for the delay between learning and test. The nature of the delay should not matter as long as participants are not rehearsing the stimuli or responses. The operation span task lasted on average 13 minutes (similar to the original study's delay of 11 minutes). The operation span (OSPAN) was selected because the task duration was similar to the visual n-back task used in the original design while also providing a measure of working memory.

Associative Learning Task During the learning phase, participants were asked to learn a set of images and pair them with a set of responses. In each block, there are either 3 or 6 distinct objects from the same category (e.g., shapes), each having one of 3 distinct correct key presses. The goal is to learn via correct/incorrect feedback the corresponding key press for each object. Each block includes 13 trials per stimulus, regardless of the number of distinct images (3 or 6) presented in the block. Each block and its respective stimuli were presented in a pseudorandom order unique to each participant, with a uniform distribution of delays (number of stimuli between repetitions of one stimulus) for each stimulus (Figure 2). There was a total of 14 blocks of images in this phase: eight blocks of categories with three images (i.e., set size 3) and six blocks of set size 6. The first and last blocks of the training phase were set size 3 and were not analyzed to control for primacy and recency. The remaining 12 blocks were presented in random order.

In the test phase, participants were presented with the same stimuli from the learning phase and asked to provide the associated keypress, without reinforcement. Participants were randomly divided into one of two conditions in the testing phase. In the random condition, all stimuli from all learning blocks were presented in random order as in the original study (Collins, 2018). In the blocked condition, stimuli in the testing phase were presented in blocks, in the same manner that they were presented in the learning phase.

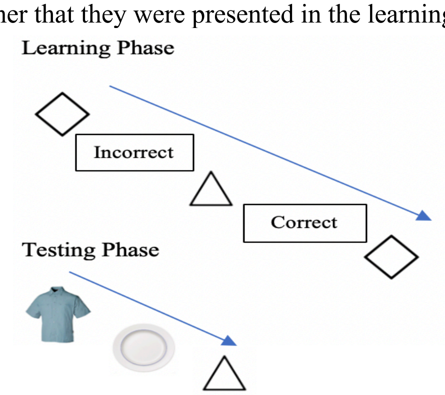

Figure 2: Stimuli are presented in a pseudorandomized order for 1500 ms. Feedback is presented for 500 ms immediately

after the response, followed by a 500 ms fixation cross. During the test phase, the trials are identical except that feedback is not displayed.

Operation Span The OSPAN was used both as a distractor task between the learning phase and test phase and as a measure of working memory capacity. The OSPAN task asks participants to remember a series of letters in sequence while completing math problems. Participants were first presented with a simple math problem, which they were asked to complete as quickly as possible. After responding to the problem, they were briefly presented with a single letter. After 3-7 problem-letter pairs, participants were asked to recall the letters in order and select corresponding boxes. Feedback was provided at the end of each trial. Letters/problems were presented in sets that ranged from 3-7 with each set size presented three times. A participant's score on the task is the number of letters recalled correctly, ranging from 0-75. A performance criterion of 80% correct was applied to the math portion of the OSPAN to ensure compliance with the task and 7 participants failed to meet this criterion. The OSPAN was used to measure working memory and yielded 139 valid responses. Scores ranged from 13 to 74 (out of 75), with a mean of 57.2 and a standard deviation of 12.2.

Analyses

Data were analyzed in linear mixed effects models to determine if response accuracy and response time were predicted by the testing condition (blocked or random), the set size of the initial stimulus presentation during learning (3 or 6), and working memory capacity. An additional set of analyses was performed using only data from the random condition to replicate the original study design (Collins, 2018). Accuracy was modeled using a logistic general linear mixed effects model to predict individual responses and change in accuracy from learning to test. All models contained random intercepts for participants and items as well as random slopes for all within-participant or within-item effects. Degrees of freedom were estimated using Satterwaite's method as implemented in the lmerTest R package.

