

# Training For Retention of Complex Mathematical Processes with a Focus on Contextual Interference Effects

By

John F. White

Department of Psychology, University of Colorado at Boulder

Defended April 3, 2014

Thesis Advisor:

Dr. Matthew C Jones, Department of Psychology

Defense Committee:

Dr. Matthew C. Jones, Department of Psychology

Dr. Alice F Healy: Committee Member, Department of Psychology

Dr. Donna J Caccamise: Committee Member, Department of Education

Dr. Ryan Bachtell, Honors council Representative, Department of Psychology

### Abstract

*In our current study, we investigated how acquisition and retention of performing a complex mathematical process is affected by the manner in which it is trained. We adapted three common training methods to promote the learning of a statistical concept and then tested to find the method of greatest immediate acquisition as well as greatest delayed retention of the concept. We decomposed the complex formula into training the simple steps, individually (blocked condition) and training the simple steps in the appropriate order (serial condition). The third condition trained the entire process in a single step (whole). Using stimuli common to all three conditions, we trained and tested learning of the complex mathematical process of computing the formula for finding variance. As this experiment conformed to the paradigm of investigating effects of contextual interference, we attempted to create conditions where contextual effects may aid in supporting long-term retention. Although this experiment failed to capture enough power to support any valid claims, we found trends that support effective learning techniques of complex processes. However, we found no effects of contextual interference.*

As we continue to explore deeper into the scientific concepts of our universe, we are forced to delve into more complex mathematical calculations in which to understand and explain those concepts. The need to effectively train the next generation of researchers to compute complex processes becomes more important as the calculations become more involved and specialized. This need for effective training practices crosses all scientific disciplines: from engineering and physics to statistical practices in biology and psychology, the need to teach sound and durable methods of mathematical computations is important to instilling good practices of future researchers.

In our experiment, we employed three instructional approaches for the training of a complex statistical formula and tested to find the approach that promotes the greatest retention of

the process. Each of the conditions used employs various degrees of intrinsic and germane cognitive load that can inhibit effective learning. The different aspects of cognitive load theory were explained by Paul Ayres, a researcher in the University Of New South Wales School Of Education, as he applied it on learning in the domain of mathematics. Ayres wrote, “Intrinsic cognitive load is the load placed on working memory by the intrinsic nature of the materials to be learned...while germane cognitive load is the load required for schema formation and automation” (Ayres, 2006, p. 287). Each of the three instructional approaches utilized in this experiment is structured to various degrees of cognitive load in an effort to observe different levels of retention performance. The process employed by our design should allow for differences in acquisition and retention of the learned material. This process of our design included five rounds of training pages, with all training followed by quiz questions that supported the subject trained. We then offered a posttest immediately after the five rounds of training and quiz questions. Our experiment’s design concludes with a retention test taken seven days after the initial part of the experiment.

All conditions in our experiment conformed to the same breakdown of the complex process; allowing for common step-wise parts to be trained differently and to varied levels of difficulty. The process of calculating variance was divided into five steps: finding the mean, calculating the deviations from the mean, squaring the deviations, summing the squared deviations, and dividing for the variance. Early versions of the experiment separated the process of finding the mean into three steps, but the data collected in those conditions found a ceiling effect in all steps for computing the mean. We decided that consolidating the three steps would increase the level of difficulty for finding the mean. With the process of calculating variance broken down into five steps, we believed we could teach each of these steps individually: we

refer to these segmented processes, throughout the paper, as “simple steps.” We then used these simple steps to build the different instructional processes that define the three conditions: blocked, serial, and whole.

In the blocked condition, the complex process of computing variance is broken down into a series of smaller tasks, each of which was trained and practiced separately. As participants trained and tested in this condition, they were allowed to concentrate on only one of the five processes at one time. Each round of training in the blocked condition had focused on one of the five simple steps that composed the formula for finding statistical variance. For example, the first round of quizzes after training focused on finding the mean from a set of scores; the second round of training focused on calculating the deviations from the mean for the scores. This continued through the fifth round of training which focuses on calculating variance from the sum of squares.

In the serial condition, the complex process of computing variance is decomposed into a series of the same simple steps as the blocked condition but trained and quizzed in sequential order in each round. Each of the training rounds offered a training slide that focused on all the steps to calculate variance, but the five quiz questions following each training page contained all five steps in order: first quiz question asks the participant to find the mean, second quiz question asks participant to calculate the deviations from the mean, through quiz question five which asks to calculate the variance.

