

EFFECTS OF CATEGORY SIZE ON CUED RECALL PERFORMANCE
FOLLOWING THE GENERATE-RECOGNIZE MODEL

APPROVED BY SUPERVISING COMMITTEE:

Reed Hunt, Ph.D., Chair

Rebekah Smith, Ph.D.

Paul Romanowich, Ph.D.

Accepted: _____
Dean, Graduate School

DEDICATION

This thesis is dedicated to mom, my best friend. You are everything, and I want nothing more than to show you what an amazing job you did with me.

**EFFECTS OF CATEGORY SIZE ON CUED RECALL PERFORMANCE
FOLLOWING THE GENERATE-RECOGNIZE MODEL**

by

KELLY DEPROSPO, B.A.

THESIS

Presented to the Graduate Faculty of
The University of Texas at San Antonio
in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE IN PSYCHOLOGY

THE UNIVERSITY OF TEXAS AT SAN ANTONIO
College of Liberal and Fine Arts
Department of Psychology
August 2015

ACKNOWLEDGEMENTS

A debt of gratitude is owed to the entire faculty and staff of the UTSA department of psychology. Specifically I would like to thank Dr. Reed Hunt for all of his time and attention in bringing this thesis to life. I would be nowhere had it not been for his guidance.

I would also like to thank each of my professors in this master's program for helping every student meet the demands of the program, for exposing us to so much knowledge in such a short span of time, and for challenging me in ways I never thought possible. I will without a doubt take pieces of each course with me as I continue on.

August 2015

EFFECTS OF CATEGORY SIZE ON CUED RECALL PERFORMANCE
FOLLOWING THE GENERATE-RECOGNIZE MODEL

Kelly DeProspero, M.S.
The University of Texas at San Antonio, 2015

Supervising Professor: Reed Hunt, Ph.D.

Memory retrieval in recall has been persistently found to follow the generate-recognize model of recall, which deems that items to be remembered must first be generated and then will be subjected to a recognition check prior to output. Additionally, the manipulation of category sizes in cued recall studies has revealed a performance advantage for small over large categories in recall tests. The present study has been designed to determine if the effects of category size on recall can be explained by the generate-recognize model, and to do so by determining the effects of category size on cued recall, generate-only, and generate-recognize test conditions. To confirm adherence to the generate-recognize model, small categories should demonstrate the same performance effects in cued recall and generate-recognize conditions. Furthermore, cued recall performance should be less than or equal to that of generation performance alone, and greater than or equal to generate-recognize performance. Because small categories have been argued to be distinctively processed in category cued recall, and because distinctive processing has been shown to have a large effect on recognition, the performance advantage with small categories should not be revealed in generate-only conditions as no recognition check is involved.

TABLE OF CONTENTS

Acknowledgements.....	iii
Abstract.....	iv
List of Figures.....	vi
Chapter One: Background.....	1
Chapter Two: Problems for the Early Model.....	3
Chapter Three: Revised Generate-Recognize Model	4
Chapter Four: Pursuing the Revised Model.....	7
Chapter Five: Method	8
Chapter Six: Results.....	12
Conclusion	16
References.....	19
Vita	

LIST OF FIGURES

Figure 1	Mean proportion of study items produced as a function of category size	12
Figure 2	Mean proportion of study and baseline items produced in the generate condition as a function of category size.....	14
Figure 3	Mean proportion of study items recognized in the generate-recognize condition as a function of category size.....	15

CHAPTER 1: BACKGROUND

In understanding the underlying mechanism of recall in memory, we all assume a psychological output process known as retrieval to be involved. Even in his early evaluation of memory, James (1890) refers to retrieval as the first step in recall, and describes it as “the revival in the mind of an image or copy of the original event” (p. 649). Various theoretical descriptions acknowledging the importance of retrieval in recall have developed over time, the most venerable of which has been the generate-recognize model.

Early Generate-Recognize Model

The generate-recognize model encompasses the basic idea that recall for events includes both the production of a response (generation) and a validation step to ensure the correct item was produced (recognition). According to the theory, during recall an unbiased list of potential responses will be generated, and then the process of recognition is deployed to select the correct response among these generated items. The principal research agenda motivated by this idea has been to establish that generation and recognition are separate processes necessary for output in recall and, by extension, that recognition is in fact necessary for successful recall. Initially research focused on the difference between generation and recognition, which led to experiments designed to show that certain variables have a different effect on recall and recognition. These different effects are understandable due to the lack of generation involved in recognition, therefore variables affecting generation will not affect recognition as they do recall.

