Chapter 3

METHODS AND PROCEDURES

Introduction

The goal of quantitative experimental research is to gather evidence that allows a reasonable conclusion to be reached as to whether or not a particular treatment causes a particular result. This should be done with a representative sample under carefully controlled conditions so that the conclusion can reasonably be generalized to a larger population.

Unfortunately, threats to data validity are commonplace. Experimental samples may be too small or may be made up of participants that do not accurately represent the larger population. Participants may be assigned to experimental groups in such a way that significant overall differences between groups can influence results. External forces may introduce interference that impact results. Natural growth and development in participants may lead toward inaccurate conclusions. Inconsistencies in the manner that treatments and assessments are conducted can produce unaccounted for variability in results. All of these threats can cast doubt upon a research study’s conclusions.

This specific research study attempted to implement good practices in experimental design in order to defend against such threats. The goal of these methods and procedures described below was to generate reproducible results that allowed for the formation of valid, generalizable conclusions in regards to techniques for long-term vocabulary acquisition. The description of these procedures will begin with the sample selection process, will continue through the research implementation plan, and will
conclude with an examination of the means by which the resulting research data were analyzed.

**Determination of Sample Size**

The researcher originally chose a minimum target number of 20 participants before conducting any *a priori* calculations that made estimates of the research study’s statistical power.

Subsequent statistical power estimates made using the application *GPower 3.1* (Buchner, A., Faul, F., & Erdfelder, E., 1992) supported the assertion that it was indeed possible to attain a statistical power of at least 0.8 given this research project’s design; even with a sample size as small as 18. This was reached by first specifying an effect size of .34 which is slightly higher than midway between the traditional large (.40) and medium (.25) effect sizes. When the power was calculated with an alpha value of .05 a value of 0.820 was predicted.

**Random Selection of Sample**

Some of the threats to validity that were mentioned earlier (such as non-representative sampling or groups created with significant differences) were mitigated in some respects by random selection and assignment. Critics’ attempts to attribute experimental results to other possible explanations (that is, reasons other than the experimental treatments) become much less viable when random sampling and designation techniques are utilized (Shadish, Cook, & Campbell, 2002). When utilized, these techniques also help protect the researcher from accusations of bias in the performance of the selection process.
Computer-based systems that select potential participants from a population and assign them to experimental groups offer speed, reliability, and efficiency. There may be some that are concerned, however, that computers’ random number generation processes are not purely random (Anthes, 2011); they are more accurately described as being pseudo-random. In real-world applications, this distinction would be much more significant in the realms of encryption and cryptography than in this realm of educational research. Computer functions producing numbers that are technically pseudo-random in nature are really more than adequate for use in research study methodology. Such numbers will from this point on simply be referred to as random numbers.

In an attempt, however, to retain the advantages of a technology-based solution while seeking to eliminate any possible random number issues in the actual selection process, the author developed a scripted *Filemaker Pro*-based utility called *ResearchDatabase*. *ResearchDatabase* leverages a computer’s ability to generate random numbers in order to take human action almost completely out of the initial stages of the process. The system then goes further by leaving the majority of the selection process to a table of numbers (Rand Corporation, 1955) that has undergone extensive statistical analysis and has been verified to be truly random (Brown, 1949).

*ResearchDatabase* allows a list of potential participants to be imported, the size of the population to be specified, and the size of the desired sample to be identified. During the importation process, this utility automatically assigns each member of the population an identification number incrementing from the number one to the total number of participants in the population.
Once the preliminary inputs were made, ResearchDatabase’s automated processes were activated. Filemaker Pro’s built-in random number generator was initially used to pick a column and location within the venerable table of one million random numbers (Rand Corporation, 1955). This initial selection is analogous to the commonly utilized method of pointing one’s finger at particular point within a print version of that random number table. The utility takes that point as a starting location and then uses another randomly generated number to determine if the system will begin to go sequentially up or down the list of random numbers. The numbers that were generated as a result of these processes were compared to the incremented identification numbers of the members of the research population. These individuals whose identification numbers matched the randomly generated numbers were, as a result, placed into the sample that was invited to participate in the research study.