The whole condition is not broken down for the participant. Each of the training pages was identical to the training pages in the blocked and serial conditions, but the quiz following the training page in the whole condition included only one quiz question which asked the participant to compute the formula for variance from a set of scores.

By intentionally allowing each instructional method to maintain different levels of complexity between the conditions, we can estimate where we should see evidence of contextual interference. William F. Battig, of the University of Colorado at Boulder, initially countered the prevailing belief that interference always led to negative transfer. Battig was convinced, however, that under certain circumstances, interference could lead to positive transfer (Magill & Hall, 1990, p. 242). Contextual interference is hypothesized to arise when a person performs multiple different tasks in succession or interleaved, in practice. Due to the complexity during practice, the context of each task (i.e., the other tasks performed before or after) interferes with the performance of that task. Richard A. Magill and Kallie G. Hall, of Louisiana State University, describe the effect of contextual interference as when “...several variations of a task must be learned during practice and where high levels of interference occur, such as randomly practicing the variations. While performance during practice is typically depressed, performance on a later retention or transfer test performance is better than for those practice conditions, in which there were lower levels of interference” (Magill & Hall, 1990, p. 242).

Magill and Hall were able to show the effect of contextual interference with their work on motor-skill learning. A later study performed here at the University of Colorado at Boulder found some effect of contextual interference when training for the cognitive learning of language acquisition (Schneider, Healy, and Bourne, 2002, p. 430). In their study, they found that the type of increased difficulty affects the amount of forgetting exhibited in a delayed retention test. They went on to support the ideas of Ayres as they found that the type of increase in difficulty can affect performance, such as increasing the difficulty of the task as opposed to increasing the difficulty of the subject matter learned. It should be added that they found, “...any manipulation

that increases the difficulty of a learning task may have different effects on initial and eventual performance” (Schneider, Healy, & Bourne, 2002, p. 439).

It is our intention to provide evidence that the amount of cognitive load affects acquisition and delayed retention. Ayres defined cognitive load as the amount of information to be processed in the working memory. He found that, “...overloading working memory inhibits learning and therefore in order to facilitate learning, cognitive load (working memory load) must be kept at a manageable level” (Ayres, 2006, p. 287). The focus of this study is to create good representations of the different structural approaches and compare the results for acquisition and retention. When describing the effects of cognitive load between conditions, we can infer several points that define load in each. The blocked condition has the least amount of cognitive load as the intrinsic load on working memory is reduced to learning several simple processes separately. The serial condition also lessens the burden of the intrinsic cognitive load when compared to the whole condition, as the process isolates each step and allows for simple calculations when quizzed after training. We believe the whole condition to contain the greatest amount of intrinsic load as participants were forced to remember the entire process in one step.

We believe that the performance means between conditions will be different, without knowing what that difference would look like. Looking at the results offered by Schneider, Healy, and Bourne (2002, p. 439. Table 5), we can infer that serial training should produce better results in posttests and retention tests than training in the blocked condition (in this supposition, I am assuming similarities between their experimental “mixed” condition and our “serial” condition). However, as the domains of their experiment and ours (language vs. mathematics) differed, we cannot assume identical trends. However, the study by Lim, Reiser and Olin (2009, pg. 71, Table 1) measured the effects of part-task versus whole-task approaches on a complex

cognitive skill and found a marginal benefit when training with the whole-task and testing with the whole-task as compared to training and testing with the part-task. Furthermore, we believe that the whole condition should show the greatest effect of contextual interference as this condition offers the greatest training complexity.

## **Experiment**

In our experiment, we trained participants in one of three instructional approaches for the complex mathematical process of finding statistical variance: whole, serial, and blocked. Immediately after training, we tested for acquisition of the material. The balance of participants returned to the lab seven days later to be tested on the complete process of finding statistical variance in a retention test without retraining.

### ***Methods***

#### *Participants*

Participants for this experiment were acquired from a pool of undergraduates in the psychology program at the University of Colorado at Boulder. The participants' level of prior experience in statistical methods ranged from none to prior training in high school and college. There were 26 participants across all six conditions, used in the final experiment, 18 of our participants performed in both sessions (part 1 and 2) whereas an additional 8 participated in training/ quiz questions and posttest only. Participants were supplied for this experiment from a pool of students taking general psychology who had a requirement to volunteer for lab participation; these students chose which experiments to participate in with little or no contextual knowledge of the type of experiments they were choosing beyond time of appointment and duration of the experiment.