Different effects on recall and recognition. Word frequency has been manipulated to evaluate the different effects variables can have on recall and recognition performance. For example, Hall (1954) provided participants with study lists consisting of both high-frequency (i.e. more commonly encountered) and low-frequency (i.e. more rare) words. Examples of high-

frequency words used in the study were “forward” or “because”, and low-frequency examples included words such as “winsome” or “bagpipe”. Upon testing free recall performance, high-frequency words were shown to have better performance outcomes than low-frequency words. For recognition tests, Shepard (1967) used a similar procedure to demonstrate that low-frequency words result in better recognition performance than high-frequency words. To jointly exhibit these effects, Gregg (1976) included both recall and recognition tests in his study using a word frequency manipulation. The previous findings were confirmed: high-frequency words result in better recall, and low-frequency words result in better recognition.

Categorization of study lists has also been shown to differently affect recall and recognition performance. In research by Kintsch (1968), participants were presented with study lists either in a highly organized or an unorganized structure, and performance was assessed using free recall followed by a recognition test. The category norms of Cohen, Bousfield, and Whitmarsh (1957) were used to develop these study lists. For the highly organized lists, the top 10 most common responses from 4 different categories were selected, whereas the unorganized lists consisted of the 10 least common responses from 4 categories. High organization of study lists was found to improve recall but have no effect on recognition performance. This demonstrates how categorization of study material helps to enhance the generation process, with a more organized set of items coming to mind during generation. Since recognition does not involve generation, no effect was observed in recognition testing when the study lists were organized.

CHAPTER 2: PROBLEMS FOR THE EARLY MODEL

Tulving and Thomson (1973) sought to demonstrate the importance of encoding on retrieval performance, and postulated that both recall and recognition have the same retrieval process involved. In three experiments, participants were given study lists consisting of weak cue-target relationships (e.g., spider-bird), then were presented with different, stronger cues (e.g., eagle instead of spider) when testing generate and generate-recognize performance. Cued recall was also tested using the initial study cues, and performance was found to be relatively good. However, comparing across conditions they found that using a different cue in testing than what was previously studied led not only to diminished recall performance, but furthermore to poor recognition performance even when words were accurately recalled.

The effects of retrieval cues on recognition performance seriously undercut the generate-recognize model, which was formulated on the premise that recognition tests do not involve a retrieval process. If recognition tests did not involve retrieval, performance would not have been reduced when the cue was changed for testing from what was studied. Also contrary to the model's assumptions, Tulving and Thomson (1973) showed that generation is heavily influenced by the particular prior encoding experience. According to their encoding specificity principle, the authors argued that it is not a list of items related to the cue brought to mind in retrieval, but rather the cue elicits the specific representation with which it was encoded. They further explained that given the proper cue, effective retrieval in both recall and recognition will be observed. Due to the recognition failure of recallable words observed in their research, the authors rejected the generate-recognize model and concluded the encoding specificity principle to better explain recall.

CHAPTER 3: REVISED GENERATE-RECOGNIZE MODEL

Responding to the criticisms presented by Tulving and Thomson (1973), the goal of generate-recognize research moved away from demonstrations of different effects on recall and recognition and began to focus on demonstrating that recall data could be accounted for through the combined influence of generation and recognition. In an important revision to the generate-recognize model proposed by Jacoby and Hollingshead (1990), the authors accepted that particular prior experience does have an effect on generation. They attributed the ability to recall unrecognizable items to a certain fluency in generation, which can lead to bypassing recognition altogether. Thus, they concluded the importance of recognition in recall would vary in different circumstances.

