Sample Psychology Internal Assessment Assignment

Examiner's Marks and Comments

Mark	**Introduction (max 6 marks)**
5	The theory upon which the student's investigation is based is described and the link to the student's investigation is made, but could be more fully explained. (5-6) The aim of the investigation is stated, and its relevance to teachers and to student learning is explained. (5-6) The independent variable (IV) and dependent variable (DV) are stated and fully operationalized in the null or research hypothesis. The IV is operationalized in the research hypothesis and the DV is operationalized just below the research hypothesis, which is acceptable and still achieves a mark in the top markband. (5-6)
	Exploration (max 4 marks)
4	Independent samples research design is explained. (3–4) Opportunity sampling technique is explained. (3–4) The choice of participants is explained. (3–4) Controlled variables are explained. (3–4) The choice of materials is explained. (3–4)
	Analysis (max 6 marks)
5	Descriptive and inferential statistics are appropriately and accurately applied. (5–6) The graph addresses the hypothesis and is correctly presented, although it contains a minor error in that the y-axis should go from 0–10, as that was the possible range of answers. (5–6) The statistical findings are interpreted with regard to the data and linked to the hypothesis. (5–6) Note that the student has interpreted the data correctly, but the term "critical value" could cause some confusion. The online program used states that the observed value should fall above or below certain parameters. The correct term for the upper parameter is the upper limit, not the critical value. (Critical values are found in tables of critical values, and for the Mann–Whitney U test the observed value should fall below the critical value as shown in the table.)
	Evaluation (max 6 marks)
5	The findings of the student's investigation are discussed with reference to the levels of processing model. (5–6) The strengths and limitations of the design, sample or procedure are stated and explained and relevant to the investigation. (5–6) The strengths and limitations of the procedure could be made more relevant to the student's investigation—for example, lack of ecological validity in an experiment is a general comment. Modifications are explicitly linked to the limitations of the student's investigation and mostly fully justified. (5–6)
	General Comments
Total Marks 19/22	A clear and coherent report with very few errors. References and appendices are appropriate. Evidence is presented of ethical guidelines being followed.

A Partial Replication of Craik and Tulving's Level of Processing Study (1975)

Psychology May

Word Count: 2157

Table of Contents

Introduction

The Levels of Processing (LOP) Model of Memory was proposed by Craik and Lockhart in 1972 and suggests that recall depends on the depth of processing that information undergoes during encoding (as cited in Popov, Parker, & Seath, 2017, p. 136). There are two main levels of processing: shallow processing and deep processing. Shallow processing only involves taking superficial features of a stimulus into account and is what is occurs during rehearsal. Structural processing (taking the physical properties of the stimulus into account) and phonetic processing (taking the acoustic properties of the stimulus into account) are both types of shallow processing. Deep processing, however, occurs in the form of semantic processing and involves making meaningful connections between the stimulus and prior knowledge. Craik and Lockhart therefore suggested that the deeper that information is processed, the longer its trace in episodic, long-term memory. The theory also serves as an extension to the Multi-Store Model of Memory as it suggests that rehearsal is not the only way to consolidate information from the short-term store to the long-term store, and that long-term memory is dependent on how the information was processed during encoding.

In 1975, Craik and Tulving decided to test the LOP model through a lab experiment with a repeated measures design (Craik & Tulving, 1975, pp. 268-294). The aim of their study was to investigate how shallow processing (structural or phonetic processing) and deep processing (semantic processing) affect memory recall and recognition. Participants were shown yes-or-no questions about words and then shown the word for 200 milliseconds. There were three types of questions, each encouraging a different kind of processing (structural, phonetic, or semantic). Participants then had to indicate their response to the question by pressing either a yes-button or no-button. This process repeated itself 48 times. Afterwards, participants were either asked to complete a free-recall task or a recognition task for the words. The researchers

found that memory recall and recognition of the words was significantly better for those that were preceded by a question that encouraged semantic processing. For example, the average percentage of words correctly recalled was 8% for structural processing, 9% for phonetic processing, and 22% for semantic processing. These results therefore supported the LOP model.