### *Materials*

Unlike the initial pilot which was performed as a computer program, this experiment was conducted and scored on paper to maximize the opportunity to secure all work performed in the calculations when answering the training quizzes as well as the posttest and retention tests. The importance of capturing the participant's computations is detailed in the analysis subsection.

We decided to train and test participants on a subject that is relevant to our department, but conforms to our understanding of a mathematical process. In this experiment, we chose to teach and test on finding variance. We initially broke variance down into seven simple steps (finding  $n$ , summing scores, dividing for the mean, finding the deviations, squaring the deviations, summing the squared deviations, and dividing the sum of squares by  $n-1$ ), but adjusted the process to five steps as we considered "finding the mean" (steps 1-3) to be a simple process that compared better in difficulty to the other steps. This reduced the number of steps for calculating the formula for calculating variance to five.

The platform used to create the training, quiz, and testing slides was MS Office PowerPoint. All training slides, across all conditions, provided the definition and use of variance in statistical calculations and illustrated the entire process of finding variance from a set of scores. Each training slide also supplied an example of computing variance, broken down to the five steps that the quiz questions were built upon (1. finding mean, 2. finding the deviations, 3. squaring the deviations, 4. summing the squared deviations, and 5. dividing the sum of squares by  $n-1$ ).

For uniformity between conditions, five different training slides were created using two different layouts and a unique example for each. These five training slides were used in all conditions.



Variance $\sigma^2$	Example:
<p><b>What is Variance?</b>  <b>Variance</b> indicates how different the scores in a set are from each other. The lower the variance, the more similar the scores.</p> <p>Examine the formula for variance, then notice how the procedure is broken down:</p> $\sigma^2 = \frac{\sum(X-M)^2}{n-1}$ <p style="text-align: center;"><math>\sum = \text{Sum}</math>  <math>X = \text{Scores}</math>  <math>n = \# \text{ of scores in set}</math></p> <p>Looks pretty complex, but the individual steps are actually simple:</p> <ol style="list-style-type: none"> <li>Find the mean for the scores. (<math>M = \sum X/n</math>)</li> <li>Subtract the mean from each score, to find its deviation: <math>(X-M)</math>.</li> <li>Square each of the deviations: <math>(X-M)^2</math>.</li> <li>Sum the squared deviations together: <math>\sum(X-M)^2</math>.</li> <li>Divide by the number of scores, minus 1: <math>\frac{\sum(X-M)^2}{n-1}</math>.</li> </ol>	<p><b>Five mice ran a maze. Their times are below:</b>  <math>X = (4, 6, 8, 10, 12)</math></p> <ol style="list-style-type: none"> <li><b>Find the mean:</b>  <math>4+6+8+10+12 = 40/5 \Rightarrow 40/5=8</math></li> <li><b>Subtract the mean from each score to find their deviations.</b>  <math>4 - 8 = -4</math>  <math>6 - 8 = -2</math>  <math>8 - 8 = 0</math>  <math>10 - 8 = 2</math>  <math>12 - 8 = 4</math>  <b>(-4, -2, 0, 2, 4)</b>            (deviations from the mean)</li> <li><b>Square each deviation:</b>  <math>(-4)^2 = 16</math>   <math>(-2)^2 = 4</math>   <math>0^2 = 0</math>  <math>2^2 = 4</math>   <math>4^2 = 16</math></li> <li><b>Sum the squared deviations:</b>  <math>16 + 4 + 0 + 4 + 16 = 40</math> (sum of squares, SS)</li> <li><b>Divide the sum of squares by <math>n - 1</math></b>            (<math>n</math> being the number of scores in the set):  <math>\frac{40}{n-1} \Rightarrow \frac{40}{4} = 10</math>   <b>So, your variance for the set above is 10</b></li> </ol>

Figure 1 Training page #1 for all three conditions. Training slides within conditions varied by design, and each used different data in the examples.

