Journal of Psychological Inquiry
2012, Vol.17, No. 2, pp.# 17-25

Elizabeth A. Dahl, Ph.D., Award for
Excellence in Undergraduate Research

Judgments of Memory: Do the Number and Presentation of Cues Available Help?

Lauri E. Foster *
Missouri State University

ABSTRACT—In this experiment, participants were given word-pairs, such as LOST-FOUND, and asked to judge those word-pairs for their semantic or associative strength. Semantic strength is how strongly words have the same meaning; whereas associative strength denotes how often words are used in the same context. Although participants report these tasks to be easy, previous research shows their judgments of strength are fairly poor when compared to database answers (Maki, 2007a). To aide participants, repeated cues were given the first word in a word-pair), similar to a previous study (Maki, 2007b). To examine the effect of these repeated words, some participants saw all cue words grouped together, whereas other participants saw cue words mixed. Participant judgments were found to be better than random guessing. Both semantic (Maki, McKinley, & Thompson, 2004) and associative (Nelson, McEvoy, & Schreiber, 2004) database answers were used to examine which memory source, semantic or associative, helped participants with the task and the interaction with mixed or blocked repetitions of cues.

Keywords: Judgments of Memory, Associations, Semantics, Memory

INTRODUCTION

Memory was once thought of as one solid cohesive entity, but that thinking is no longer supported. Memory is now divided into multiple systems (Schacter & Tulving, 1994). The focus of this paper is on the differences between two types; associative and semantic memory. Associative memory is based on the relationships of words that occur together frequently in text and speech, such as OLD-NEW, and is thought to be episodically based (Tulving, 1993). Associative memory is also dependent upon the culture in which individuals are immersed. An example of cultural influences on the associations made between words is ROCK-ROLL, an association that would not exist if not given its notoriety in American popular culture. Associative memory is normally studied through tasks such as word norming. For example, Nelson et al. (2004) compiled a database of associative word norms, all of which are scaled in percentages to determine how commonly words are paired together in speech and text.

The Nelson et al. (2004) database of free association norms is a set of cues, their targets, and the probability of those cue-target pairings. Associative links between cues and targets were determined by instructing participants to respond to a cue with the first word (target) that came to mind. The frequency of that cue-target response was then recorded to determine the forward strength (abbreviated: FSG). Therefore, forward strength is the probability of a cue eliciting that target as the response. An example of a possible cue-target associative pair is LOST-FOUND (FSG = .75) or CAT-MOUSE (FSG = .25). The backward strength for each associative cue was also recorded, which is the probability of the target for a specific cue eliciting that cue. For example, the word-pair CAT-MOUSE have a backward relationship of MOUSE-CAT (BSG = .54). The scale for forward and backward strength ranges from 0-1, with 0 being completely unrelated and 1 indicates the high probability of a target when shown the cue word.

Semantic memory, on the other hand, is generally described as a mental dictionary or the set of facts and world knowledge we have obtained

*Erin M. Buchanan served as Faculty Sponsor.

through life experiences (Tulving, 1993). For instance, we know that CATS and DOGS both have fur, tails, and four-legs, but they make very different noises. Based on this information, concepts can be measured on how related they are, or the semantic distance between pairs. Semantic distance is conceptualized as the overlap in meaning between words. Therefore, CAT and DOG would have a high overlap because they have many of the same basic features. Database information on word relationships can be collected in the same way as with associative norms, namely with participants listing word meanings and features (McRae, Cree, Seidenberg, & McNorgan, 2005). Alternatively, the online dictionary WordNet's hierarchical structure was examined for the distance between words (Maki et al., 2004). For example, CAT would be connected to the concept of ANIMAL, which would in turn be connected hierarchically to the concept of LIVING THINGS. Jiang and Conrath (1997) created a distance scale for this hierarchy ranging from 0-32. A score of zero means that the words can be found on the same tier and are very related (RANK-RATE or POLICE-LAW), whereas a score of 32 means the words are extremely unrelated (CELLO-JELLO).