To test their revised model, Jacoby and Hollingshead (1990) devised a paradigm consisting of study list presentation followed by a cued recall, generate, or generate-recognize test. In cued recall, participants responded to cues presented only with words that had previously been studied. For the generate test, participants would simply respond to the cues presented with the first word that came to mind. The generate-recognize test involved these same generate instructions, but also included a request to recognize any generated items that were on the study list. The authors then laid out several predictions for the revised model: that variables affecting recall performance would have the same effect on generate-recognize performance; that certain variables would have a different effect on generation and recognition; and that recall performance should never exceed generate-only performance. This final prediction stems from the fact that the recognition step in recall can only lower, not improve, overall performance if the correct target has been generated and is then unrecognized. The authors tested their predictions through manipulating several variables and assessing their effects on recall tests compared to

tests explicitly requiring generation and generation-recognition. With each of their predictions ultimately confirmed, this lent strong support for the revised model in arguing that generation and recognition are different processes yet both necessary for recall.

Problems for the Revised Model

Recently, data were reported that appear to contradict the final prediction put forth by Jacoby and Hollingshead (1990). In the course of studying implicit memory, Mulligan (2012) compared cued recall and category generation as a function of category size at study. For purposes of this discussion, category size is intended only to define the specific number of items on a given list, and should not be interpreted as all possible items that fall within a category. Previously, category size has been shown to affect retrieval in that cued recall performance is better for small categories than it is for large categories (Earhard, 1967; Hunt & Seta, 1984). In Mulligan's (2012) research, better performance with small categories was observed for cued recall over generate conditions. While this aligns with prior findings of the performance advantage for small categories in cued recall, it directly conflicts with the idea that recall performance should never exceed that of generate-only.

Although it was not Mulligan's (2012) goal to test the generate-recognize model, it is important to look at how performance was measured to help understand these contradictory findings. In his research, generate performance was determined by subtracting the proportion of non-target items produced from the proportion of target items produced. This method may account for the performance advantage he noted in cued recall over generate tests as more study items are available in large categories, therefore more targets would understandably be produced. Perhaps a more appropriate method for measuring generate performance would be to calculate the proportion of target items produced, without subtracting new items. Bearing these possible

methodological drawbacks in mind, we cannot dismiss the assumptions of the revised generate-recognize model.

CHAPTER 4: PURSUING THE REVISED MODEL

The purpose of the present research was to examine the effect of category size following the Jacoby and Hollingshead (1990) paradigm. Category size was manipulated within subjects using 2, 8, and 12 item categories, and performance between subjects was measured using cued recall, generate, and generate-recognize tests. Although it was not fully known what would occur within the 8 item categories, the main focus of this research was placed on the effects of the small (2 item) and large (12 item) categories. We anticipated small categories to result in better performance on cued recall tests, and the other test conditions were then used to analyze that effect in terms of the generate-recognize model. Contrary to Mulligan's (2012) findings, we predicted no difference in the proportion of items generated across various category sizes as generation is an implicit test of memory. If, however, generation is subject to strategic control, a possible category size effect was expected to mirror that of cued recall. A category size effect was also expected in recognition tests, a prediction derived from the distinctive processing research of Hunt (2013). Distinctive processing occurs when individual item processing, engaged upon studying small categories, is coupled with providing relational information (i.e., category names) to participants. Because distinctive processing has been shown to facilitate memory for recognition tests, better performance with small categories was anticipated.

CHAPTER 5: METHOD

Participants

The present study utilized 247 student volunteers participating for partial course credit in Introduction to Psychology at the University of Texas at San Antonio. As previous research has demonstrated that 40 participants per test condition is more than adequate for detecting differences among these types of conditions (Hunt, in press; Jacoby & Hollingshead, 1990), each condition contained a minimum of 42 participants in order to accommodate counterbalancing of the study lists. Participants were excluded upon failing to produce a minimum of 80% of the total requested words, or when more than four adjectives (rather than category examples) were produced for a given category. A total of 69 participants were excluded and replaced (5 from recall, 17 from recall+baseline, 21 from generate, and 26 from generate-recognize).

Materials

A total of 144 words were selected from the category norms of Battig and Montague (1969). The norms presented in their research were ordered based on the overall frequency of respondents naming an item for a given a category. From each of twelve categories, twelve instances were selected from the 5th, 8th, and 10th through 19th frequency positions, and the average frequency of the selected instances across all categories was 0.21. Two separate study lists were created by arbitrarily assigning 6 of the categories to each study list. Each of the study lists contained 44 words: two categories contributed 2 instances; two categories contributed 8 instances; and two categories contributed 12 instances. Three versions of each master list were created by counterbalancing which categories supplied instances to specific values of category size. This counterbalancing ensured that each category was represented equally often at each

category size across participants. A minimum of 7 participants saw each of the 6 versions of the study lists.