The aim of our study was to investigate whether or not semantic processing increases the recall of a list of 10 words among international, multilingual teenagers in comparison to structural processing. The importance of memory in education makes the LOP model important for students and educators. Understanding which type of processing leads to the best recall will allow students to improve their study techniques and allow educators to use more effective teaching methods.

The null hypothesis was that there will be no significant difference in the number of words that participants recall (from the set list of 10 words) after structural processing in comparison to after semantic processing.

The experimental hypothesis was that participants will recall significantly more words (from the set list of 10 words) after semantic processing than after structural processing. Structural processing was achieved by asking participants questions relating to the physical properties of words (e.g. Is this word written in red?) whereas semantic processing was achieved by asking participants questions that encourage them to make connections with prior knowledge (e.g. Is this word a type of body part?).

The independent variable in this study was the type/level of processing (either structural or semantic) that participants underwent prior to seeing a word and the dependent variable was the number of words recalled in a free-recall task.

Exploration

The research design of this experiment was independent measures. Our sample was therefore equally split into two conditions. Each participant hence only completed the experiment once for their allocated condition. It was most appropriate to use independent measures for our study because if participants took part in both conditions, the temporal separation of the conditions would make the differences between the types of questions obvious, thus causing participants to be affected by demand characteristics and decreasing the validity of the results. In addition, independent measures allowed us to use a standardized list of 10 five-letter long words in both conditions (see Appendix 1) and prevent order effects from impacting participants. By creating a standardized word list, we were able to eliminate the possibility that different word lists accounted for the differences in participant recall between the two conditions and thus isolate our independent variable. The word list was created with a random word generator (see Appendix 1) in order to prevent researcher bias. All words were 5 letters long to ensure that differences in word length did not cause our results.

Through opportunity sampling, we obtained 34 participants (20 females; 14 males). The participants were 16-17 year old IB students at an international school who were in their humanities class at the time when we conducted our experiment. By having teachers provide their students, we were able to easily find participants. Opportunity sampling was the best for our use as it was convenient in terms of finding participants under time constraints, and effective as we were able to collect all of our results in a short amount of time. Using a random group generator, we split our participants into two groups and randomly allocated a condition to each group. As a result, the structural processing condition had 9 females and 8 males, and the semantic processing condition had 11 females and 6 males. No participants were color-blind since the ability to perceive color was necessary to answer the structural

processing questions relating to word color. Although we were not assessing the accuracy of the participants' responses to the questions, by only having participants who could perceive color, we were able to ensure that differences in color perception did not impact the processing of words and recall. All of the participants were fluent in English, which was important as the experiment was conducted in English only.

Whilst conducting our experiment, we read our Standardized Instructions (see Appendix 2), which ensured that our results occurred due to varying levels of processing and not differences in instructions across both conditions. After collecting Informed Consent Forms (see Appendix 3), we handed out Yes/No Answer Sheets (see Appendix 4) for participants to fill out during the question session. We then started our PowerPoint for the given condition. Each PowerPoint contained 10 questions that lead to either structural or semantic processing (see Appendix 5 and 6). Each question was followed by a word from our word list (see Appendix 1). The PowerPoints were programmed to make a ringing sound every time a new question appeared on the screen so participants knew to look up and read the next question. They were also programmed to show the question slide for 8 seconds, the word slide for 2 seconds, and the slide telling participants which question to answer for 8 seconds in order to eliminate human error and ensure that time inconsistencies did not affect our results. Once participants answered all 10 questions, we collected the answer sheets and handed out blank paper for the free-recall task. Participants were given 2 minutes to complete this task, allowing us to ensure that differences in recall were not caused by differences in the time for the free-recall task. We then collected the sheets and gave each participant a debriefing form (see Appendix 7 and 8) and a contact sheet (see Appendix 9) containing our emails and our supervisor's email for further inquiries. We then repeated this process for the other condition.