Minor additions were included in the training slides which allowed the participant to focus on learning key aspects of the complex process relative to the training quizzes that followed it (if the quizzes following the training slide all asked for finding the mean, the training slide asked the participant to focus on the part of the training slide that explained finding the mean). Training quiz questions, posttest questions, and the delayed retention questions each included a word problem to supply context and a set of scores to base calculations upon as well as an area for the participant to supply a short, open answer. The posttest questions and delayed retention questions were identical in each condition; each offered a contextual basis in the form of a word problem and asked the participant to find variance from a set of scores. The word problems of the training quiz questions were identical across all conditions; however, each condition had a unique structure, which created different variations of how we organized the quiz questions:

Whole condition quiz questions were similar to the posttest and delayed retention test questions as each asked to find variance from a set of scores. Blocked and serial condition quiz questions broke the process of computing variance into five consecutive simple steps (mean, deviations of the mean, squaring deviations, sum of squares, and variance). Each of the simple steps had its own quiz questions and each part offered the correct answer from the previous question (when a quiz question asked for deviations from the mean, the mean was provided in the question).

**As part of a probability exercise, several participants were asked to flip a coin 100 times. Their scores were based on how many times their outcome was "heads." Sum the squared deviations from the mean for the scores below :**

**Andy: 44  
Charlene: 58  
Evan: 51  
Yvette: 52  
Valerie:47**

**For this question, we will provide the correct squared deviations from the mean:**  
 $(X - M)^2 = (25,81,4,9,4,49)$

**The sum of the squared deviations from the mean for the raw scores is:**

---

Figure 2. This training quiz question is found as the fourth quiz question in the first round of training in the serial condition and the first quiz question in the fourth round of the blocked condition.

The elegance in our design allowed us to use identical quiz questions for both the blocked and serial conditions, grouped differently: the quiz question that asked the participant to find the sum of squares for the first word problem could be found as the fourth question after the first training page in the serial condition and as the first quiz question after the fourth training page in the blocked condition. This was utilized to balance the level of difficulty between the conditions.

Included across all conditions were various transition pages and descriptor questions. Transition pages were used to prepare the participant for whatever task that followed. Transition

pages that followed training slides prepared the participant for the next set of quiz questions.

Transition pages that followed quiz questions informed the participant of the next training slide or of the following posttest. Descriptor questions were included at the end of the first part of the experiment and included prior experience in statistics in high school or college, prior experience with college level math, participant's major, and participant's gender.

### *Design*

Each of the conditions was assembled differently to adhere to how each condition would train statistical variance. The overall structure of the first part of the experiment, between the conditions, was consistent: written instructions, five rounds of training and training quizzes, posttest, and relevant descriptor questions.

The instructions, posttest, and descriptor questions were the same across all conditions. The greatest differences between conditions were the structure of the quiz questions:

The serial condition provided five quiz questions after each training slide. These five questions used a common word problem, and then broke down the process of finding variance into the five questions, asking participants to find the mean, deviations from the mean, squared deviations, sum of squares, and variance. After the first quiz question (to find the mean of the scores), the subsequent questions provided the correct answer for the prior step. This ensured that each quiz question prompted the correct answer even if the participant answered the prior quiz question incorrectly.

The blocked condition also provided five quiz questions after each training page. These five questions concentrated on one of the five simple steps used, by the experiment, to break down the process of finding variance. Specifically, after the first training page, the participant would be asked to calculate the mean on each of the five quiz questions, each with its own word

problem and set of scores. The following rounds of training pages and quiz questions focused on the remaining simple steps that compute variance, in their normal sequence (deviations from the mean, squared deviations, sum of squares, and variance).

The whole condition provided a single quiz question after each training page and asked the participant to calculate variance from a set of scores. The quiz questions in the whole condition employed the same word problems as the other conditions to minimize unintended differences between conditions.