At test, all participants were provided with a booklet with one category label on each page of the booklet. For the recall condition, the booklet had 6 pages containing only the labels for the study list categories the participant had seen. In the generate and generate-recognize conditions, the booklets contained 12 pages. To ensure no demand effects would occur for participants receiving the 12 versus 6 page booklets, an additional recall+baseline condition was incorporated. For these latter three conditions, the 6 study list category labels were given along with the 6 labels corresponding to the categories from the other study list, which was not seen by the participant. The non-studied categories provided base-rates for generation of category instances.

Procedure

Participants were run in groups no larger than 5 at a given time. Upon entering the lab, each was given information about the nature of the study and was asked to give their consent prior to participation. Instructions for the study lists were then presented on a computer. First, participants were informed that they would see words presented one at a time on the monitor, and that they should rate how pleasant each word is. The pleasantness rating task was added to ensure attention to each item. Pleasantness was rated on a scale of 1 (very unpleasant) to 4 (very pleasant). The scale was presented along with each study word, and participants recorded their ratings using the number keys on the computer. The list words were randomly presented with respect to category. Following completion of the study list presentation, a 3 minute distractor activity was presented. The distractor activity displayed a letter of the alphabet on the monitor, and participants were instructed to press the key for the following letter of the alphabet.

Following the distractor activity, instructions were presented on the computer for one of the four test conditions to which the participant was assigned.

Cued Recall and Cued Recall+Baseline Instructions

In a moment the experimenter will bring you a test booklet. On each sheet of your test booklet, you will see a series of blank spaces and the name of a category at the top of the sheet. On each sheet, please write all of the words you saw from that category in the list we just showed you. Be careful to only write words that you remember being in the list. When you think you have written all of the words you remember from the category that were in the list, continue to complete the blank spaces by writing words from the category that come to mind. Fill out all of the blank spaces before going on to the next page of the booklet. If you do not remember seeing the category on a test sheet, please list the first 12 instances of that category that come to mind.

Generate Instructions

In a moment the experimenter will bring you a test booklet. On each sheet of your test booklet, you will see the name of a category at the top of the sheet and a series of blank spaces beneath. Please list the first 12 instances of that category that come to mind. There are no correct or incorrect answers. Just list the instances that come to mind when you think of the category.

Generate-Recognize Instructions

In a moment the experimenter will bring you a test booklet. On each sheet of your test booklet, you will see the name of a category at the top of the sheet and a series of blank spaces beneath. Please list the first 12 instances of that category

that come to mind. If the instance you write is a word that you recognize as having been in the list you just studied, circle it. The instances that you write should be the ones that just come to mind when you think of the category, but if the instance happens to be one that you encountered in the list you just saw, draw a circle around it.

Upon completion of the test phase, participants were then debriefed and thanked for their time participating in the study.

The booklets were then scored according to their test condition. For cued recall, the number of study list items correctly recalled from each category size were counted. In the generate condition, study list items generated from each category size were counted as correct. For the generate-recognize condition, study list items which were both correctly generated and recognized (i.e., circled) from each category size were counted. Baseline categories were also scored for the generate, generate-recognize, and recall+baseline conditions, and any items listed from the non-studied baseline lists were counted.

CHAPTER 6: RESULTS

The mean proportion of studied items produced in each test condition as a function of category size are shown in Figure 1. The proportions shown for the generate-recognize condition are the mean proportion of studied items that were both generated and recognized. These data were subjected to an overall 3 (test condition) x 3 (category size) analysis.