One use of these databases is to examine participants' judgments of association between word pairs. In several judgments of associative memory (JAM) experiments, Maki (2007a) gave participants a cue-target pair with instructions on how to rate them. Participants were instructed to rate how often college students would give the second word (target) if given the first word (cue). For example, if participants were given OLD-NEW, they should rate the pair around "50," because 50 people out of a 100 would list NEW when shown OLD. By using this technique, Maki was able to view some significant trends with regard to participants' abilities to make associative judgments. When compared to database information (forward strength) from the Nelson et al. database, JAM ratings were both elevated and insensitive to the differences between low and high frequency pairs. Sensitivity was shown in the slope of the function relating forward strength to JAM. The slopes of the JAM function were different from zero, but significantly lower than perfect (slope = 1.0) at slopes of 0.2 - 0.3. Participant's ratings were then compared with semantic relatedness database scores (WordNET), and the JAM function was not affected by semantic relatedness. Thus, participants are able to use some information about word frequency from memory (otherwise slopes would not significantly different from zero) and are able to ignore competing semantic memory information when asked to make an associative judgment in the JAM task.

Assessments were also done to determine if training could improve JAM (Maki, 2007b), specifically, if training would influence bias (the intercept of the JAM function) or sensitivity (the slope of the JAM function). Several participant groups were given feedback training on associative judgments in which they were shown JAM performance. The trained groups were less biased in their judgments than were the untrained groups, but training did not affect the slope of the JAM function when all participants were tested on new word-pairs. In short, participants were able to adjust their judgments when given feedback, but they showed a clear inability to transfer their training to new word pairs. As with Maki's (2007b) research, Koriat and Bjork (2006) have also found that the JAM function was resistant to several different manipulations, such as varying study-test conditions.

Maki (2007b) also tested the effect of multiple targets for the same cue on judgments, which is the main concern of this research study. In his experiment, two groups of participants were shown a cue and four targets associated with that cue. Participants were asked to rate the cue-target pairs through self or other reference. However, this distinction did not change the JAM function. For the cue four-target pairings, judgments were restricted so that the total of the four pairs' ratings must equal 100. Maki found that by limiting the numerical ratings for the cue-target pairs lowered the numerical ratings given (hence, bias), but the slope of the JAM function remained shallow. Even with this rating constraint, participants still showed bias towards overestimation in their judgments with regard to the cue-target pairs' forward strength. Therefore, it appears that people have difficulty judging their context-based memory, even when experiments are designed to improve judgments. While biased, it is important to note that these judgments are still better than random guesses as slopes would not differ significantly from zero.

In comparison to judgments of associative memory, the overlap in semantic features can correspondingly be judged by participants to see if they show the same insensitivity and bias. For example, Maki, Krimsky, and Muñoz (2006) asked participants to provide estimates of the extent that they felt the two concepts shared features. Results showed participants were able to judge semantic relatedness with a high inter-rater reliability, as raters showed remarkable agreement as to which pairs shared many, few, or all features. Their mean ratings were shown to accurately predict semantic dictionary relationship and did so significantly better than other database norms.

The present study combined the experimental paradigms mentioned above to examine differences in memory judgments of strength. First, participants completed both an associative judgment task (JAM) and a semantic judgment task on different word-pairs in one experiment, so comparisons between judgment types could be made. Next, we examined the effect of multiple targets with repeated cue words. Because no research to date has shown the effect of repeated cues on semantic judgments, participants were randomly assigned to trial conditions where repeated cues were shown together (blocked) or mixed throughout the experiment (mixed). The blocked condition mimics Maki's (2007b) experiment by grouping repeated cues with their targets, whereas the mixed condition will examine if multiple targets throughout the experiment will impact judgments. Their judgments were then compared to the database scores for association (Nelson et al., 2004) and semantics (Maki et al., 2004) to examine judgment performance. Matching scores (i.e., associative judgments to associative database) were used to examine how well participants are able to judge the right memory connections. Conversely, non-match scores (i.e., associative judgments to semantic database) were used to examine interference from the opposing memory connection. Participant sensitivity was examined by testing if slopes were greater than zero, indicating that they could judge memory connections better than guessing. Experiment hypothesis are listed below.

Hypotheses

- Hypothesis 1: Block and mixed trial conditions will show different patterns of judgment ability across semantic and associative judgments.

- Hypothesis 2: Participant's judgment scores will be significantly related to the database scores in match conditions (i.e., semantic judgments-semantic databases, associative judgments-associative databases). This hypothesis examines if participants are sensitive to the differences in associative or semantic relatedness when judging those relationships.