Analysis

Descriptive Statistics

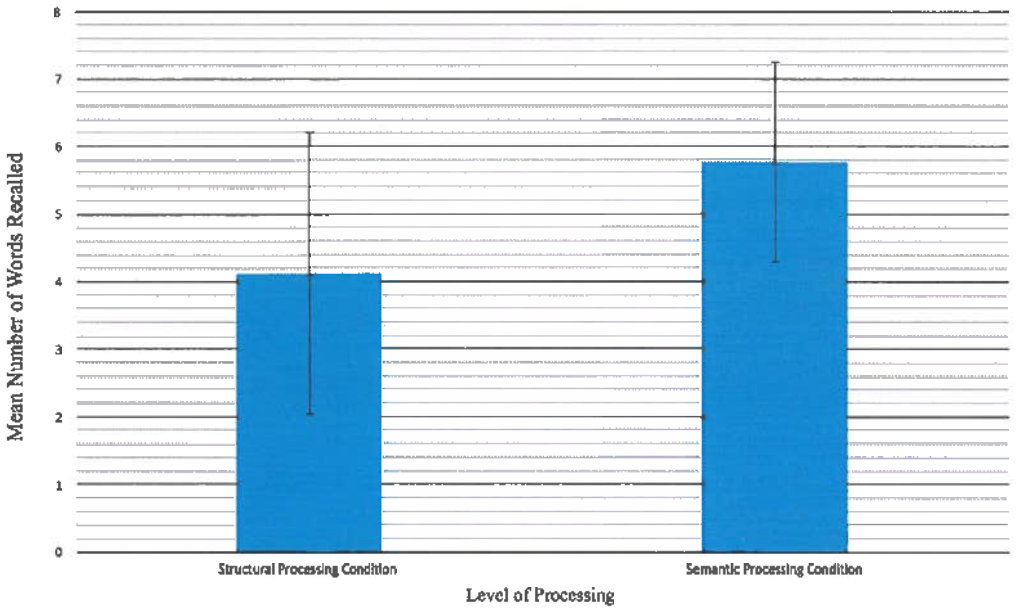
Table 1: Showing the Mean Number of Words Recalled and the Standard Deviation for the Structural Processing Condition and the Semantic Processing Condition

Level of Processing	Structural Processing	Semantic Processing
Mean Number of Words Recalled	4.12	5.76
Standard Deviation (# of words)	2.09	1.48

(see Appendix 10 for raw data and Appendix 11 for calculations)

The processed data shows that the mean number of words recalled in the structural processing condition was less than the mean number of words recalled in the semantic processing condition (4.12 compared to 5.76). The structural processing condition has a higher standard deviation (2.09) than the semantic processing condition (1.48). The number of words recalled by participants in the structural processing condition was therefore more spread around the mean in comparison to the semantic processing condition and hence less reliable. Since the means of the conditions are so close in value together (difference of 1.64) the standard deviations (2.09 and 1.48) suggest that there is overlap between the data sets.

Graph 1: Showing the Mean Number of Words Recalled for the Structural Processing Condition and the Semantic Processing Condition with Error Bars Representing the Standard Deviation for Each Condition



Inferential Statistics

In order to further process the data, a Mann-Whitney Test was carried out. This test uses ordinal data so upon inputting our ratio data into the online calculator, it transformed our ratio data into ordinal data. Once the test was carried out (see Appendix 12), the following conclusions could be made. As the observed value (UA=213.5) is greater than the critical value (=212) for a one-tailed test, we can reject the null hypothesis at the $p \leq 0.01$ level. This means we are 99% certain that the number of words recalled was only affected by the level of processing instead of pure chance.

As a result of our study, we can conclude at a $p \leq 0.01$ level of significance that semantic processing of words has a positive effect on international teenagers' recall of 10 given words.

Evaluation

Overall our study ultimately reached the same conclusion as Craik and Tulving's study: semantic processing leads to better recall than structural processing (Craik & Tulving, 1975, pp. 268-294). By controlling several variables, we were able to reduce the impact of extraneous variables on our dependent variable. It is therefore more likely that our conclusion was caused by the changes in the level of processing, hence supporting the LOP model. With relation to the model, participants in the structural processing condition only took in superficial characteristics of the words that had to do with appearance (e.g. color, typography) in order to answer the questions. Participants in the semantic processing condition had to make connections between the words and previous knowledge in order to answer the questions. This led the participants in the semantic processing condition to process the words at a much deeper level than those in the structural condition, which was evident by higher memory recall.