Results

Accuracy

Mean response accuracy over the course of the 13 presentations of a stimulus is shown in Figure 3. Mean accuracy on small set size blocks was higher than large set sizes. The slope of the accuracy curve leveled out earlier in the small set size blocks, however both slopes approached 0 by the end of the block. These qualitative learning results replicate the same pattern of results observed in the original study (Collins, 2018). Analysis of the learning phase yields similar results as the original study. Given the focus of the current study is on the testing phase, we report analyses on the testing phase only. Mean accuracy at the end of learning and at test is shown in Figure 4.

Replication of Prior Study First we focus on the random testing condition and assess to what degree the results replicate those of the original study. One result from the prior study is that asymptotic learning performance in set size 6 was a significant predictor of performance at test for both set sizes and that asymptotic set size 3 performance was not a significant predictor of test performance. We found similar results with asymptotic set size 6 performance predicting both set size 6 test performance ($z = 5.56$, $p < .001$) and set size 3 test performance ($z = 2.10$, $p = .04$). However, set size 3 was not a significant predictor of test performance in either set size ($z < 1$).

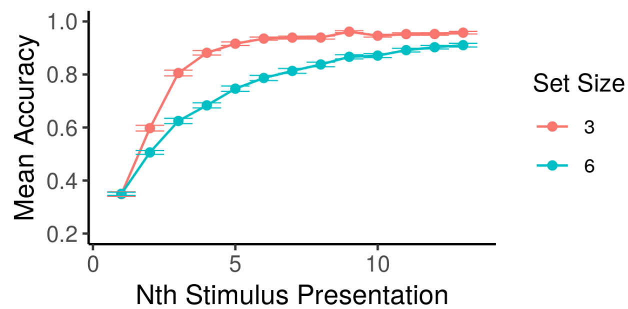

Figure 3: Mean accuracy across learning phase for set size 3 and set size 6.

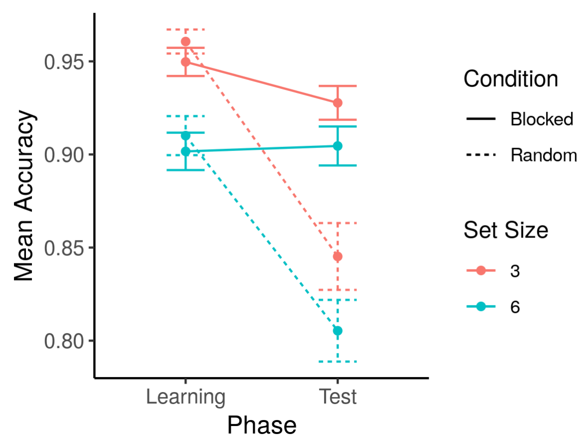

Figure 4: Mean accuracy at the end of the learning phase and during the test phase.

Next, the original study found that at test, items learned in set size 6 were recalled better than set size 3, but this effect was not replicated in our data using a simple model with only set size predicting test performance, $z = -0.39$, $p = .7$. However, a model with set size, asymptotic learning performance, and learning block showed two interactions: a set size by learning block interaction, $z = 3.33$, $p < .001$ and a set size by asymptotic learning accuracy interaction, $z = 5.31$, $p < .001$. The block by set size interaction was due to block recency having a greater impact on set size 6 accuracy than set size 3. The learning accuracy by set size interaction was due to set size 3 having higher testing accuracy at lower levels of learning while set size 6 had slightly higher accuracy at the highest levels of learning accuracy.

The learning accuracy by set size interaction suggests that the failure to replicate the set size difference may be due to variability in learning accuracy. When the simpler model predicting test accuracy from set size during learning is restricted to items with perfect asymptotic learning accuracy, the set size effect is significant, z = 3.26, p = .001 with set size 6 having higher accuracy (M = .89, SE = .01) than set size 3 (M = .85, SE = .01). In comparison to Figure 4, it can be seen that the set size 6 mean shifts to a greater degree when restricting the stimuli to only those well learned. Most of the results from the original study were replicated with the exception that the set size effect was dependent on learning accuracy, which was not observed in the original study.

