**Hypotheses**

* Hypothesis 1: Block and mixed trial conditions will show different patterns of judgment ability across semantic and associative judgments. This hypothesis will be tested with a mixed factorial ANOVA.
* 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 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. Both Hypothesis 2 and 3 will be analyzed by examining standardized beta values against zero with a single sample t-test.
* 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. Hypothesis 4 will be analyzed with independent t-tests comparing the between-subjects trial condition variable.

**Results**

The overall subject pool contained 102 participants in total. 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 (associative database-judgment, semantic database-judgment) and for non-matching conditions (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 that 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, *η2*=.409. Overall, participants were better at semantic judgments (*Mbeta* = 0.107, *SE* = 0.015) than associative judgments (*Mbeta* = 0.065, *SE* = 0.014). The main effect of the matching database to judgment was significant, *F*(1,98)=109.051, *p*<.001, *η2*=.527. Participants performed better in the match conditions (*Mbeta* = 0.049, *SE* = 0.017) than the non-match conditions (*Mbeta* = 0.007, *SE* = 0.012). Finally, the blocked versus mixed between groups main effect was not significant, *F*(1,98)=1.435, *p* = .234, *η2*=.014.

All three of the two-way interactions were significant: judgment type and blocked or mixed condition, *F*(1,98)=11.275, *p* =.001, *η2*=.103; database match and blocked or mixed condition, *F*(1,98)=6.750, *p* = .011, *η2*=.064; judgment type and database match, *F*(1,98)=6.358, *p*=.013, *η2*=.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, *η2*=.044. Consequently, since the three-way interaction was significant, we analyzed only this interaction. First, Hypotheses 2 and 3 are examined as a post hoc analysis, where match condition beta weights are expected to be greater than zero and non-match condition beta weights are 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.

**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 that 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 was not significantly different from zero, *t*(45)=2.605, *p*=.012, *d*=.777. Semantic judgments related to associative database scores was 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. Since 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.