- Hypothesis 3: If our study replicates previous research (Buchanan, 2009), associative database scores will be related to semantic judgments, but semantic database scores will not be related to associative judgments (i.e., the non-match condition). This hypothesis examines the extent to which participants rely on the opposite memory information they are not being asked to rate.

- Hypothesis 4: If Hypothesis 1-3 are supported, we will examine the strength of judgment beta weights across mixed and blocked trial conditions as a post hoc test. First, Hypothesis 1 will examine if differences across experimental conditions occurred. If supported, Hypothesis 2 and 3 examine if judgments are better than participant chance guesses (i.e., zero). This hypothesis will examine the non-zero judgment combinations to portray which condition (blocked or mixed) participants were able to perform more accurately.

METHOD

Participants

Participants (N = 102) were recruited from the undergraduate participant pool at a large southern University for course credit. Their age range was approximately 18-24 years old. Power analyses indicated approximately 82 subjects were needed for an $\alpha = .05$ and $\beta = .80$.

Apparatus

The computers used for the experiment included IBM clones, Dells, and HP personal computers. The computers all had 15-inch monitors and were set to a display rate of 60 Hz.

Materials

Associative Pairs. Associative pairs were selected using the Nelson et al. (2004) free association norms. Associative word pairs were assembled based on the cue forward strength, which is the probability that the first word shown (cue) will cause people to think of the second word shown (target). Word-pairings were selected based on the following procedure:

a. All cues with more than four cue-target pairings were selected from the database. For example, COMPUTER has 30 associated targets, such as KEYBOARD, GAME, and PROGRAM.
b. A random set of 25 cue words were chosen from the list created above. These cues were sorted by their target forward strength. Therefore, COMPUTER targets would be ordered PROGRAM, KEYBOARD, and GAME due to their forward strength probabilities.
c. The top four strongest forward strength cue-target pairs from these sets were used.

By selecting 25 cue words with four targets each, we created 100 associative judgment pairs.

Semantic Pairs. We used the same procedure described above to create the semantic judgment pairs with one exception:

a. Cues with at least four related targets were selected from the Maki et al. (2004) database of semantic dictionary relationship. For example, ACHE would be related to HURT, PAIN, HEAD, and BACK.
b. A second random set of 25 of cue words were chosen from the list created in step 1, so that none of the cues overlapped the associative cues. These cue-target pairs were sorted by their semantic dictionary strength. ACHE is most related to HURT, followed by PAIN, then HEAD, and BACK.
c. Finally, the top four semantic cue-target relatedness values were used.

This selection procedure created 25 cue words with four targets each, thus generating 100 semantic judgment pairs.

Stimuli. All cue-target pairings were unique for both the 100 associative judgment pairs and 100 semantic judgment pairs. However, 25 of the target words repeated (e.g., SOUR-GRAPE and VINE-GRAPE). Target words were allowed to repeat because we decided to use the strongest associative and semantic database values. Each cue-target pair has a different relatedness value (SOUR-GRAPE should be rated low for association and semantics, while VINE-GRAPE should be rated high for association and semantics), which should minimize the effect of repeated target words.

Procedure

Participants were brought into the lab and asked to sign a consent form. They were given a participant number and placed at a computer. Each participant was randomly assigned to one of two trial conditions: mixed or blocked. Participants who were assigned to the mixed trial condition were shown the set of cue-target pairs in a randomized order that varied from subject to subject. Participants who were assigned to the blocked trial condition were given the set of cue-target pairs in a specified order. For the blocked trial condition, cue words were first randomized using Microsoft Office Excel's random number generator. Then the matching target words were randomized within those cues. Whereas the cues and targets were randomized into blocks of cues, the blocked order did not change across participants. Participants were then randomly assigned to rate associative or semantic judgments first, and this order was counterbalanced across subjects to control for judgment order effects.