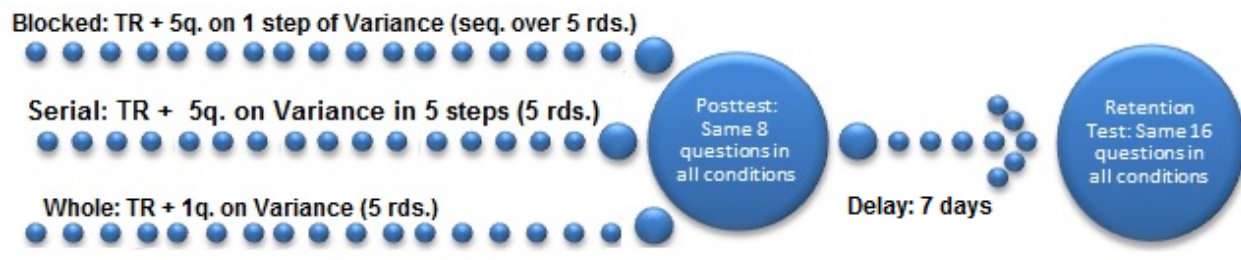


Figure 3. Illustration of the experiments progression. (TR = training page, 5q.= five questions, rds.= rounds). The 8 posttest questions and 16 retention test questions were identical between conditions, but questions were different between tests.

### *Procedure*

As the medium for facilitating this experiment was on paper, all copying and compiling of the experiment was performed prior to the participant's arrival. Each stack of pages (one experiment) was placed in an envelope and numbered for matching to the second part of the experiment. The envelopes were then stacked together, with an equal number of each condition represented.

For the first part of the experiment, participants arrived at the lab at their chosen appointment time and were given the appropriate informed consent form to initial and sign. Their names were checked against the online appointment system and their initials were logged next to their running number on the count sheet. These initials were used to ensure participants received the same running number when they returned to the lab for the second part of the

experiment. Participants were seated at a clean table with only a calculator and several pencils provided. Participants were offered instructions as a stack of half-sheets of paper was placed in front of them, “You will be studying a concept in math that is taught in this building. Please study the top sheet of the stack of papers in front of you. When you are finished with the top sheet, place it upside-down in a new separate pile. Once you have turned over a sheet, you cannot go back to it. The calculator and pencils are provided to make any calculations necessary to perform the experiment. When you have gone through the stack, you are finished with the experiment.” Participants were then allowed to proceed at their own rate until each had finished the experiment while monitored.

The second part of the experiment was performed almost identically as the first part for the participants. Once in the lab, participants again were seated at an open table, given the experiment (a small stack of papers from a manila envelope) and offered verbal instructions, “This part of the experiment is similar to the first part. You will be making calculations on the type of math you learned in part one. Feel free to show any of your work on the page you are working on. A calculator and pencils are provided for you.” Participants then completed the 16 questions, each computing variance from a set of scores provided. When the participant finished the experiment, the papers were placed back in the manila envelope for later scoring.

### *Analysis*

Overall, this experiment is of mixed design depending on which aspects we are analyzing. To study the data from the posttest, we employed a one-way analysis of variance (ANOVA) with a single between-subjects variable (condition). The same analysis was used to understand the data for the retention test. We then studied the interaction between condition and

test using a two-way 3x2 mixed factorial ANOVA, with one between-subjects factor (condition) and one within-subjects factor (test).

The data for the posttest and retention test were scored on whether the participant correctly performed the process of computing variance. In cases where the incorrect or no answer was recorded on the open answer line, but correctly exhibited the process in their written computations, then they were scored for a correct answer. To be clear, answers were scored as correct if either, a) the participant answered the question correctly, or b) the participant's notes and written work on any given question showed their complete understanding of the complex process of finding the answer. We chose this scoring method to identify whether the answers to the questions were computed with the correct knowledge of the process of computing variance. Those instances where the participant showed acquisition of the process, but committed rounding errors or supplied an incorrect answer in the space provided, could still be scored as correct as the objective of the experiment was to ascertain whether the complex process of computing the formula for variance was being learned. Data collected for the training quiz questions were scored in a similar manner but were only used to track learning trends from the training.

### ***Results***

Results were mixed as compared to our expectations, but all data are subject to being underpowered.

#### *Posttest*

We found almost no difference in the means between conditions in the posttest ( $n=26$ : blocked, .828, serial .875, whole .750:  $F(2,23) < 1$ ).

*Retention test*

The differences of the means between conditions in the delayed retention test were more varied ( $n=18$ : blocked .802, serial .531, whole .333:  $F(2,15) = 1.557$ ), but displayed high variability within conditions and lacked sufficient power for significance. Using Fisher's PLSD for a paired comparison, we found marginal significance in the paired comparison of the blocked vs. whole condition with a  $p$ -value of .0992.