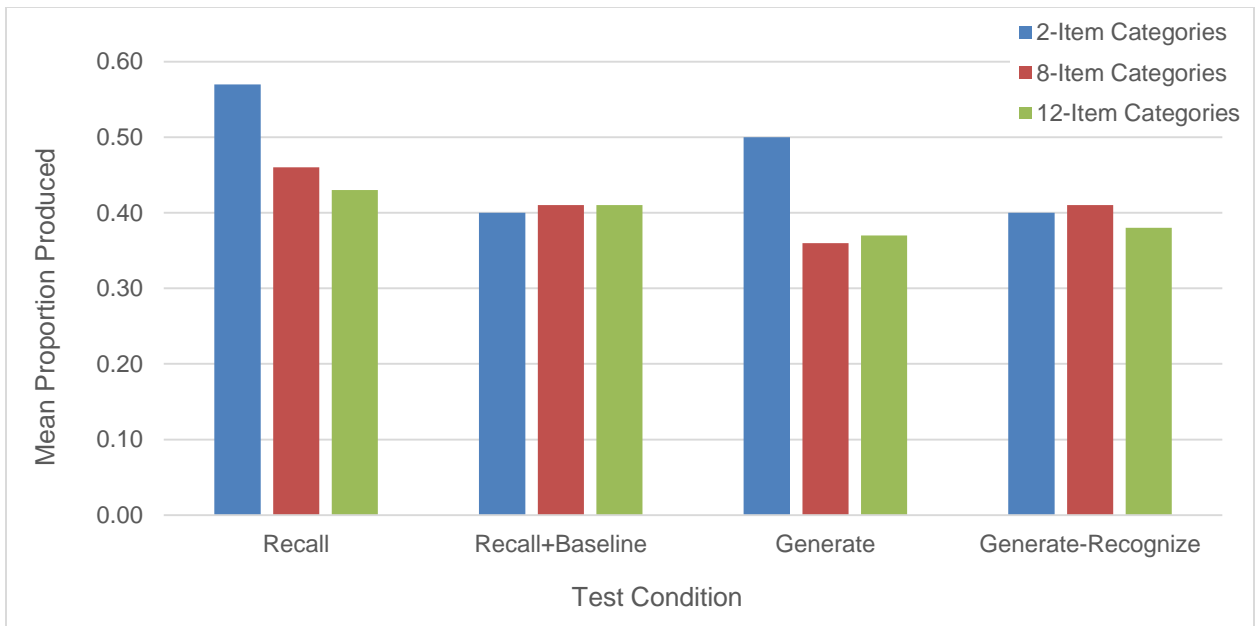


Figure 1. Mean proportion of study items produced as a function of category size.

The overall analysis revealed a significant difference between test conditions, $F(3, 174) = 2.98$, $MSE = .07$, $p = .03$, $\eta_p^2 = .05$. Tukey's HSD comparisons demonstrated that this difference was significant between the recall and generate-recognize conditions ($p = .04$). No significant differences were seen between the recall and generate ($p = .13$), recall and recall+baseline ($p = .09$), generate and generate-recognize ($p = .97$), generate and recall+baseline ($p = .999$), or generate-recognize and recall+baseline ($p = .99$) conditions. These results are generally consistent with the generate-recognize model in that there was no reliable difference between the

recall and generate conditions, and recall performance was higher than generate-recognize performance.

As is obvious from Figure 1, there was a reliable effect of category size, $F(1.51, 263.43) = 9.14$, $MSE = .04$, $p = .001$, $\eta_p^2 = .05$, with small categories apparently yielding higher production than large categories. However, this effect of category size varied as a function of test condition, as was indicated by the significant interaction, $F(4.54, 263.43) = 3.19$, $p = .01$, $\eta_p^2 = .05$.

In order to diagnose the source of this interaction, each test condition was separately analyzed as a function of category size. Significant effects of category size were observed in both the recall condition, $F(1.50, 64.61) = 8.41$, $MSE = .04$, $p = .002$, $\eta_p^2 = .16$, and the generate condition, $F(1.52, 63.69) = 8.95$, $MSE = .04$, $p = .001$, $\eta_p^2 = .18$. In recall, this difference was significant between the 2- and 8-item categories, $t(43) = 2.56$, $p = .01$, and the 2- and 12-item categories, $t(43) = 3.74$, $p = .001$. In the generate condition, a significant difference was also seen between the 2- and 8-item categories, $t(42) = 3.14$, $p = .003$, and the 2- and 12-item categories, $t(42) = 3.40$, $p = .002$. No effect of category size was observed in the recall+baseline or generate-recognize conditions. Analysis of the interaction reveals a serious discrepancy from the generate-recognize model in that the generate-recognize condition was not affected by category size in the same way as the recall condition.

