Predicting Recall of Words and Lists

Aka, A. and Kahana, M.J.

University of Pennsylvania

Author Note

The authors gratefully acknowledge support from National Institutes of Health grant MH55687. We thank Jonathan Miller and Isaac Pedisich for assistance with designing and programming the experiment, and Patrick Crutchley, Elizabeth Crutchley, Kylie Hower, Adam Broitman, Logan O'Sullivan, and Nicole Kratz, for assistance with data collection.
Correspondence concerning this article should be addressed to Michael J. Kahana (kahana@psych.upenn.edu).

Abstract

For more than a half-century, lists of words have served as the memoranda of choice in studies of human memory. To better understand why some words and lists are easier to recall than others, we estimated subject-specific multivariate models of word and list recall. In each of 23 sessions, subjects (n=88) studied and recalled the same set of 576 words, presented in 24 study-test lists. Fitting subject-specific multivariate models to these data revealed positive effects of word frequency and emotional valence on both word and list recall. Word- and list-level recall models, however, exhibited distinct effects of semantic structure. Whereas words were best recalled when they were similar to many other words in the lexicon, lists were best remembered when their constituent words were similar to one another but different from out-of-list words. These findings point to the importance of network-based measures in understanding the role of similarity on recall performance.

Predicting Recall of Words and Lists

Ever since Ebbinghaus introduced the consonant-vowel-consonant (CVCs) in his seminal experimental analysis of serial learning, students of memory have sought to measure, manipulate, and control the mnemonic difficulty of list materials. Glaze (1928) demonstrated that CVCs varied substantially in their meaningfulness, and Hull (1933) showed that these differences predicted the ease of learning individual items. The fact that CVCs varied considerably in their memorability, and that such variation was often idiosyncratic to individual subjects, was one reason that post-war-era scholars abandoned CVCs in favor of common words as the memoranda of choice in studies of learning and memory. Using words, researchers could draw upon a much larger sample of memoranda, and although words also vary in their memorability, they exhibit less interpretive ambiguity across individuals and require less response learning than do CVCs. Nonetheless, it is of vital importance that researchers have good models for estimating the mnemonic difficulty of learning both individual words and entire lists. The goal of the present study is to construct parallel statistical models to account for variability in free recall of both words and lists (as measures of list-level performance are common), and to assess the stability of these models at the level of individual subjects.

Earlier studies establish the importance of several variables as predictors of word-level recall. Below we discuss findings involving six specific word properties: concreteness, frequency, word length, valence, arousal and meaningfulness. In the case of concreteness[1], prior work has demonstrated that concrete words exhibit a mnemonic advantage over abstract words in a variety of memory paradigms including free recall (Dukes & Bastian, 1966; Paivio, 1967; Hamilton & Rajaram, 2001), item recognition (Gorman, 1961), paired-associates (Epstein, Rock, & Zuckerman, 1960), and immediate serial recall (Walker & Hulme, 1999).

---

[1] Concreteness is defined in terms of directness of reference to sense experience (Paivio, Yuille, & Madigan, 1968)

By contrast, word frequency[2] exhibits a mixed pattern of results, with low frequency words possessing an advantage on recognition tests (Gorman, 1961) and high frequency words possessing an advantage in free recall of pure lists (Hall, 1954). In free recall of mixed lists, both low and high frequency words exhibit superior recall to words of mid-frequencies (Lohnas & Kahana, 2013).

Researchers have found word length[3] to influence memorability of studied items. In the case of immediate serial recall, short words boast a substantial mnemonic advantage (Baddeley, Thomson, & Buchanan, 1975), whereas in free recall results have been mixed (Hulme, Suprenant, Bireta, Stuart, & Neath, 2004; Katkov, Romani, & Tsodyks, 2014; Tehan & Tolan, 2007).

Previous work also found a positive relationship between the emotionality[4] of a word and its memorability. In free recall, emotional words (positively or negatively valent or arousing) are generally better remembered than neutral words (Kensinger & Corkin, 2003; Dolcos, LaBar, & Cabeza, 2004; LaBar & Cabeza, 2006). More recently, some researchers introduced a controversy to the relationship between emotionality and memory. Talmi and Moscovitch (2004) highlighted that emotionality may be serving as an organizing principle, and Hunt, Trammel, and Krumrei-Mancuso (2015) demonstrated that emotion may impair memory for overall meaning for items. More specifically, in Hunt et al.'s study, memory for semantically related emotional lists was worse compared to semantically related neutral lists, suggesting that arousal enhances memory for individual words but interferes with the use of semantic associations as a recall cue.