The steps described above produced a representative sampling of the population that was presumed to be free of bias. It was presumed that reasonable generalizations were obtained about the target population in regards to responses to experimental treatments.

*Population and Sample*

The proposed study was conducted with students from a large midwestern public high school in the heart of Omaha, Nebraska. The age of individuals in this population ranged from age 14 years to 18 years within the school’s four grades (9-12).

From a list of the approximately 1900 students at this school, a sample of 60 students was randomly selected (see process described earlier). While 20 students was
the minimum target sample size, additional participants were sought for enlistment. This over-recruitment was done in an attempt to increase the probability that the target sample of 20 participants was reached.

**Institutional Review Board Procedures**

The author has received approval from the University of Nebraska-Lincoln’s Institutional Review Board (IRB 20120212385EP). Approval to conduct this research was sought and granted by both the assistant district superintendent and the high school principal. Procedures for obtaining individual permission from each participant and their parent/guardian were also approved by the Institutional Review Board.

These research activities took place outside of and independent of any other coursework at the participants’ high school. Within their Modular Schedule, nearly all of the students attending this school have Independent Study Time during which they have freedom of choice as far as activity and choice of location. Appointments for participation in the research activities were scheduled to take place in a designated room at the building either during students’ Independent Study Time, before school, or after school.

Since the research activities were conducted outside of class time with no relationship to any courses on students’ class schedules, there should have been no possible direct positive or negative impact upon any participant’s grades or standing at the participants’ high school. These conditions made participation truly voluntary. As a result, there was a concern that it would not be possible to recruit enough fully-participating individuals to reach a high enough level of statistical power. Although
every effort was been made to keep the number of participation sessions and the required
time of commitment to a minimum, issues with recruiting emerged. Efforts were made to
make participation in treatment and assessment activities as relatively convenient and
non-threatening as possible. It was hoped that these conditions would have eliminated
many of the obstacles that inhibited selected students from participating.

Students at this school are required to complete at least sixty hours of service
learning (community service) in order to graduate. Permission to offer two hours of
service learning to students who fully completed the research activities was granted by
the school’s Service Learning Coordinator.

Selected students were encouraged to participate for the sake of the study’s
validity, for the opportunity to gain new knowledge in regards to the study’s content of
focus, and for the availability of the two hours of service learning credit. Students and
their guardians were free, however, to make decisions about student participation without
any undue coercion or fear of academic repercussion.

Content Focus

This research study focussed upon a set of twenty-one vocabulary words. These
twenty-one words were chosen from a database of 5014 SAT review words that was used
by permission (Baba, 1999). The selected words consisted of ones that the researcher
himself did not immediately recognize during a scan of the database’ contents. The
assumption was made that uncommon words such as these that were not recognized and
immediately known by a doctoral candidate would not likely be initially familiar to a
random sample of high school students. Therefore participants’ baseline performance on the study’s pre-assessment was generally expected to be low.

It was hoped that this use of relatively rare vocabulary words would make it less likely that the participants would accidentally encounter them during the elapsed time between sessions. The chance that this type of interference took place during this intervening time should have been low.

For each participant a function in the online research system used computer-based random processes to assign seven of the twenty-one words to the retrieval-intensive treatment condition, seven to the review-intensive treatment condition, and seven to the control (no treatment) condition. The condition to which each word was assigned for a particular participant was recorded and tracked throughout the entirety of the research study. Performance on pairs within each condition were used to make the controlled comparisons of the investigated treatments.

Treatments

The research questions identified earlier focus upon a comparison of two different strategies for the long-term acquisition of new vocabulary words; a review-intensive strategy and a retrieval-intensive strategy. The treatments that each study participant experienced consisted of four precisely-timed cycles of exposure to fourteen vocabulary definition/word pairs using various mixtures of these two strategies. The order in which the pairs were presented during each of the four cycles was determined by computer-based random processes for each participant in an attempt to negate serial order effects.
During Cycle #1, each participant was shown each of fourteen vocabulary definition/word pairs one by one for twelve seconds in a web browser window. These pairs consisted of the words assigned to each participant that were within both the retrieval-intensive and review-intensive conditions. The display of each definition/word pair was separated by three seconds of a blank page prior to the next pair being shown. Cycle #2 was identical to Cycle #1.