Once they were seated and the program was started, they were given instructions on how to judge the associations or semantics of the word pairs. They were given 15 word pairs as practice for rating the cue-target pairs. In the associative judgment condition, participants were asked to rate how many people out of 100 would say the second word (target) if shown the first word (cue). For example, if shown LOST-FOUND, participants were to rate the number of times someone would say FOUND, if given the word LOST. The judgment scale was based on a scale from 0-9, where 0 indicated 0 to 9 people, 1 indicated 10 to 19 people, etc. For the semantic judgments, participants were asked to rate the cue-target pairs on how much their dictionary definitions overlapped. The judgment scale was a similar 0-9 scale; however, with this scale a rating of 0 meant that the word pairs had no semantic overlap, a rating of 3 meant that the word pair had some semantic overlap, and a rating of 9 meant that there was complete semantic overlap. Participants rated the 100 judgment pairs for each associative and semantic block.

Word-pairs were manipulated so that each cue-target pairing was rated in either the semantic or associative judgment condition, but not both for one participant. Across participants, cue-target pairs were rated in both conditions. Judging the same cue in both associative and semantic judgment pairs allowed the associative and semantic context of that cue to be examined without having a single individual judge the same cue in both contexts, thus counterbalancing each judgment made on an individual cue. Once a participant had completed the experiment, they were debriefed and instructed that they would receive course credit in approximately one week for their participation in the experiment.

Design

The procedure for this experiment results in a 2 (group: mixed, blocked) X 2 (judgment type: associative, semantic) X 2 (database match: match, non-match) design. The group independent variable was between subjects, as participants only rated word pairs in the blocked or mixed trial condition. The judgment type independent variable was within subjects because participants performed both the semantic and associative judgment rating tasks. Finally, the database match independent variable was within subjects. Participant ratings for associative judgments were compared to both associative and semantic database scores and vice versa for semantic judgments. The dependent variable calculated was the standardized regression coefficient (beta) for each judgment and match combination (matches: associative judgment – associative database, semantic judgment – semantic database; non-matches: associative judgment – semantic database, semantic judgment – associative database). The standardized regression coefficient indicated how well a participant was able to judge word pairs for their associative or semantic relationship. Scores close to zero indicate participants are not able to discern the difference between no to little relationship and high relationships (little sensitivity), whereas scores close to one indicate participants are able to judge the different relationships in word pairs (high sensitivity).

RESULTS

The overall subject pool contained 102 participants. Forty-eight individuals were assigned to the blocked-cue group, however data from two were eliminated for failing to follow instructions. Fifty-four participants were assigned to the mixed-cue condition, which made the total number of participants in this experiment N = 100. In all analyses, data were screened for statistical assumptions and multivariate outliers.

Hypothesis 1 – Group Differences

Participant judgments were compared to the associative and semantic database scores for each judgment type. Using the database norms, we calculated the standardized beta weights for judgments matching conditions (i.e., associative database-judgment, semantic database-judgment) and for non-matching conditions (i.e., semantic database-associative judgment, associative database-semantic judgment). These standardized weights give an indication of how well participants were able to discriminate between word-pairs with different strength relationships (low versus high), where larger beta weights portray better judgment performance. We expected to find blocked and mixed conditions would show different patterns of judgment weights.

A 2 (associative versus semantic judgments) by 2 (database match versus non-match) by 2 (blocked versus mixed condition) mixed factorial ANOVA was used to analyze the data. First, the main effect of judgment type was significant, $F(1,98) = 67.753$, $p < .001$, $\eta^2 = .409$. Overall, participants were better at semantic judgments ($M_{beta} = 0.107$, SE = 0.015) than associative judgments ($M_{beta} = 0.065$, SE = 0.014). The main effect of the matching database to judgment was significant, F

(1,98) = 109.051, p <.001, η^2= .527. Participants performed better in the match conditions (M_{beta} = 0.049, SE = 0.017) than the non-match conditions (M_{beta} = 0.007, SE = 0.012). Finally, the blocked versus mixed between groups main effect was not significant, F (1,98) = 1.435, p = .234, η^2= .014.

All three of the two-way interactions were significant: judgment type and blocked or mixed condition, F (1,98) = 11.275, p = .001, η^2= .103; database match and blocked or mixed condition, F (1,98) = 6.750, p = .011, η^2 = .064; judgment type and database match, F(1,98)=6.358, p=.013, η^2=.061. However, the three way interaction between judgment conditions, database match, and blocked or mixed conditions was significant, F (1,98) = 4.522, p = .036, η^2= .044. Consequently, because the three-way interaction was significant, we analyzed only this interaction. First, Hypotheses 2 and 3 were examined as a post hoc analysis, where match condition beta weights were expected to be greater than zero and non-match condition beta weights were expected to be greater than zero in the associative judgment condition. If beta values were zero, then participants could not use the extra cues to judge the relationship between words. We analyzed this data separately for the mixed and block conditions, outlining the different pattern of results from Hypothesis 1. We tested each database match by trial condition combination against zero, resulting in eight post hoc t-tests. Therefore, we used single sample t-tests with a Bonferroni correction experiment-wise to control for Type 1 error rate. The corrected Bonferroni alpha was set to α = .006, which kept α <.05 for eight comparisons experiment-wise. Average beta weights are shown in Figure 1.