Conducting our study with an independent measures design was advantageous in several ways. Firstly, it allowed us to use a standardized list of words for both conditions, thus enabling us to eliminate the possibility that our results were caused by differences in words across the two conditions, and increasing the internal validity of our study. This design also prevented our participants from being affected by order effects, thus increasing the reliability of our results. In addition, since participants only took part in one condition, they would not see the differences between the types of questions, thus eliminating demand characteristics. However, a limitation of the design is that individual differences (e.g. differences in memory abilities) could have caused the differences in our results, thus decreasing the internal validity. Furthermore, even though participants were randomly allocated into conditions, there is a chance that the participants in one condition ended up having a stronger memory

than those in the other, hence decreasing reliability. A modification would therefore be to use a matched pairs research design, pairing participants according to their memory abilities (determined from a preliminary test) and splitting them accordingly. Although this would be more time consuming, it would increase reliability.

A strength of our sample was that all of the participants were fluent in English, thus allowing us to ensure that they all understood our instructions and the questions/words in the PowerPoint. We therefore know that our results were not caused by the lack of understanding or confusion of the participants. However, since our sample only consisted of 16 to 17-year-old students from a single international school, it is not representative, thus limiting generalizability. In order to rectify this, the study could also be conducted at other schools (IB and non-IB) with participants in the same age range. In addition, since our sample consisted of humanities students, those who were familiar with psychology and memory experiments could have guessed that a free-recall task would be conducted. As a result, they could have made a greater effort to remember the words, hence leading to inconsistencies in our results. A modification would be to pre-screen participants, ensuring that they are unfamiliar with psychology and memory experiments.

Overall, our procedure had several strengths as it allowed us to produce results that were applicable to our aim and ended up supporting our experimental hypothesis. Since we did not reveal our complete aim to our participants, they were not affected by demand characteristics, hence increasing internal validity. Another strength was that our standardized instructions were clear and seemed to be understood by all participants judging by their performances. This also increased internal validity as we were able to ensure that differences in directions did not affect recall. Despite these strengths, a limitation of the procedure was that the task

was artificial and not representative of real-life learning situations, hence reducing ecological validity. A modification would be to have participants do a more realistic task such as answering questions about ads.

In general, it is difficult to study memory as it cannot be directly observed. Therefore, we cannot be sure that participants actually underwent the level of processing of their condition and that they only underwent this level of processing. Nevertheless, by controlling several variables, we were able to reach the conclusion that semantic processing led to better recall in our sample and hence support the LOP model.

References

Craik, F. I. M., & Tulving, E. (1975). Depth of Processing and the Retention of Words in Episodic Memory. Journal of Experimental Psychology, 104(3). Retrieved from http://alicekim.ca/CraikTulving1975.pdf

Free Calculators and Convertors. (n.d.). Retrieved September 25, 2018, from https://www.easycalculation.com/statistics/mean-median-mode.php

Lowry, R. (1998). Website for Statistical Computation. Retrieved September 27, 2018, from VassarStats website: http://vassarstats.net/index.html

Popov, A., Parker, L., & Seath, D. (2017). Psychology Course Companion (2nd ed.). Oxford, UK: Oxford University.

Random Word Generator. (n.d.). Retrieved September 18, 2018, from https://randomwordgenerator.com/

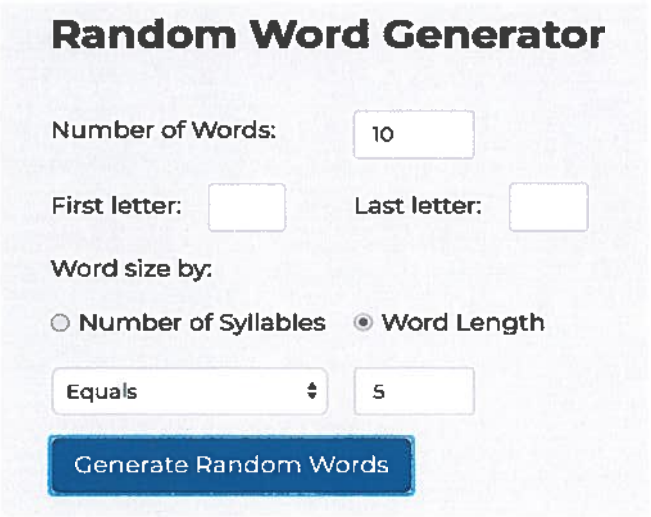
Appendices

Appendix 1: Word List

Word List

raise
wagon
shark
apple
study
track
beard
total
rumor
relax

This word list was obtained by using the website https://randomwordgenerator.com/ and setting it to the following settings:



Appendix 2: Standardized Instructions

Good morning. Our names are xxxxxx, xxxxxx, xxxxxx, and xxxxxx and we are conducting an experiment for our Psychology IA.