Comparison of Blocked and Random Conditions Under the skill acquisition framework, performance for set size 3 should be greater than set size 6 in the blocked condition. First, we examined whether asymptotic learning performance was a predictor of test performance in the blocked condition. Asymptotic performance during the learning phase was used as a covariate to determine if reward history was predictive of performance at test. Results showed that asymptotic performance in the learning phase for set size 6 is predictive of performance in set size 6 at test, z = 5.79, p < .001, but not set size 3 associations. Asymptotic performance for set size 3 associations was not predictive of performance for set size 3 or set size 6 associations at test, z < 1.

Mean accuracy at the end of the learning phase was also compared to accuracy in the test phase by including phase (learning or test), set size, and condition as predictors of accuracy. The decrease in accuracy from learning to test was greater in the random condition than in the blocked condition, $t(137)$ = -6.49, p < .001. The larger set size also dropped less than the smaller set size from learning to test in both conditions, $t(137)$ = 4.19, p < .001. The three-way interaction between phase, set size, and condition was not significant. The steeper slope for the smaller set size is consistent with the original study, but given that the smaller set size was also recalled better than the larger, this result seems inconsistent with interference between working memory and reinforcement learning.

Contrary to our predictions for the blocked condition, associations learned in the larger set size showed a shallower decline in accuracy when compared to the smaller set size. In other words, there was not a three-way interaction with condition. The blocked condition decreased much less than the random condition, providing partial support for proceduralization of both set sizes. However, it is possible that the additional category cue in the blocked condition benefitted the retrieval strategy as well as the proceduralized rules.

Higher working memory scores were associated with higher accuracy during both phases, $t(136)$ = 3.82, p < .001). However, higher working memory also interacted with set size and phase such that higher working memory was associated with increased set size 6 accuracy more than set size 3, $t(137)$ = 2.65, p = .009, and increased accuracy more in the testing phase than the learning phase, $t(137)$ = 2.87, p = .005. This result seems inconsistent with the theory that working memory is interfering with reinforcement learning.

Because the delay from learning to test did vary across participants depending on OSPAN completion time, it may be possible to discriminate between the retrieval strategy and a proceduralization strategy by examining the effect of delay between learning and test. The time between the last presentation of the stimulus in the learning phase and the time it was presented during the testing phase was included in a model predicting testing accuracy. Delay should impact retrieval but not execution of a proceduralized rule. Also, from the interfering reinforcement learning and working memory theory, delay is not explicitly theorized to have any impact on associations learned via reinforcement learning.

In a model examining end of learning and test performance, delay was included along with phase, condition, set size, and working memory. In this case, the model was a logistic mixed effects model predicting the binary outcomes for individual stimuli because delays were stimulus specific. The model revealed that in both conditions, associations learned in set size 3 were affected less by the duration of the interval between the last stimulus presentation at learning and the first presentation at test, z = 5.44, p < .001. The mediating effect of stimulus presentation delay further suggests that associations learned in large set sizes are more of a function of memory retrieval while associations encountered in smaller sets are retrieved using a more robust proceduralized rule.

Response Time

Mean correct response time during learning is shown in Figure 5. The initial couple of presentations are faster likely due to a "guessing" stage until the correct response is learned. After that, response time was decreased slightly as the correct response became more practiced.

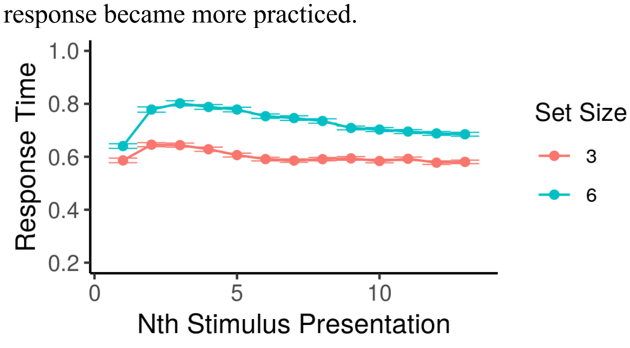

Figure 5: Mean correct response time in seconds across the learning phase for set size 3 and 6.