Whereas the previous five measures pertain to the intrinsic properties of words, one might expect the similarity relations among words to be particularly predictive of item

---

[2] Word frequency refers to the estimate of the frequency usage in the English language

[3] Word length refers how many letters the word consists

[4] Two classical emotionality measures exist: Valence refers to the degree of pleasantness of the word and arousal refers to the degree to which a physiological reaction is elicited by the word.

recall. The highly cue-dependent nature of recall (Kahana, 1996), and the substantial evidence for semantic organization in recall (Howard & Kahana, 2002; Klein, Addis, & Kahana, 2005) indicates that remembering a word will tend to retrieve related words from memory. In a classic study, Noble (1952) defined the meaningfulness (m) of an item as the number of strong associations that a given word possesses with other words, as measured using a free association task (Noble, 1952; Paivio et al., 1968). The positive relationship between meaningfulness and recall performance has been demonstrated in multiple studies (Christian, Bickley, Tarka, & Clayton, 1978; Paivio et al., 1968). More recently, D. Nelson and colleagues systematically investigated the effects of associative networks on performance in a wide range of memory tasks. In cued recall tasks, they found that increasing the number of semantic associations to either the cue or the target word resulted in lower recall rates. In free recall, however, they found less consistent results (Bruza, Kitto, Nelson, & McEvoy, 2009; Nelson, McEvoy, & Pointer, 2003; Nelson & McEvoy, 1979; Nelson, Schreiber, & McEvoy, 1992).

While these previous studies focused on characterizing the effects of each of these variables in isolation, Rubin and Friendly (1986) took a multivariate approach to predict the mnemonic difficulty of individual words. They considered measures of orthography, pronunciability, imagery, concreteness, meaningfulness, availability, familiarity, frequency of occurrence, goodness, and emotionality. Rubin and Friendly (1986)'s results demonstrated that free recall of 925 nouns can be best predicted based on the words' availability, imagery, and emotionality. Additionally, contradicting the findings from traditional literature, Rubin and Friendly questioned the role of meaningfulness, frequency, and pronunciability.

Whereas Rubin and Friendly (1986) sought to predict the memorial difficulty of individual words, here we sought to model both the difficulty of individual words and of entire word lists. We also sought to re-examine the role of word meaningfulness when defined using modern computational linguistic metrics of word similarity (e.g., the Google Word to Vector metric). To examine the role of meaningfulness on both word and list

memory we created two measures: list meaningfulness, computed by measuring semantic similarity of each word to other words in its list, and pool meaningfulness, computed by measuring semantic similarity between each word to all other words in the word pool. Finally, by conducting our analyses of word and list difficulty in a unique multi-session experiment in which each of 88 subjects saw the same pool of 576 words (24 words × 24 lists per session) in each of 23 sessions we were able to evaluate our multivariate model of at the level of individual subjects. This latter feature of our approach allows us to establish the stability of our word and list recall models across individual subjects.

Methods

The data reported here comes from Experiment 4 of the Penn Electrophysiology of Encoding and Retrieval Study (PEERS). The primary goal of PEERS is to assemble a large public database on the electrophysiological correlates of memory encoding and retrieval. Data from Experiments 1-3 have been reported in several prior publications (e.g., Healey, Crutchley, & Kahana, 2014; Healey & Kahana, 2014, 2016; Lohnas & Kahana, 2013, 2014; Lohnas, Polyn, & Kahana, 2015a) and a subset of data from Experiment 4 has been reported in Kahana, Aggarwal, and Phan (2018). Subjects consisted of 88 young adults (ages 18-35) who were recruited from among the students and staff at the University of Pennsylvania and neighboring institutions. All subjects were right-handed and native English speakers.

Subjects performed a delayed free recall experiment consisting of 23 experimental sessions[5]. Each session consisted of 24 trials, with each trial containing a list of 24 words, presented one at a time on a computer screen. A random half of the lists (excluding the first list) were preceded by a 24-second, distractor-filled delay, and all lists were followed by

[5] Subjects participated in a 24th experimental session during which they studied lists composed of both old words (drawn from the pool of 576) and new words matched on the word attributes. Because the focus of this paper is on recallability of words under constant conditions, our analyses do not include data from this last session.

a 24-second distractor period. A free recall test followed the post-list distractor on each list.

Each trial began with a 10-second countdown, which was displayed onscreen. Subjects were permitted to pause and resume this countdown at any time by pressing a key. After the countdown was complete, a fixation cross appeared on the screen for 1500 ms. For trials without a pre-list distractor, the fixation cross was immediately followed by the presentation of the first word. For trials with a pre-list distractor, this fixation cross was instead followed by a 24-second distractor period. After the distractor period, the screen went blank for a jittered 800–1200 ms (uniformly distributed), after which the first word was presented. Each word was presented on the screen in white text on a black background for 1600 ms, and was followed by a jittered inter-stimulus interval of 800–1200 ms (uniformly distributed). Following the inter-stimulus interval after the final word in each list, subjects performed a distractor task for 24 seconds. This post-list distractor task was followed by a 1200–1400 ms (uniformly distributed) delay, after which a tone sounded and a row of asterisks appeared onscreen for 500 ms, indicating the start of the free recall period. subjects were given 75 seconds to recall aloud as many of the words from the current list as possible, in any order. A fixation cross was displayed onscreen for the duration of the recall period followed by a blank screen was displayed for 1000 ms, after which the 10-second countdown for the next list began. Subjects were also given a short break (approximately 5 minutes) after every 8 lists in a session.