First, we will give you an informed consent form, which you should carefully read. If you agree to participate in this study, please sign it. If you do not want to participate, please leave now. If you have any questions concerning the experiment, please do not hesitate to ask. When you are done filling out the form, please put your pens down and look up.

(wait until everyone is done)

We will now collect the informed consent forms.

(collect the forms, look through and check that nobody is color-blind and not fluent in English)

The aim of our experiment is to investigate memory and it should take about 10 minutes. We will now hand out sheets of paper that contain 10 yes or no answer options. Please do not write anything on them until instructed to do so.

(hand out the answer sheets)

This is how to experiment will proceed. First when we start the PowerPoint, a ringing sound will occur. This sound will indicate that there is a question on the screen. The question will be displayed for 8 seconds and will ask you about a word. After these 8 seconds have passed, the word in question will be displayed on the screen for 2 seconds. Once these 2 seconds pass, the screen will indicate which question you should answer on your answer sheet. You will have 8 seconds to circle a yes or no response to the question on your answer sheet. After these 8 seconds have passed, a ringing sound will occur. When you hear this ringing sound,

make sure to look up at the screen as a new question will be displayed. This process will repeat 9 additional times until all 10 questions are answered. Please remember that all of this must be completed individually and in silence.

Does everyone understand the process? If you are confused, please let us know as it is vital that everyone knows what they are supposed to be doing.

(answer any questions if relevant)

The first question will momentarily show up on the screen. Please remember to look up at the screen whenever you hear a ringing sound as this will indicate the start of a new question and circle your answers on your sheet. It is important that you complete this task on your own and in complete silence. We will now start the experiment.

(start the PowerPoint; experiment takes place)

(PowerPoint is now over)

We will now collect your answer sheet and provide you with a sheet of paper. Please remain quiet and do not write anything on it yet.

(collect the answer sheets and hand out blank sheets of paper)

You will now complete a free-recall test. In order to complete this task, you will have to write down as many words from the previous task as you remember. If you are unsure, please give your best guess. You will have approximately 2 minutes to complete this task. Please remember to complete this task individually and in silence. When you are done, please put down your pens, flip your papers over, look up, and remain seated and silent. Does anyone have any questions?

(answer questions if any)

Your time starts now.

(start timer for 2 minutes)

(if timer rings and people are still writing, wait for a maximum of 1 minute for them to finish)

We will now come and collect your papers and give you a debriefing form, which gives you more information about this experiment. Please remain silent seated while you fill this out. When you are done filling it out, put your pens down and look up.

(collect the free-recall task papers and hand out debriefing forms)

(wait until everyone is or seems to be done)

We will not collect the debriefing form and give you a slip of paper that contains our emails as well as our supervisor's email should you wish to contact us in the near-future. Thank you all for taking part in our experiment.

(collect the debriefing forms and hand out the contact sheets)

Appendix 3: Informed Consent Sheet

Informed Consent Form

We are performing an experiment as part of our Psychology Internal Assessment. We are investigating memory and would like to ask you if you want to take part in our experiment. The experiment will take about 10 minutes and requires you to answer 10 yes or no questions about different words.

If you agree to take part in this experiment, it is important to know that all data we obtain will be kept confidential and anonymous. You will have the right to withdraw from this experiment at any time (beginning, during, or after) for any reason. If you withdraw, your results will be destroyed. You will receive more information about the nature of the experiment after data collection is completed.

Thank you,
xxxxxx, xxxxxx, xxxxxx, and xxxxxx

I, _____ understand the nature of this experiment and agree to voluntarily participate. I give the researchers permission to use my data as part of their experimental study (unless I withdraw after the experiment).