*Across all conditions*

Main effect of condition:  $n=18$ :  $F(2,15)=1.274$ ,  $MSE= .171$ ,  $p=.3084$ . Interaction of conditions, tests:  $F(1,15)=1.439$ ,  $MSE= .102$ ,  $p=.2681$ . Main effect of test:  $F(1,15)= 4.363$ ,  $MSE=.102$ ,  $p=.0542$ .

As we compare performance between sessions (posttest and retention tests), we find a strong trend in the differences between tests ( $F(1,15)= 4.363$ ,  $p=.0542$ ) mostly from participant's poorer performance in the serial and whole conditions on the delayed retention test. Comparing the sessions (tests) in the blocked condition provided unexpected results as performance improved on the delayed retention test (blocked: posttest .771, retention test .802). Even though the differences in the means are too small for significance, it clearly displays no forgetting by the participants in this condition. Even with the lack of power to provide any significance to the results at present, we found this trend interesting enough to investigate how those in the blocked condition performed on the training quiz questions: data suggests that each of the participants in the blocked condition learned proficiently in all steps.

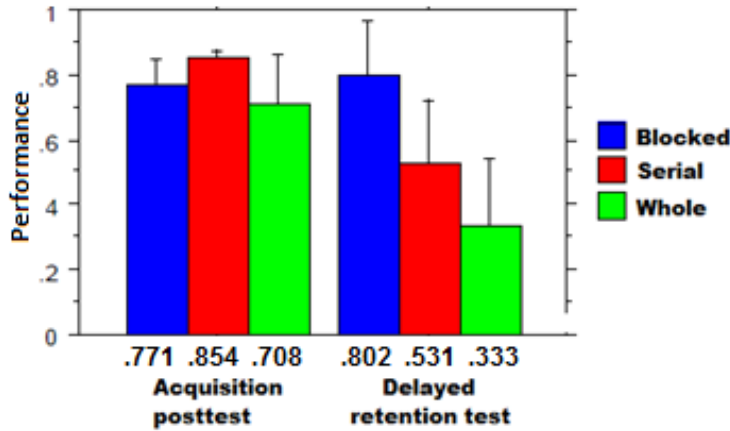


Figure 4: Comparisons of performance between conditions and tests with error bars. (  $n=18$  )

When we compare this to the results of the serial condition, we see that those of the blocked condition consistently performed better across all quiz questions, as opposed to serial training, where participants improved performance as they continued training (Figure 5,6). From this comparison, we can infer that acquisition of the correct process came later for those in the serial condition than those in the blocked condition.

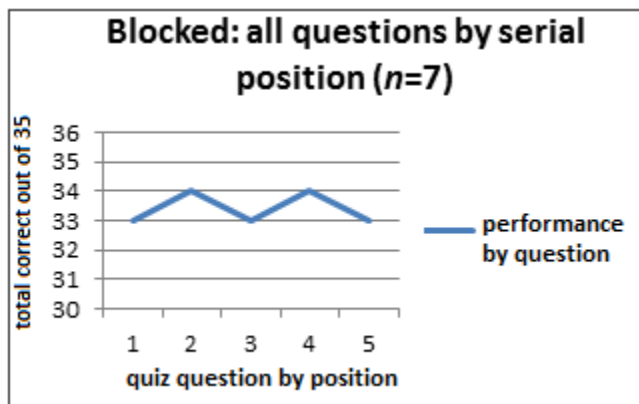


Figure 5. Total correct by quiz question number in the blocked condition



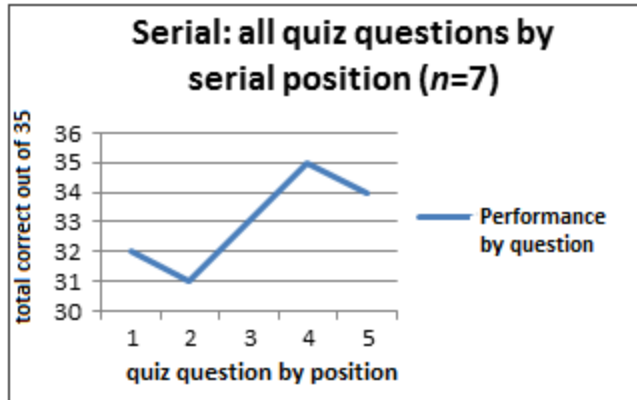


Figure 6 Total correct by quiz question number in the serial condition.