In addition to these overall analyses comparing the 3 test conditions, specific effects within test conditions were also analyzed to examine some of the predictions made by the revised generate-recognize model. First, the revised model posits that generation is influenced not only by general prior knowledge, but also by immediate prior experience. To evaluate this prediction, the number of study items produced was compared with the number of baseline items

produced in the generate test condition. The mean proportion of study and baseline items produced at each category size is presented in Figure 2. A significant difference was observed between the numbers of study versus baseline items produced, $F(1, 42) = 28.35$, $MSE = .04$, $p < .001$, $\eta_p^2 = .40$. No significant interaction was observed between category size and the type of item produced, $F(1.27, 53.49) = 1.05$, $MSE = .05$, $p = .33$, $\eta_p^2 = .02$, but the effect of category size was reliable, $F(1.69, 70.95) = 13.54$, $MSE = .02$, $p < .001$, $\eta_p^2 = .24$. The total proportion of study and baseline items produced were then collapsed over item type for each category size, and analyses again revealed significant differences between the 2- and 8-item $t(42) = 4.16$, $p < .001$, and the 2- and 12-item categories, $t(42) = 4.04$, $p < .001$.

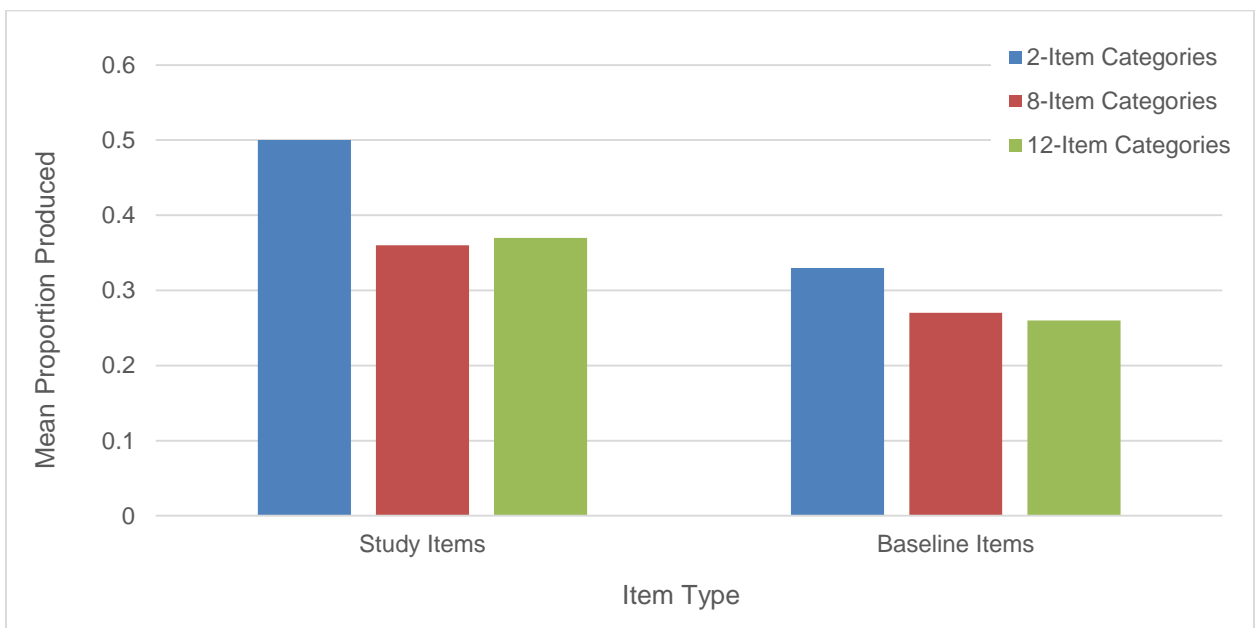


Figure 2. Mean proportion of study and baseline items produced in the generate condition as a function of category size.

Another prediction concerned the effect of category size on recognition memory, therefore the recognition component of the generate-recognize condition was also separately analyzed. The mean proportion of study items recognized at each category size is presented in

Figure 3. No significant effect of category size was supported for the recognition component of the generate-recognize condition, $F(1.63, 73.26) = .398$, $MSE = .05$, $p = .63$, $\eta_p^2 = .01$.