Signature: _____ Date: _____

Please check the following boxes if they apply to you:

- ☐ I am not colorblind.
- ☐ I am fluent in English.

Appendix 4: Yes/No Answer Sheet

1. YES / NO
2. YES / NO
3. YES / NO
4. YES / NO
5. YES / NO
6. YES / NO
7. YES / NO
8. YES / NO
9. YES / NO
10. YES / NO

Appendix 5: List of Structural Processing Questions

The following questions were displayed on our PowerPoint for the structural processing condition:

1. Is this word written in capital letters?
2. Is this word written in red?
3. Is this word written in lowercase letters?
4. Is this word written in green?
5. Is this word written in capital letters?
6. Is this word written in black?
7. Is this word highlighted in yellow?
8. Is this word written in lowercase letters?
9. Is this word underlined?
10. Is this word written in orange?

Appendix 6: List of Semantic Processing Questions

The following questions were displayed on our PowerPoint for the semantic processing condition:

1. Does this word fit in this sentence: "Are you this dog's _____?"
2. Does this word fit in this sentence: "Can I ride in your _____?"
3. Is this word a type of animal?
4. Is this word a type of furniture?
5. Is this word a body part?
6. Does this word fit in the sentence: "Are you on _____ with your studies?"
7. Can this word be found on someone's face?
8. Does this word fit in this sentence: "I went to the _____."
9. Is this word a color?
10. Is this word something you can do at a spa?

Appendix 7: Structural Processing Condition Debriefing Form

Debriefing Form

Thank you very much for taking part in our study today. We hope you found it to be an interesting experience. This experiment was a partial replication of a study by Craik and Tulving in 1975. We were investigating the Levels of Processing Model of Memory, which says that the deeper you process something, the better you remember it. You were in a condition where the questions you were asked regarding the words led to shallow processing. Participants in the other condition were asked questions that led to deep processing.

If you found any aspect of the experiment upsetting, please let us know so we can provide details of sources of support.

It is your right to withdraw your results from the experiment should you wish so.

☐ Please tick the box if you would like us to destroy your data.

If you would like to receive a copy of the results when they are available, please leave your email address here:

_____

Thank you again,
xxxxxx, xxxxxx, xxxxxx, and xxxxxx

Appendix 8: Semantic Processing Condition Debriefing Form

Debriefing Form

Thank you very much for taking part in our study today. We hope you found it to be an interesting experience. This experiment was a partial replication of a study by Craik and Tulving in 1975. We were investigating the Levels of Processing Model of Memory, which says that the deeper you process something, the better you remember it. You were in a condition where the questions you were asked regarding the words led to deep processing. Participants in the other condition were asked questions that led to shallow processing.

If you found any aspect of the experiment upsetting, please let us know so we can provide details of sources of support.

It is your right to withdraw your results from the experiment should you wish so.

☐ Please tick the box if you would like us to destroy your data.

If you would like to receive a copy of the results when they are available, please leave your email address here:

_____

Thank you again,
xxxxxx, xxxxxx, xxxxxx, and xxxxxx

Appendix 9: Contact Sheet

Contact Sheet

Please feel free to contact any one of us (xxxxxx@xxxxxx, xxxxxx@xxxxxx, xxxxxx@xxxxxx, or xxxxxx@xxxxxx) or our supervisor (xxxxxx@xxxxxx) if you have any questions or concerns regarding the experiment.

Appendix 10: Raw Data Tables

Raw Data Table Showing the Number of Words Correctly Recalled by the 17 Participants in the Structural Processing Condition

Participant Number	Number of Words Correctly Recalled by the Participant
1	7
2	4
3	5
4	5
5	4
6	4
7	6
8	5
9	2
10	6
11	4
12	8
13	0
14	2
15	1
16	3
17	4

Raw Data Table Showing the Number of Words Correctly Recalled by the 17 Participants in the Semantic Processing Condition

Participant Number	Number of Words Correctly Recalled by the Participant
1	4
2	6
3	8
4	8
5	5
6	6
7	6
8	5
9	7
10	3
11	8
12	4
13	5
14	6
15	5
16	7
17	5