The primary effect of an item-specific proceduralized rule for responding would be to significantly increase the speed of response. For example, in ACT-R a production rule is executed in 50 ms, while retrieval from declarative memory would take about an order of magnitude longer (Anderson, 2004). It was hypothesized that a behavior produced by a rule would operate much more quickly than memory, given that declarative memory retrieval is not necessary. In the present study this translates into the prediction that response times would be substantially faster in the blocked testing condition, where a production rule might be utilized, than for the random condition. Similarly, it follows that associations learned in the smaller set size condition would result in a faster response than those learned in a larger set in the blocked condition, because it is more likely that a production rule would be compiled for the smaller set size than for the larger set size. In the random condition, we would not expect to see any significant differences between small and large set sizes, since the task demands are identical for the small and large set sizes and both would require retrieval from declarative memory.

Mean response time at the end of learning and at test was examined with a model including phase (learning or test), condition (random or blocked), and set size. Mean correct response time for correct responses once asymptotic performance was reached in the learning phase and correct responses in the test phase is shown in Figure 6. Consistent with the hypothesized results, response times increased more in the random than in the blocked testing condition from learning to test but the lower set size was affected the most as shown by an interaction between condition, phase, and set size, $t(135) = -7.62$, $p < .001$.

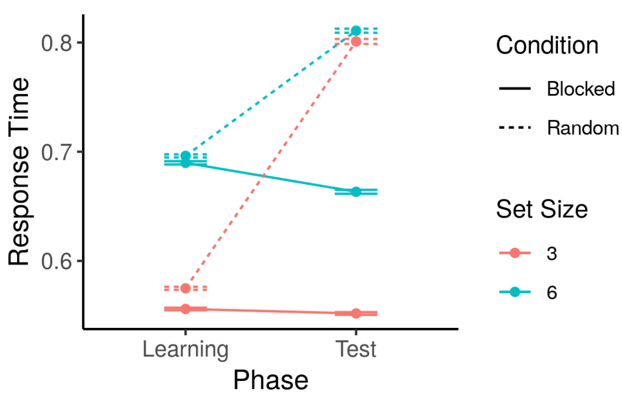

Figure 6: Mean correct response times in seconds at the end of learning as during the test phase.

An analysis similar to that done for accuracy was also carried out by predicting testing response time with asymptotic learning accuracy, set size, and block. In this model, higher learning accuracy led to faster response times, $t(187) = -6.76$, $p < .001$. The larger set size was associated with slower response times, $t(193) = 12.39$, $p < .001$. Stimuli learned in more recent blocks were responded to faster, $t(128) = -4.67$, $p < .001$. Stimuli in the random testing condition were also responded to slower, $t(172) = 9.49$, $p < .001$. However, again there was a larger effect of set size in the blocked condition than in the random condition, $t(162) = -13.1$, $p < .001$.

Working memory had a greater impact on response time in the random condition than the blocked condition ($t(126.5) = -2.70$, $p = .008$), however the interaction of working memory and set size was not significant. This finding is potentially consistent with a retrieval-based strategy being utilized more in the random condition as discussed below.

Discussion

The current study has significant implications for how we conceptualize the interaction between working memory, reinforcement learning, and skill acquisition. In the random testing condition, the current study replicated many of the findings from prior research (Collins, 2018) which posited that reinforcement learning and working memory are separate, competing processes. The main difference is that the lower set size only led to better test performance if the analysis was restricted to stimuli that were at ceiling at the end of learning. It is likely that there was simply higher variability in learning accuracy in our data than in the original study even though all participants met the performance criteria of 75% mean accuracy at the end of learning.