Both the pre-list and post-list distractor tasks consisted of answering math problems of the form $A + B + C =?$, where A, B, and C were positive, single-digit integers. Math problems were displayed onscreen one at a time in white text on a black background, and subjects were instructed to type the answer to each equation as quickly and accurately as possible. New problems continued to appear until the full 24 seconds had elapsed, at which point the final problem was immediately removed from the screen. Subjects were given a monetary bonus based on the speed and accuracy of their responses.

Each session required $24 \times 24 = 576$ words. The word pool for this experiment thus

consisted of a 576-words. Each of these 576 words appeared exactly once in each experimental session (24 lists × 24 items), so each subject saw the same set of words 23 times. Within each session, words were randomly assigned to lists following certain constraints on semantic similarity, as described in our earlier PEERS papers. With this experimental design of our multi-session study, we attempted to wash out all the idiosyncratic reasons why certain words might be poorly or well remembered such as whether a word falls into a favorable or unfavorable list position. It should be noted that words in the pool did not have extreme values along dimensions of word frequency, concreteness, and emotional valence as these words are usually omitted from experiments with controlled word pools and we wanted to create a word pool similar to those used in other memory studies.

All previously published raw behavioral data from the PEERS studies, as well as the new data reported in the present manuscript, may be freely obtained from the authors' website, `http://memory.psych.upenn.edu`.

Variables

We created parallel regression models to predict word-level and list-level recall performance. Based on previous work, we identified five properties of words that would be expected to predict recall performance: concreteness, word frequency, word length, valence, and arousal. In addition, we included two variants of Noble's (1952) classic "meaningfulness" index: pool meaningfulness and list meaningfulness, which we define below. While one value for each word is used in the word recall model, an average value computed using each of twenty-four list words' properties is used in the list recall model. For the list recall model we also included variables of trial number and session number to account for proactive interference and practice effects.

Concreteness. Concreteness measures of the 568 words in the word pool is obtained from Brysbaert, Warriner, and Kuperman (2014). These authors collected

concreteness ratings through an Internet crowd-sourcing website by asking subjects to indicate how concrete the meaning of each word is (i.e. can be experienced directly from one of the five senses) using a 5-point rating scale going from abstract to concrete. For example, while the word "apple" is a concrete word that has a concreteness value of 5, the word "patient" is an abstract word that has a concreteness value of 2.5. Eight words in our word-pool do not have a concreteness value reported in Brysbaert et al. (2014), thus we excluded them from our analyses.

Word Frequency. Word frequency measures are obtained through the CELEX2 database (Baayen, Piepenbrock, & Gulikers, 1995), which defines word frequency as counts per million in the Birmingham corpus. In our word pool, for example, while "world" is a common word with high word frequency and "freckle" is a rare word with low word frequency. All words in our word pool has a corresponding word frequency measure in the database.

Word Length. Word length is calculated by counting the number of letters in each word. In our word pool, "playground" is the longest word with ten letters while "ox" is the shortest word with two letters.

Valence. Emotional valence is the degree of pleasantness of the word. Our word pool's emotional valence ratings are obtained from Long, Danoff, and Kahana (2015) . Authors collected ratings through an Internet crowd-sourcing website by asking subjects to give each word a numerical value from 1 to 9, with 1 being unpleasant and 9 being pleasant. For example, while "puppy" is the word with the highest emotional valence (7.78), "poison" is the word with the lowest emotional valence (2.27).

Arousal. Arousal is the degree to which a physiological reaction is elicited by the word. Arousal ratings are obtained from the study mentioned above (Long et al., 2015). Each word has a numerical arousal value from 1 to 9. For example, while "cyclone" is the most arousing word (7.34), "sunset" is the least arousing word (2.75).