In an effort to understand potential reasons why the blocked condition, between tests, performed so well, we chose to examine what proportion of participants, within each condition, reported having prior training in statistics. We found that five of the eight (62.5%) participants had claimed prior statistical training in high school or college in the blocked condition as opposed to six of nine (66.7%) participants in the serial and five of nine (55.6%) in the whole conditions. When we compared the means of correct answers on the posttest for those who claimed prior training in statistics to those with no prior training (across all conditions) we found a significant difference between the two groups ( $F(1,24) = 4.644, p = .0414$ ).

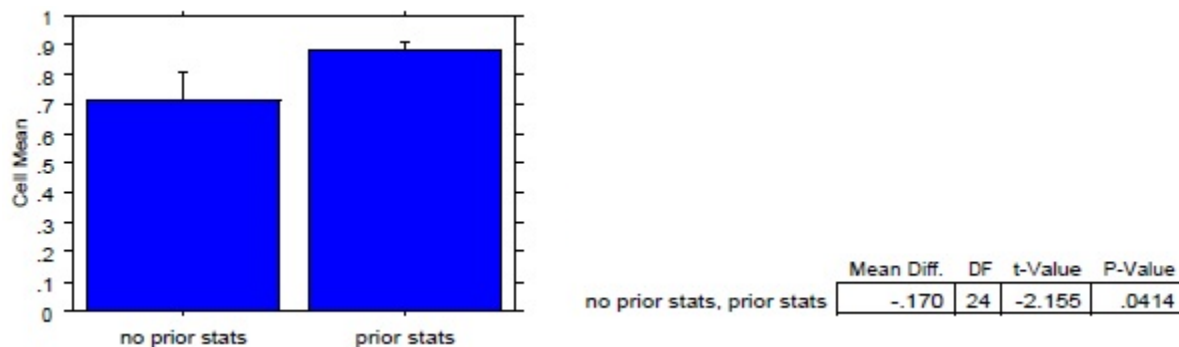


Figure 7. Differences between those who reported prior stats learning

But, as we consider the blocked condition with a smaller percentage of prior statistical learners as opposed to the serial condition, we see that if we control for prior experience, the differences

between the two conditions would increase slightly, which would not contribute to the unexpected results displayed by the blocked condition.

### *Discussion*

We relied on the participant pool for the spring semester in the psychology department to create the power we needed to provide enough participants for our experiment. Unfortunately, the turnout for the spring semester was much smaller than anticipated. Our initial estimates for power required that we would need at least 45 participants to facilitate a good experiment. As we were not able to get enough participation (at the time of reporting), we must discuss any conclusions as that of statistical trends without significance.

We found from the data that the greatest amount of retention when learning statistical variance was displayed in the blocked condition, as those who participated in the blocked showed no forgetting after a week delay between tests. Some speculation from our team considered that participants in this condition might have studied independently during the delay; although we cannot deny that this may have happened, we would be curious why participants in only one condition would have taken this extra step (considering that performance only increased between sessions in the blocked condition). We also noted that such actions by participants are usually minimized when conforming to good power standards (as individual tendencies are marginalized by greater participation in the experiment). In relation to the performance between sessions in the blocked category, it is also supposed that as we increase the participant totals in this condition, the trend of greater accuracy in the retention test would reverse to display some forgetting, thus providing a more similar trend to that of the other conditions.

As we compare our results to prior work on contextual interference, we find additional unexpected results. Prior investigations on contextual interference would have led us to believe

that participants would have performed best in the blocked condition on the posttest and better in the mixed condition on the delayed retention test. However, for our experiment, the data indicates a trend in the opposite direction: we found a slightly better performance on the posttest in the serial condition ( $M=.875$ ) than in the blocked condition ( $M=.828$ ) and better performance on the delayed retention test in the blocked condition ( $M=.802$ ) as opposed to the serial condition ( $M=.431$ ). We believe these results to indicate participants gained an advantage as they learned each step through repetition before being moved on to the following steps. This depth of acquisition supported more durable learning of the process when tested after a delay in the blocked condition. It is important to note that contextual interference requires a level of difficulty great enough to contribute to poorer performance in training or on the posttest during an experiment. With the effect of contextual interference, diverse training and poor results on the posttest lead to better results on a delayed retention test. With our experiment, we could not produce the effects of contextual interference. We can conclude that one of the conditions to create the effect of contextual interference may not have been met in our experiment. We believe that the difficulty of our training did not rise to the level of negatively affecting participants on the posttest. This supposition is made primarily in the blocked and serial condition as these two conditions are most similar to what is normally tested for contextual interference (the whole condition also displayed no effects of contextual interference as well).