For accuracy, the results are largely consistent with the predictions from the skill acquisition framework. In the blocked condition, participants should be both faster and more accurate than in the random condition due to the use of proceduralized rules. In the random condition, participants should be relying more on a retrieval-based strategy and therefore take longer and be more sensitive to the delay between the last presentation of the stimulus during learning and test. Furthermore, the opposite effect of set size was observed in this study as in the original study (Collins, 2018) even when controlling for asymptotic learning accuracy. As noted, the original set size effect is only found when limiting analyses to stimuli learned perfectly. One possible interpretation of this set size effect is that these are exactly the stimuli that are well proceduralized at the end of learning. When stimuli are still being learned, they are still reliant on a retrieval-based strategy; therefore the larger set size lags behind due to the increased delay between stimuli [Figure 1b].

Particularly compelling are the differences in response times between set sizes in the blocked and random conditions, which may imply that smaller set sizes are being executed via a proceduralized rule rather than retrieved from declarative memory. At the very least, the response times suggest that the blocked testing condition is similar enough to learning to yield little difference in response times. Whether that is a function of proceduralization or an alternative theoretical explanation is still unclear.

Working memory, as measured by the OSPAN, had an overall effect on accuracy across all conditions and set sizes. In addition, working memory had a stronger relationship with response time in the random condition than in the blocked condition. Working memory as measured by the OSPAN is likely composed of multiple cognitive processes including attentional control and ability to more easily access long-term memories to bring them into working memory (Unsworth, 2016). It could be that the working memory relationships observed here are due more to the ability to access long-term memories. Some models of working memory have modeled individual differences as increased activation to spread in long-term memory to improve the accessibility of memories (Lovett et al., 2001). This kind of working memory mechanism would be consistent with the improved ability to retrieve the correct associations from declarative memory that was observed here in both learning and test behavior.

The proposed theoretical explanation for differences between the two conditions is informed by the ACT-R cognitive architecture. Parameters obtained from these behavioral data will be used to create an ACT-R model of the association learning task, allowing us to further explore whether differences in performance at test can be attributed to procedural knowledge. At the moment, results seem to indicate that when the context is similar enough, a proceduralized and reinforced rule is learned and used at test. However, it will be important to verify that the proposed mechanism actually accounts for the data.

In addition to developing a cognitive model of the task, future work will attempt to further examine the reinforcement learning process by introducing probabilistic reinforcement to the learning phase. It is hypothesized that associations with a higher value assigned in the learning phase will be retrieved more quickly in the test phase. If reinforcement prediction error is indeed present in associations learned in smaller set sizes, it would suggest that the association learning task is performed using reinforcement learning. Additionally, a repeated measures modification of the design might be used to address relatively unexplored issues such as skill decay.

Acknowledgements

This research was supported by an Office of Naval Research award (N00014-17-1-2324).

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An Integrated Theory of the Mind. *Psychological Review*, *111*(4), 1036–1060.

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. E., Albrecht, M. A., Waltz, J. A., Gold, J. M., & Frank, M. J. (2017). Interactions Among Working Memory, Reinforcement Learning, and Effort in Value-Based Choice: A New Paradigm and Selective Deficits in Schizophrenia. *Biological Psychiatry*, *82*(6), 431–439.

Collins, A. G. E., Ciullo, B., Frank, M. J., & Badre, D. (2017). Working Memory Load Strengthens Reward Prediction Errors. *The Journal of Neuroscience*, *37*(16), 4332–4342.

Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis: Working memory in reinforcement learning. *European Journal of Neuroscience*, *35*(7), 1024–1035.

Lovett, M. C., Daily, L. Z., & Reder, L. M. (2001). A source activation theory of working memory: Cross-task prediction of performance in ACT-R. *Cognitive Systems Research*, *1*(1–4), 99–118.

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*(2), 123–155.

Turner, M. & Engle, R. (1989). Is working memory capacity task dependent? *Journal of Memory and Language* 28(2), 127-154.

Unsworth, N. (2016). The Many Facets of Individual Differences in Working Memory Capacity. In *Psychology of Learning and Motivation* (Vol. 65, pp. 1–46). Elsevier.

Unsworth, N., Heitz, P., Schrock, J., Engle, R. (2005). An automated version of the operation span task. *Behavioral Research Methods*, 37(3), 498-505.