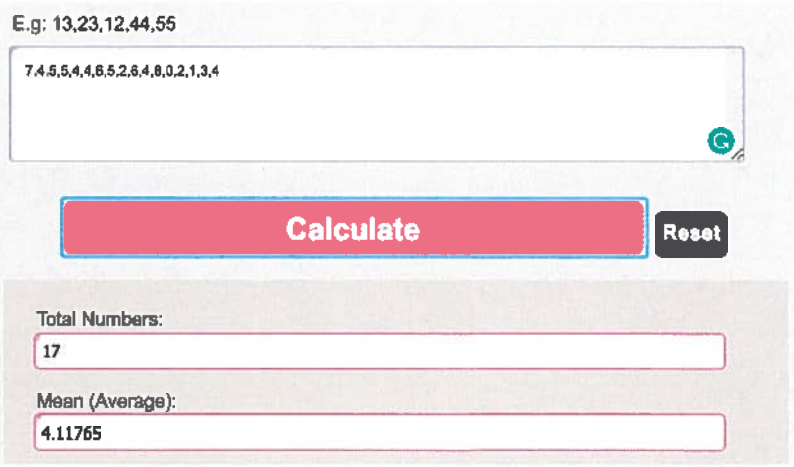
Appendix 11: Descriptive Statistics Calculations

All of these calculations were taken from the following website:

https://www.easycalculation.com/statistics

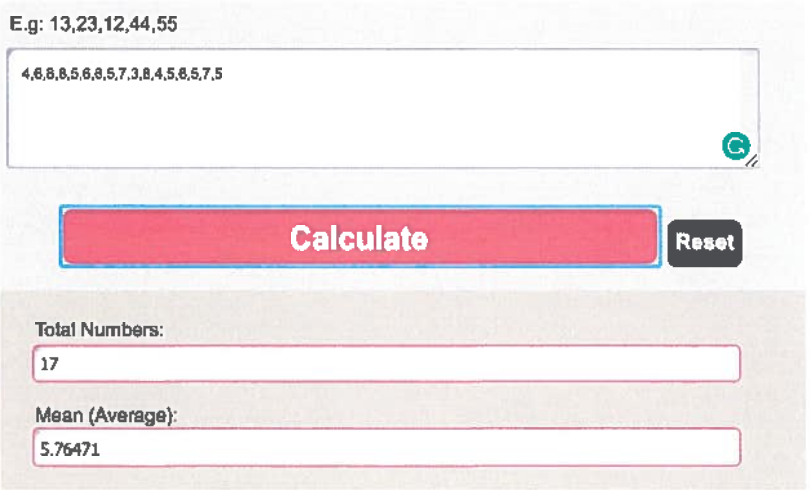
Screenshot Taken from an Online Calculator that Calculated the Mean Number of Words Recalled by Participants in the Structural Processing Condition

(https://www.easycalculation.com/statistics/mean-median-mode.php)

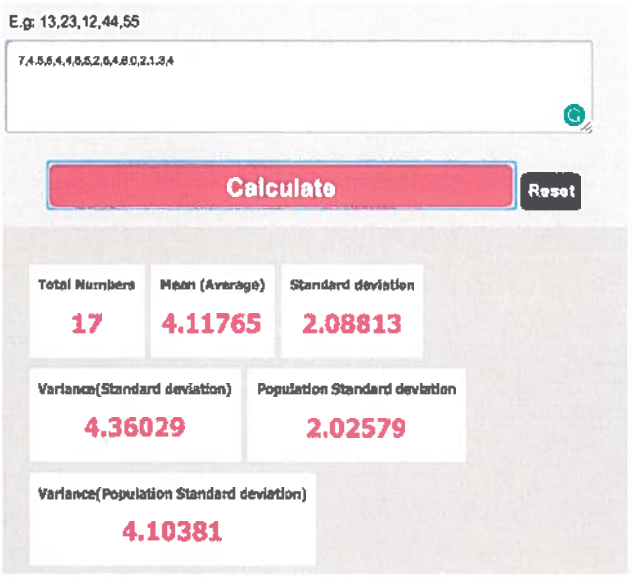


Screenshot Taken from an Online Calculator that Calculated the Mean Number of Words Recalled by Participants in the Semantic Processing Condition