When comparing our results to the effects of cognitive load, we found our data to support the theory that the intrinsic cognitive load of the blocked condition may be considerably lower than that of the other two conditions. As Paul Ayres defined, "Intrinsic cognitive load is the load placed on working memory by the intrinsic nature of the materials to be learned." (Ayres, 2006, p.287). He further explained that intrinsic load is one of three types of cognitive load and is

generated mostly by element interactivity. Fred Paas, of the University of the Netherlands, Heerlen, and his collaborators defined element interactivity as, “the way the individual elements of a task interact with other tasks, low in element interactivity, contain elements that do not react with each other, can be learnt in isolation, and require relatively low working memory load” (Paas, Tuovinen, Tabbers, & VanGervin, 2003). As the blocked condition of our experiment taught the participant to learn segments of the task in isolation, as opposed to the other conditions which trained to the entire process, we can conclude that some of the success of the blocked condition is directly contributable to a lower amount of intrinsic cognitive load within and specific to the condition.

### ***Future Directions***

It is important to complete this study by reaching a level of power which would allow for significance between means by condition and session. It is our intent to continue this study during the fall 2014 semester to at least double the participant count (based on the F-value of the ANOVA between conditions on the delayed retention test ( $F(2,15) = 1.557$ )).

As we found no measurable effect for contextual interference in this experiment, it is our intention to format a similar experiment to adhere to the conditions which would help to potentially trigger a contextual interference effect. Our primary focus will be to increase the complexity of our participant training. It has been a constant concern that our desired difficulty levels in training and training complexity may not provide an environment where contextual interference could be produced. Potential changes that are being considered include changing the focus of training to teaching standard error or replacing the serial condition to the more traditional method of a mixed condition. By training to standard error, we would increase the difficulty of the training as the number of steps to complete the complex process would increase.

This change may address the want to increase desired difficulty of training, but may contribute little to increasing the complexity of training as the processes in computing variance and standard error contain the same functions. To increase complexity of training, we will continue to investigate the feasibility of a more “mixed condition.” With prior studies on contextual interference, training could range in length and randomness of varied stimuli. In our experiment, we trained a specific mathematical process, with common steps for training in all conditions, trained in serial order in all conditions. If we tried to adopt more randomness to our training, such as potentially training variance in random order, we may create conditions that increase delayed retention of the entire process. However, a key component to teaching a linear process is teaching the linear order. More discussion and study will need to be given to this direction.

Another avenue which we want to pursue is adapting our experiment from an open answer format to that of multiple-choice. The original concept for this experiment was to be able to identify the best method for training complex mathematical processes using Iclickers in a classroom environment. These initial directions were based on experiments explained in an NSF grant proposal by Dr. Alice Healy and Dr. Matt Jones. Twice, we changed directions as we repeatedly faced a ceiling effect in our results. One of these directional changes forced us to adapt our experiment to a different answer format. As we fine-tune other aspects of our design, we will reopen the possibility of running this experiment in a multiple-choice format for its relevance to Iclicker use and adaptation to a classroom setting.

## References

- Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology, 20*, 287-298. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/acp.1245/pdf>
- Lim, J., Reiser, R. A., & Olina, Z. (2009). The effects of part-task and whole-task instructional approaches on acquisition and transfer of a complex cognitive skill. *Education Tech Research and Development, 57*, 61-77. doi: 10.1007/s11423-007-9085-y
- Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science, 9*, 241-289. doi: 10.1007/s11423-007-9085-y
- Paas, F., Tuovinen, J. E., Tabbers, H., & VanGervin, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71. Retrieved from <http://coursesite.uhcl.edu/HSH/PeresSc/Classes/PSYC6419seminar/Measurecogload.pdf>
- Schneider, V. I., Healy, A. F., & Bourne Jr., L. E. (2002). What is learned under difficult conditions is hard to forget: contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language, 46*, 419-440. doi: 10.1006/jmla.2001.2813