During Cycle #3, the treatments diverged. Each participant was shown the seven definition/word pairs assigned to them in the review-intensive condition in precisely the same way they were shown them during the first two cycles (each pair for twelve seconds with a separating, blank three second interval). This was the review-intensive treatment.

For the seven remaining pairs assigned to the retrieval-intensive condition, the participant was shown the definition, given a blank editable text field, and be prompted to type in the word that correctly matched the definition displayed. This is the retrieval-intensive treatment. The intent of this was to force a retrieval attempt in each of these seven instances. The fact that recall attempts were triggered for some pairs but not for other pairs was the crux of this research study.

While the contents that the participant entered into the editable text field was not assessed and was not considered a measured variable, all of those entries were transmitted and saved in a MySQL database. These records were captured with a view to giving insight to the researcher regarding outlying data-points.

Cycle #4 consisted of exactly the same treatment as each participant experienced during Cycle #3.
The precisely-timed cycles, controlled user interface, randomized presentation order, and data verification checks were efforts to provide all participants with a uniform treatment experience. These precautions guarded the results against accusations of suspect validity.

**Variables, Measures, and Instruments**

This research study measured changes in one dependent variable that hypothetically resulted from manipulations of two independent variables. The dependent variable that was measured was the number of definition/word pairs that participants matched correctly. The two independent variables were the type of treatment and the amount of time elapsed from the start of participation in the research activities.

Changes in the dependent variable were measured by means of a web-based instrument developed by the author. The *Drag and Drop Matching Assessment* allowed the participant to choose any one of twenty-one vocabulary word objects, click upon it, drag it, and then release it in an area of their choice that corresponds to any one of twenty-one definitions (Figure 3.1).
The researcher chose this format of measurement and developed this instrument because of its concrete, quantifiable nature; that is, either the correct word was chosen to match a definition or it was not. There was little left open to subjective interpretation.

This matching format was chosen over a more traditional multiple-choice structure because of the increased level of variability. While a multiple-choice assessment will usually have three or four incorrect answers that have the possibility of being chosen, this format presented the participant with twenty possible incorrect responses thereby drastically decreasing the chances of merely selecting a correct answer at random.
Participants had 10 minutes to choose and place a vocabulary word next to each of the twenty-one definitions. A countdown clock was displayed within the instrument. If the participant did not answer prior to the elapse of these seven minutes, the instrument automatically did so. Any items left uncompleted were recorded with no answer.

The same assessment instrument was used by participants three times: for a preliminary assessment, for an assessment that was conducted immediately after the treatment cycles, and for a final assessment that was conducted at least one week after the treatment. The words and definitions, however, were placed in a different randomized order for each of these instances of assessment. The answer that each participant provided for each individual item during each of the assessments was transmitted and stored in an online MySQL database for later analysis.

Controls

One significant part of the experimental design was the use of seven control definition-word pairs. Every participant was given seven of the twenty-one words that were assigned to this condition. Every participant’s performance on these seven words was measured and tracked at each of the three assessment points; yet they were not referenced in any of the treatments for any participants. Performance on these seven words was expected to remain relatively constant throughout the course of the three assessments.

If significant changes in performance on these seven had been observed, this would have indicated that some type of interference or some type of maturation effect was unduly influencing results. Therefore, instances of purposeful outside study or
instances of participants randomly encountering words from the research study’s content focus should have been revealed by an examination of performance on these control words.

In an effort to help make sure that differences in the words themselves were not the cause of a change in the dependent variable, controls were also in place in the implementation of the other two (treated) groups of words. As described previously in the treatments section, each participant had words assigned randomly to the three treatment conditions. This random assignment of words to conditions controlled for differences in difficulty between the words in the set of twenty-one.

The following chart and key (Table 3.1) provides a visual summary of the treatments and controls utilized throughout the four timed cycles that each participant experienced. For example, during Treatment Cycle #3 participants received review-intensive treatment on seven randomly assigned word pairs, retrieval-intensive treatment on seven randomly assigned word pairs, and no treatment on seven final randomly assigned word pairs. An alternate view of this treatment schedule is in Appendix C.