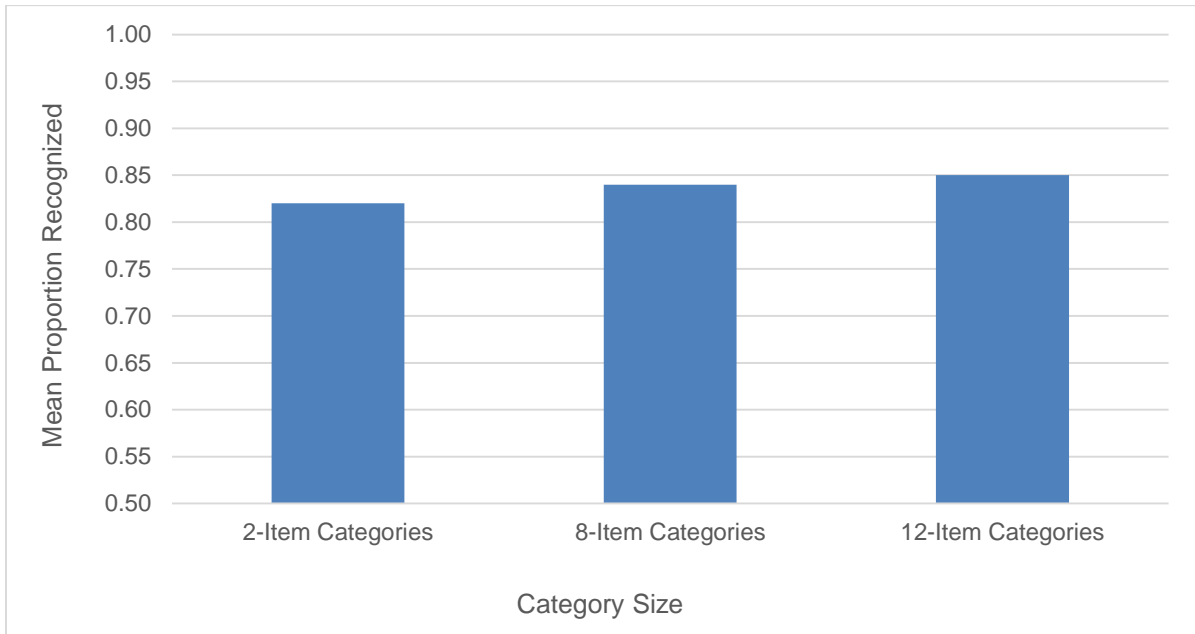


Figure 3. Mean proportion of study items recognized in the generate-recognize condition as a function of category size.

CONCLUSION

The primary purpose of this research was to test the revised generate-recognize model by manipulating category size. To accomplish this, 4 test conditions (cued recall, cued recall+baseline, generate, and generate-recognize) were established to align with the Jacoby and Hollingshead (1990) paradigm. The revised model assumes that generation performance can be affected by recent experiences, such as studying a word list. Accordingly, the revised model predicted that there would be an effect of study list on generate performance. In the present research, the difference between the number of study list items generated compared to the number of baseline items generated across conditions was reliable, indicating that prior study did in fact affect the production of those items.

More centrally, the model predicted that recall performance should be less than or equal to generation performance as the recognition check involved in recall can only lower performance if a correctly generated item is not recognized. In fact, the present data are consistent with this prediction as no difference was seen between recall and generation performance. In light of Mulligan's (2012) findings, it was expected that any difference in the proportion of items produced across category sizes in the generate condition would mirror that of the recall condition. This was also confirmed with small categories demonstrating the best performance in both the generate and recall conditions, indicating that generation is subject to strategic control as postulated.

The model also predicted that recall performance should be greater than or equal to generate-recognize performance. This prediction is due to the fact that fluently generated items in recall may bypass the recognition phase, whereas the recognition check will be applied to every item in the generate-recognize condition. Because the recognition stage can only lower performance on correctly generated study list items, performance in the generate-recognize

condition can only be equal to or lower than performance in the recall condition. Again the data were consistent with this prediction in that recall performance was higher than generate-recognize performance.