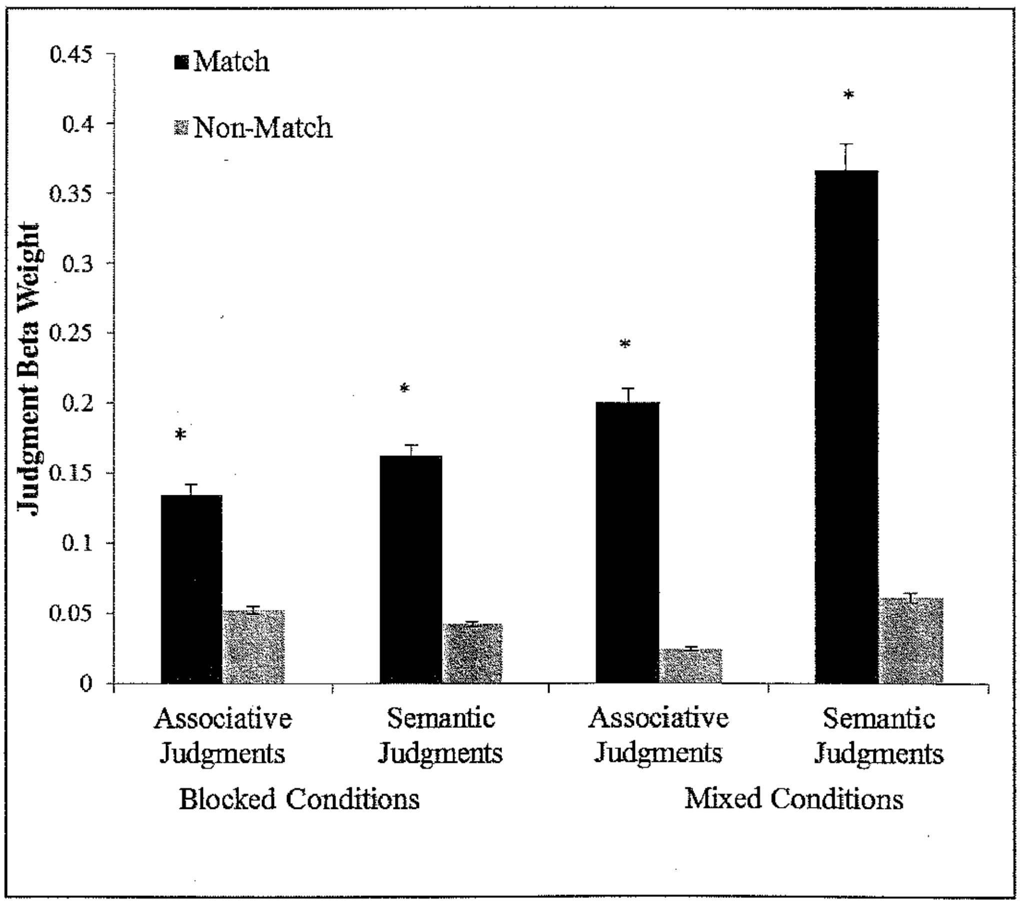

Figure 1. *Associative and Semantic betas for each judgment condition averaged across subjects. Larger beta weights indicate better judgment ability. Match conditions are when judgment type and database comparison match. Error bars are standard error. Starred values represent beta values significantly greater than zero.*

Hypothesis 2 – Database Match Condition Performance

Mixed Trial Condition. As seen in Figure 1, both match conditions were significantly greater than zero, supporting our hypothesis. Associative judgments compared to the associative database scores were greater than zero, t (53) = 7.452, p < .001, d = 2.221. Semantic judgments compared to semantic database scores were significantly greater than zero as well, t (53) = 10.287, p < .001, d = 3.070. When multiple cues were available (i.e., seeing the same first word several times) and mixed together, participants were able to judge word-pairs better than random guessing.