<table>
<thead>
<tr>
<th>Participants</th>
<th>Treatment Cycle #1</th>
<th>Treatment Cycle #2</th>
<th>Treatment Cycle #3</th>
<th>Treatment Cycle #4</th>
</tr>
</thead>
<tbody>
<tr>
<td>S₇S, S₇T, N₇C</td>
<td>S₇S, S₇T, N₇C</td>
<td>S₇S, T₇T, N₇C</td>
<td>S₇S, T₇T, N₇C</td>
<td></td>
</tr>
</tbody>
</table>

S=Study(Review-intensive) T=Test(Retrieval-Intensive) N=No(No Treatment)
7S= 7 Words Assigned to each participant’s Review Condition
7T= 7 Words Assigned to each participant’s Retrieval Condition
7C= 7 Words Assigned to each participant’s Control Condition.

Table 3.1 Treatment and Controls Summary
Procedures

A listing of the current students at the school was exported out of PowerSchool, the district’s student information system. Each student’s name, grade level, gender, and home room advisor was imported into the previously described ResearchDatabase system. Once a target sample size of sixty was specified, the system was activated and randomly selected the participants for the study.

From the sample generated, a list was made that facilitated the distribution of informed consent forms (Appendix A and B) to participants and their parents/guardians through home room advisors. When participants returned the completed forms in the provided pre-addressed, stamped envelopes contact was made with the students through the district email system. It was through this email system that times and locations were agreed upon for each of the two research activity sessions that made up each individual’s participation.

Once participants received the informed consent forms, the ResearchDatabase system was used to generate single-component login credentials for each participant. A different login credential was created for each participant for each of the two sessions.

This information was imported into the MySQL database that served as the back end for the web-based measurement/treatment system. The data uploaded to this web database did not contain student names. There was no link between student results and student the student login credentials that were stored online. This procedure was designed with student privacy and safety foremost in mind.
The Research Database system generated the aforementioned student login credentials for the two sessions in barcode form. The system printed student names, the barcode version of session one login credential, and session two login credential into a layout formatted for the 3 by 10 sheets of Avery 5160 labels (Figure 3.2).

<table>
<thead>
<tr>
<th>Name</th>
<th>Session #1</th>
<th>Session #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>STUDENT NAMES OBSCURED</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
</tbody>
</table>

*Figure 3.2 Screenshot of Login Credentials*

These sheets of labels were placed in a 3-ring binder and were kept in a secure location when not being used by the researcher. The contents of this binder was the only link between participant names and participant research results that existed outside of the password-protected *ResearchDatabase* system.

When a participant arrived at the research session location for their first session, the label containing the barcode version of their Session #1 credentials was removed...
from the 3-ring binder and was placed on a 3 inch by 5 inch index card that was given to the participant. This action accomplished two purposes: it provided students with their needed login credentials in a secure form and it also produced an easily-checked visualization of the individual participant’s status in the research process. That is, the label was missing for a participant who had already participated in that session.

This label and index card combination was taken by the student to an iMac assigned to them for use during the research activities. The iMacs were automatically loaded up to an introductory page containing an overview of the study and an invitation to continue (Figure 3.3).

![A Study on Learning](image)

**Figure 3.3 Screenshot of Session 1 Intro Screen**

Once the participant read the intro and advanced to the next screen, they were given onscreen instructions on how to log in to the rest of the online research system (Figure 3.4). The text field for the entry of login credentials was pre-selected; the participant did not need to click into the text field. The participant used the provided USB barcode reader to scan the login credential barcode found on their 3” X 5” card. As
shown below, that text field was designated with a password attribute that prohibited content from being visually displayed. Inputs to this field were shown as only a series of dots for security reasons.

![Let's Start](Let's Start)

*Figure 3.4 Screenshot of Session 1 Login Screen*

Once the participant was logged on to the system for Session #1, a brief introductory screen was displayed. This screen provided the participants with instructions and also the time parameters under which this initial measure was to be completed (Figure 3.5).
A Study on Learning

In this research study, we are going to take a look at different ways to learn a set of vocabulary words. The first step will be to find out which of the vocabulary words you might already know. This will take a little time (you'll have 10 minutes to complete this part).