In spite of these positive outcomes, certain important comparisons were not consistent with the generate-recognize model. As outlined by Jacoby and Hollingshead (1990), one of the main predictions of the revised model was that variables affecting recall would have the same effect on generate-recognize performance; however, the present data did not confirm this prediction. In recall, a significant effect of category size was revealed, yet this effect was not observed in the generate-recognize condition. Because category size did not affect recall and generate-recognize performance in the same way, the generate-recognize model does not appear to be the retrieval strategy used.

While the effects of category size were different for generate-recognize versus recall performance, the pattern of results for the generate-recognize condition does parallel the pattern for the recall+baseline condition. It is possible that the recall+baseline group did use a generate-recognize strategy that for some reason was not used by the recall group. However, this suggestion is not viable as performance in the generate condition was affected differently by category size than performance in both the generate-recognize and recall+baseline conditions. If generate-recognize is the retrieval strategy, then it is expected that recall and generate-recognize performance will show the same effects of a variable as is seen in generation performance. The only way a generate-recognize model can accommodate a violation of this pattern is if the effect of the variable on recognition is complementary to its effect on generation. In this case, for example, recognition should have been poorer for the 8 and 12 item categories, yet the data showed that category size had no effect on recognition. Therefore, the similarity between

recall+baseline and generate-recognize performance is not a convincing argument for a generate-recognize strategy.

Equally perplexing is the difference in the effects of category size on recall and recall+baseline performance. At first glance, the outcome suggests that the effect of category size on the production of study items is eliminated by the requirement to also produce baseline items, a speculation that is consistent with the lack of category size effect on generate-recognize performance. However, category size did affect generate performance, a condition which also required the production of baseline items. The disparity between recall and recall+baseline performance is puzzling.

Although this research did confirm several important aspects of the revised model, a major inconsistency was revealed in that the recall condition was affected by category size, yet the generate-recognize condition was not. While a comprehensive explanation of these results remains unclear, these findings effectively eliminate the generate-recognize model as the retrieval strategy involved in recall.

REFERENCES

- Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, 80(3), 1-46.
- Cohen, B. H., Bousfield, W. A., & Whitmarsh, G. A. (1957). *Cultural norms for verbal items in 43 categories* (Tech. Rep. No. 22). Storrs, CT: University of Connecticut.
- Earhard, M. (1967). Cued recall and free recall as a function of the number of items per cue. *Journal of Verbal Learning and Verbal Behavior*, 6, 257-263.
- Gregg, V. (1976). Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183-216). London, England: Wiley.
- Hall, J. F. (1954). Learning as a function of word-frequency. *The American Journal of Psychology*, 67, 138-140.
- Hunt, R. R. (2013). Precision in memory through distinctive processing. *Current Directions of Psychological Science*, 22, 10-15.
- Hunt, R. R. (In press). Using process dissociation procedure to establish boundaries of process dissociation theory: The case of category cued recall. D. S. Lindsay, C. M. Kelley, A. P. Yonelinas, & H. L. Roediger III (Eds.). In *Remembering: Attributions, processes, and control in human memory*. New York, NY: Psychology Press.
- Hunt, R. R., & Seta, C. E. (1984). Category size effects in recall: The roles of relational and individual item information. *Journal of Experimental Psychology*, 10, 454-464.
- Jacoby, L. L., & Hollingshead, A. (1990). Toward a generate/recognize model of performance on direct and indirect tests of memory. *Journal of Memory and Language*, 29, 433-454.
- James, W. (1890). *The principles of psychology*. New York, NY: H. Holt and Company.

Kintsch, W. (1968). Recognition and free recall of organized lists. *Journal of Experimental Psychology*, 78, 481-487.

Mulligan, N. W. (2012). Differentiating between conceptual implicit and explicit memory: A crossed double dissociation between category-exemplar production and category-cued recall. *Psychological Science*, 23, 404-406.

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, 6, 156-163.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352-373.

VITA

Kelly DeProspero is originally from Belleville, Illinois. She received her associate's degree from Northwest Vista College prior to obtaining a Bachelor of Arts degree in psychology from Texas A&M University in San Antonio. In the last two years she has focused on obtaining a Master of Science degree in psychology from The University of Texas at San Antonio. Her future plans are to primarily explore job opportunities outside of Texas.