Blocked Trial Condition. The same pattern of results was found for the blocked trial condition. Associative judgments compared to associative database scores were significantly greater than zero, t (45) = 4.520, p < .001, d = 1.347, and semantic judgments compared to semantic database scores, t (45) = 3.266, p = .002, d = .974. Therefore, blocked multiple cues helped participants judge word-pairs better than chance estimation, which implies participants are able to judge specific memory relationships as described in the introduction. The differences in judgment beta weights between blocked and mixed trial conditions will be examined Hypothesis 4 as described above.

Hypothesis 3 – Database Non-Match Condition Performance

Mixed Trial Condition. Analysis of beta weights for opposing judgments and database scores (i.e., associative judgments to semantic databases) illustrated that participant judgments are not influenced by other memory information when making judgments. With the Bonferroni correction, associative judgments related to semantic databases were not significantly different from zero, t (45) = 2.605, p = .012, d = .777. Semantic judgments related to associative database scores were not greater than zero, t (45) = 1.521, p = .135, d = .453.

Blocked Trial Condition. The blocked trial conditions showed the same results as the mixed trial conditions, where the opposing memory information did not influence participant judgments. Associative judgments related to semantic databases was not significant, t (53) = 1.026, p = .309, d = .306, as well as semantic judgments related to associative databases, t (53) = 2.264, p = .028, d = .675 with the Bonferroni correction. Therefore, this hypothesis was not supported.

Hypothesis 4 – Differences Across Trial Conditions

Hypothesis 1 indicated an interaction between judgments, match conditions, and trial conditions. Hypothesis 2 and 3 showed that participants are able to judge memory connections in blocked and mixed trials when compared to the matched database scores. This hypothesis examined the interaction further to indicate if one of the trial conditions showed better judgment scores. Because non-match database beta weights were not significantly different than zero, we only examined the match database beta weights comparing blocked to mixed conditions. Trial condition did not differ for associative judgments, t (98) = 1.615, p = .110, d = 0.326. Therefore, participants are able to make associative judgments about the use of words together with multiple cues, but the presentation of the cues (together or random) did not impact their performance. However, cue presentation did change semantic judgments, where participants did better in the mixed condition (see Figure 1 for mean beta weights) over the blocked condition, $t(98)$ = 3.416, p = .001, d = 0.690.

DISCUSSION

Results from our experiment partially support our hypothesis. First, we found participants' judgment scores, when compared to matching or non-matching databases, were significantly different across mixed and blocked trial conditions (Hypothesis 1 – interaction). As a follow up, we first analyzed if these judgments were better than arbitrary guesses on a Likert scale. Or, more simply, can participants accurately perform the task? Using single sample t-tests, we found participant judgments were better than zero when examining matching judgment and database conditions (Hypothesis 2), but not when examining non-matching judgment to database conditions (Hypothesis 3). Lastly, we assessed if the order of the word-pairs affected participant judgment performance. This analysis revealed t mixing cue words did not change associative judgments, but semantic judgments were improved in the mixed trial conditions over blocked trial conditions (Hypothesis 4).

Maki's (2007a) and Buchanan's (2009) previous studies on judgments were supported by our findings in Hypothesis 2. Maki showed associative judgments are related to the associative database scores, and participants are able to judge word-pairs related by context in memory better than guessing. Buchanan's research portrayed the same results for semantic judgments, where judgments are related to semantic databases. Therefore, this hypothesis was supported. However, Buchanan also showed opposing (as measured by the databases, the non-match conditions) influences on judgments. This prediction was tested in Hypothesis 3 but was not supported. Participants were able to ignore other memory information when instructed to judge only one type of memory – context or meaning. This mechanism could be part of activation suppression, which was suggested by Hutchison and Bosco (2007). Activation suppression occurs when the activation of word linkages is suppressed due to the tasks demands. Each judgment task only requires the use of corresponding information (associative - associative, semantic - semantic); therefore the opposing information activation was dampened during that judgment condition.