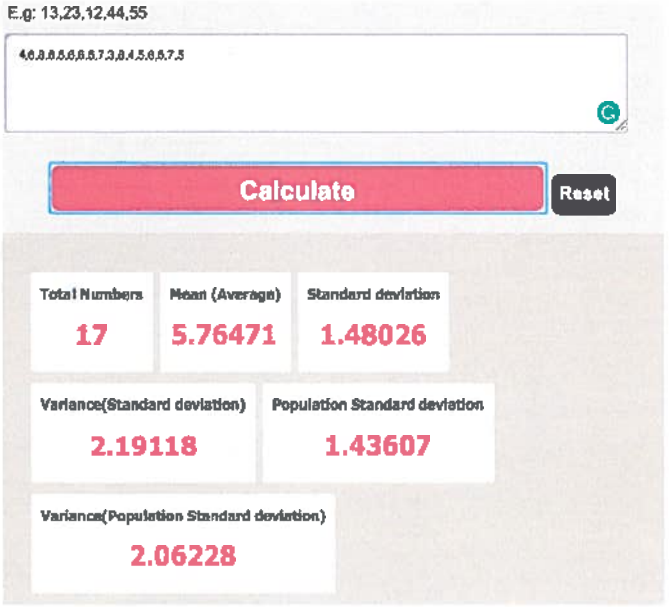
(https://www.easycalculation.com/statistics/mean-median-mode.php)



Screenshot Taken from an Online Calculator that Calculated the Standard Deviation of the Number of Words Recalled by Participants in the Structural Processing Condition (https://www.easycalculation.com/statistics/standard-deviation.php)



Screenshot Taken from an Online Calculator that Calculated the Standard Deviation of the Number of Words Recalled by Participants in the Semantic Processing Condition (https://www.easycalculation.com/statistics/standard-deviation.php)



Appendix 12: Inferential Statistics Calculations

All of these calculations were carried out by undergoing a Mann-Whitney Test on the website

http://vassarstats.net/index.html

Screenshot of the Raw Data that was Inputted into an Online Mann-Whitney Test and the Resulting Ranked Data that was Calculated

Data Entry:

	Ranks for		Raw Data for	
count	Sample A	Sample B	Sample A	Sample B
1	29	10	7	4
2	10	24.5	4	6
3	17.5	32.5	5	8
4	17.5	32.5	5	8
5	10	17.5	4	5
6	10	24.5	4	6
7	24.5	24.5	6	6
8	17.5	17.5	5	5
9	3.5	29	2	7
10	24.5	5.5	6	3
11	10	32.5	4	8
12	32.5	10	8	4
13	1	17.5	0	5
14	3.5	24.5	2	6
15	2	17.5	1	5
16	5.5	29	3	7
17	10	17.5	4	5
Reset	Calculate from Ranks		Calculate from Raw Data	

Screenshot of the Results from the Online Mann-Whitney Test

Mean Ranks for					
Sample A	Sample B	U_A =	213.5	$P_{(1)}$	$P_{(2)}$
13.4	21.6	z =	-2.36	0.0091	0.0183

Note that mean ranks are provided only for descriptive purposes. They are not part of the Mann-Whitney test. ~ Note also that the z-ratio is calculated only if n_a and n_b are both equal to or greater than 5.

Critical Intervals of U_A for na=17; nb=17

	Level of Significance for a Directional Test		
	.05	.025	.01
	Non-Directional Test		
	--	.05	.02
lower limit	96	87	77
upper limit	193	202	212

The observed value of U_A is significant at or beyond the designated level if it is equal to or smaller than the indicated lower limit for that level or equal to or greater than the upper limit. It is non-significant if it is larger than the lower limit and smaller than the upper limit.

The adjacent critical intervals are calculated only if n_a and n_b both fall between 5 and 21, inclusive. For sample sizes smaller than 5, you can refer your results to a standard table of Mann-Whitney critical values, such as the following, provided by the Department of Mathematics & Statistics at the University of Saskatchewan:

http://math.usask.ca/~laverty/S245/Tables/wmw.pdf

with n_a=17, n_b=17, U_A= 213.5 , and U_B= 75.5