Look over the 21 definitions and the 21 words. Drag the word that best matches to the definition's corresponding spot. Please do all the words, even if you have to guess.

Figure 3.5 Screenshot of Assessment Launch Screen

Participants had 10 minutes to complete the Drag and Drop Matching Assessment (Figure 3.6). A countdown clock was provided in the assessment window’s title bar. During this time they needed to click upon the various vocabulary words that were provided, drag each one over the drop zone corresponding to a definition, and then release it. The vocabulary word object was turned from the initial color of blue to white when it was over a drop zone. When released, the dropped object automatically centered itself in the drop zone. A solid black line appeared around the object indicating that it had been properly placed into one of the possible definition drop zones.

The changes in appearance and behavior of the objects indicated to the participant that an object had been correctly placed into one of the possible drop areas. No feedback or indication of the correctness of the placement was provided.
At the end of ten minutes, the system automatically submitted the current state of the responses to an online *MySQL* database for later analysis. There was no feedback provided to the participant. The system immediately advanced to a screen indicating that the first assessment had been completed and that it was time to advance to the study portion of Session #1 (Figure 3.7).
The responses for the pre-test have been submitted.

It is now time to study the vocabulary words.

**Figure 3.7 Screenshot of Pre-Assessment Completion Screen**

The following screen (Figure 3.8) provided participants with what they would be encountering during this portion of the study as well as what they were expected to do.

**A Study on Learning**

Over the next few minutes, a number of different screens will automatically load on your computer’s web browser. These pages will sometimes show you a definition and then the word that goes along with the definition. Please pay attention to them and study them carefully while each of those pages appears on your screen for 12 seconds.

These pages will sometimes show you JUST a definition of a word followed by an empty box. When you see an empty box, try to remember which word goes along with the definition that is being shown. Then, as best you can, type in the one word that best matches the definition shown.

You have 12 seconds to type in the word before the screen will automatically advance to the next page. If you get done early, just relax and wait.

**Figure 3.8 Screenshot of Treatment Introduction**
When the treatment portion of the study actually began, the participants saw fourteen definition-word pairs one at a time as they were displayed in the browser window. See Figure 3.9 for an example of one of the screens. Each pair was visible for twelve seconds at a time with three seconds of blank screen being shown before the next one appeared. Cycles #1 and #2 were complete when all of the 14 pairs of words from that participant’s retrieve-intensive and review intensive condition had been shown to each participant.

![A luxurious person.](image)

**Figure 3.9 Screenshot of Review Example**

In each of Cycle #3 and #4, the treatment system showed seven pairs in the format displayed in Figure 3.9. The seven other pairs were displayed in the format shown in Figure 3.10 below. The gray-colored prompt “Type the Word” was displayed for only a brief amount of time. It disappeared automatically after just a little over a second when the cursor automatically was placed within the editable text field; ready for the participant to enter the results of their retrieval attempt. There was no need for the participant to click any type of submit button to have their entry recorded. When the
twelve seconds for that pair had elapsed, the entry was automatically submitted to the
MySQL Server.

It took approximately 15 minutes for all four treatment cycles to run their course. When
this was completed, the system automatically notified the participant of the completion of
this portion of the research activities (Figure 3.11)

The STUDY portion has ended. Time for one more test.

Figure 3.10 Screenshot of Retrieve Example

Figure 3.11 Screenshot of Treatment Conclusion
The first session of research participation concluded with another Drag and Drop Matching Assessment. Identical in content and format to the earlier one (see previous Figure 3.6), this assessment attempted to measure the immediate effects of the treatments the participant had experienced.

The order of both the definitions and the matching words was in a different order than the previous measure. The allotted time was reduced by three to seven minutes. This reduction was based on the assumption that once a familiarity with the assessment format had been established participants would not need as much time to complete the matching of the twenty-one pairs.

When the seven minutes of allotted time had elapsed, participants’ choices were again automatically submitted to the MySQL database for later analysis. Students were reminded to return one week later to complete their participation in the research (Figure 3.12). The 3” x 5” index cards with login credential were shredded.