Finally, we found partial support for Hypothesis 4, which examined the differences in judgment ability across mixed trial and blocked trial conditions. Maki (2007b) showed participants' judgments did not increase when they were shown four cue-target pairs with the same cue word (e.g., ACHE with HURT, PAIN, HEAD, BACK). In his experiment, participants saw all four cue-target pairs together and made associative judgments on the set of word-pairs. In our experiment, judgment ability was the same in mixed and blocked conditions for associative judgments. This result showed participants were approximately equal at estimating context relationships from memory, regardless of the order the word-pairs were presented.

However, semantic judgments were better in the mixed trial conditions. Participants were better at judging word meaning information when word-pairs were mixed so that the same cue words did not repeat in order. The reason behind this phenomenon could be related to the structure of associative and semantic information in memory. When trying to judge meaning, concept features are activated (i.e., DOG has the feature BARK, TAIL). Then the features are compared for correspondence. When word-pairs were shown together, the previous word processing could interfere with the new word processing for the current judgment. As more words are related in memory, the processing of those concepts slows (Anderson & Reder, 1999) and putting all cues together hindered semantic judgments. In the mixed trial conditions, cue words were randomized, and therefore, the previous word-pair would not interfere with the current word-pair judgment process.

The experiment presented here showed judgments of associative and semantic memory are comparable to previous research even with presentation of four cue-target pairs. Overall, judgments are task appropriate: greater than chance when judgment type is matched to database information, and essentially zero when compared to the opposing database information. Activation suppression is suggested to be a mechanism that aids in the judgment process by creating task appropriate processing. Furthermore, context memory connections (association) were judged equally well in blocked and mixed conditions, indicating that judgments are not affected by word order. Judgments on word meaning (semantics) are positively influenced by mixing words, possibly because the previous information does not interfere with the current task. Further research could examine why differences are found with semantic judgment conditions and not associative judgment conditions, which would not only elucidate judgment processes, but the structure of the underlying memory network as well.

REFERENCES

Anderson, J., & Reder, L. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General, 128*, 186-197. doi:10.1037//0096-3445.128.2.186

Buchanan, E. (2009). Access into memory: Differences in judgments and priming for semantic and associative memory. *Journal of Scientific Psychology*,1-8.

Hutchison, K., & Bosco, F. (2007). Congruency effects in the letter search task: Semantic activation in the absence of priming. *Memory & Cognition, 35*(3), 514-525. doi:10.3758/BF03193291

Koriat, A., & Bjork, R. (2006). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(5), 1133-1145. doi:10.1037/0278-7393.32.5.1133

Jiang, J. J., & Conrath, D. W. (1997). *Semantic similarity based on corpus statistics and lexical taxonomy*. Proceedings of International Conference Research on Computational Linguistics (ROCLING X), Taiwan.

Maki, W. (2007a). Judgments of associative memory. *Cognitive Psychology, 54*(4), 319-353. doi:10.1016/j.cogpsych.2006.08.002

Maki, W. (2007b). Separating bias and sensitivity in judgments of associative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 231-237. doi:10.1037/0278-7393.33.1.231

Maki, W., Krimsky, M., & Muñoz, S. (2006). An efficient method for estimating semantic similarity based on feature overlap: Reliability and validity of semantic feature ratings. *Behavior Research Methods, 38*(1), 153-157. doi:10.3758/BF03192761

Maki, W., McKinley, L., & Thompson, A. (2004). Semantic distance norms computed from an electronic dictionary (WordNet). *Behavior Research Methods, Instruments & Computers, 36* (3), 421-431. doi:10.3758/BF03195590

McRae, K., Cree, G., Seidenberg, M., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37*(4), 547-559. doi:10.3758/BF03192726

Nelson, D., McEvoy, C., & Schreiber, T. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments & Computers, 36* (3), 402-407. doi:10.3758/BF03195588

Schacter, DL, & Tulving, E. (1994). What are the memory systems of 1994? In DL. Schacter and E. Tulving (Eds), *Memory Systems*(pp. 1–38). Cambridge, MA: MIT Press.

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science, 2*(3), 67-70. doi:10.1111/1467-8721.ep10770899

AUTHOR NOTE

Correspondence may be addressed to: Erin Buchanan, Ph.D., Department of Psychology, Missouri State University, 901 S. National Ave., Springfield, Missouri, 65897, Email: erinbuchanan@missouristate.edu, Fax: 417-836-8330.