Figure 3.12 Screenshot of Session 1 Conclusion
In order for the research data to be valid it was be of utmost importance to have participants return for the final assessment to be held one week after the first session. During the intervening week, email reminders were sent to participants reminding them of the time and location of their final participation session (Figure 3.13).

![Reminder Email]

From: Paul Lindgren
Subject: Reminder
To: <name of student>
Cc:

Thank you for being willing to participate in the research study on vocabulary learning.

Please be reminded that you are scheduled to take one more brief test before you are finished.

Please show up to the Warrior room on <day of week>, <month>, <date>:

- Between 7:15 and 8:00 AM
- OR
- During one of your open mods
- OR
- After school before 4:30PM

Thanks again!
Paul Lindgren
Westside Technology Coordinator

Figure 3.13 Screenshot of Sample Reminder Email

When participants reported one week later for their concluding session, they were again given a 3” X 5” card with a label containing their login credential in barcode form. The barcode reader-equipped iMac again had its web browser pre-loaded with an introductory screen that let the participant know what they had already completed in the study and what they had left to compete (Figure 3.14).
Welcome back, and thank you for your willingness to participate in this research study focusing on learning.

Your participation in the study is nearly complete. You have finished three of the four (4) parts:
1. A test at the beginning. DONE
2. A period of study. DONE
3. A test at the end. DONE
4. A delayed test (this is what you will be doing today).

Figure 3.14 Screenshot of Session 2 Introduction

The login screen and procedure was similar to that which was used during the first session; the participants again used the barcode reader to securely enter their credentials into the password field of the login page (See earlier Figure 3.4). When logged in, participants were presented with one last introductory screen that informed them of the assessment task ahead of them and the time in which they were allowed to complete it.
A Study on Learning

In this research study, we have been taking a look at different ways to learn a set of vocabulary words. This last step will be to find out which of the vocabulary words you know right now. This will take a little time (you’ll have 7 minutes to complete this part).

As before, look over the 21 definitions and the 21 words. Drag the word that best matches to the definition’s corresponding spot. Please do all the words, even if you have to guess.

Click here to Start

Figure 3.15 Screenshot of Begin Final Assessment

Other than the order of the items, the assessment was be identical to the post-treatment assessment, giving participants seven minutes to do their best to match twenty-one word objects with twenty-one definitions (see earlier Figure 3.6). When the allotted time elapsed, the responses were automatically submitted to the MySQL database for later analysis. A screen communicating the completion of their responsibilities as a research subject and thanking them for their participation appeared (Figure 3.16).
Your responses for the post-test have been submitted.

You have finished. Thank you for your participation in this study on learning.

Figure 3.16 Screenshot of Study Conclusion

At this point the researcher provided the participant with a signed Service Learning (community service) form from the participating school. After the student completed the form, added a brief personal reaction, and signed it themselves, the form was turned into the school’s Service Learning Coordinator for credit toward their Service Learning requirement. The 3” x 5” index cards with login credential were shredded.

Data Analysis

Once participants completed the pre-assessment, treatments, post-assessment, and final assessments a two-factor repeated-measures ANOVA was conducted comparing the differences between the three treatment conditions over the course of the three assessments.

A summary table shows how the nine different sets of scores was assembled. Each of these scores indicated performance on matching seven of the twenty-one definition-vocabulary word pairs (Table 3.2).
Table 3.2 Structure for Data Analysis

The preliminary test of the three was obviously essential to laying the groundwork for data validity. It established content knowledge baselines for participants that were relied upon during the concluding statistical analysis. Statistical indications of change from those baselines in the subsequent assessments were evidence of effect in regards to the various treatment conditions. Similarly, observed changes in regards to the control conditions...
content were important evidence concerning the possibility of interference, maturation, regression, testing effect (ironically), or another validity issue.

While there was a great deal of information included in this data analysis, there were two primary points of attention. The first was the scores generated by the post-assessment for the words in both the retrieve and review conditions. The second was the scores generated by the delayed assessment in both the retrieve and review conditions. The other points analyzed in the ANOVA were noted primarily as contextual information and as checks for threats against data validity.