Pool Meaningfulness. We define pool meaningfulness as the average semantic relatedness between a given word and all of the words used in our 576 item word pool. This measure is analogous to Noble (1952)'s notion of meaningfulness as the tendency of a word to be associated with other words in the lexicon. In earlier studies (e.g., Paivio et al., 1968; Toglia & Battig, 1978), researchers measured meaningfulness as the number of free associations produced to a given item within a fixed interval (e.g., 60 seconds). In this study, we used modern corpus-based computational methods to determine similarities among words (Mikolov, Chen, Corrado, & Dean, 2013b, 2013a). These methods have been validated in extensive prior work on human memory and cognition (e.g., Bhatia, 2016; Bian, Gao, & Liu, 2014). To compute pool meaningfulness for a given word we calculate the mean of all non-diagonal rows in a vector space model matrix consisting of rows and columns for each word in the word pool. For example, "spoon" has the highest pool meaningfulness value (0.18) while "stake" has the lowest (0.02). This relationship demonstrates that there are many words that are semantically related to "spoon" such as "kitchen" or "fork", while there are very few words semantically related to "stake".

List Meaningfulness. List meaningfulness is analogous to pool meaningfulness, except that we restrict the similarities to only the other items in the target list. As a result, list meaningfulness for each word is calculated by finding the semantic relatedness of each word to twenty-three other words presented in its list. For example, "muffin" has the highest list meaningfulness value (0.13) while "passage" has the lowest (0.12). This means that, on average, words that appeared in the same lists with muffin are highly semantically related to it while words that appeared in the same lists with passage were not.

Two additional variables that were considered in our list recall model were trial number and session number. These variables were added as we think they may influence the list recall performance:

Trial Number. Trial number within a session ranging from 1 to 23.

Session Number. Session number that the data comes from for each subject, ranging from 1 to 23.

Results

A key feature of the present study is that subjects studied and attempted to recall the same set of 576 words in each of 23 separate sessions. As such, a given word would appear in 23 randomly determined list and serial positions for each of the 88 subjects, resulting in a total of 2,024 occurrences of that word across the sample. This allowed us to quantify the recallability of each individual word with a high degree of precision. Figure 1 shows the average recall probability of each word, sorted from lowest to highest, along with the standard error band around this average value. Clearly subjects found some words to be very difficult to recall while other words came to mind easily: the word "survey" exhibited the lowest recall probability (35%) while the word "queen" exhibited the highest recall probability (69%). We first sought to identify those characteristics that determine the recallability of individual words. Next, we will consider characteristics that determine the recallability of lists. Finally, we will compare the factors underlying recallability of words and lists.

To answer these questions we developed two parallel regression models to predict word and list recall performance. We included seven predictor variables that we thought may explain the recall performance of words and word lists based on properties of words in our word pool: concreteness, word frequency, word length, valence, arousal, list meaningfulness, and pool meaningfulness (see *Methods* for details, see *Appendix* for words in the word pool along with their average recall probabilities and associated word properties). For the list recall model we included two additional variables: trial number and session number. Table 2 shows the correlation matrix of the predictors for each model. Both word-level recall and list-level recall correlations are obtained from each subject's model. Although most predictors appear to have weak correlations and are thus suited for

regression analyses, concreteness & pool meaningfulness variables in word recall model, list meaningfulness & pool meaningfulness variables in list recall model, and valence & arousal variables in both models have higher correlations. In our subsequent analyses, to address interpretive problems arising from multi-colineraity, we separately carried out regression analyses on the full set of predictor variables and subsets of variables excluding each member of a highly correlated variable pair (i.e., any pair of variables with a correlation more than ± 0.45).

As we are interested in investigating the contributions of our predictor variables to recall performance of words, we first fit the complete word-level recall model to data from each of the 88 subjects. Figure 2 illustrates results for the seven predictor variables and the distribution of R^2 values from each subject's word recall model. Whereas the dots in the left panel indicate the set of subject-specific β values for each term in the regression model, the thicker lines represent the mean of subject-specific β values for the given variable of the model. Filled circles in the figure indicate those β values that exceed our FDR-corrected significance threshold ($p < 0.05$). Additionally, each dot on the right panel denotes the R^2 values from the models. Models that met our significance threshold ($p < 0.05$) are shaded in black. As may be seen from the distributions of significant coefficients, some variables exhibited consistent positive or negative effects across subjects (e.g., word frequency, and pool meaningfulness), whereas the word length variable exhibited mixed effects, with some subjects having significant positive coefficients and others having significant negative coefficients.

To evaluate which coefficients were reliably positive or negative across subjects, we examined the distribution of β values across subjects for each predictor variable (Table 3). For our model of word-level recall, word frequency, emotional valence, arousal and pool meaningfulness all exhibited highly consistent positive effects (all $t(87)$'s ≥ 9.56, p's< 0.001). List meaningfulness also exhibited a reliably positive distribution of β coefficients, $t(87) = 2.32$, $p < 0.05$.

To determine whether co-linearity among some predictor variables may have biased the estimated coefficients, we fit four reduced models to our data. First, we excluded concreteness and pool meaningfulness, one at a time. Panels in Figure 2 and Table 3 demonstrate the distribution of model β and R^2 values along with their level of significance. When we removed concreteness from the model, word frequency, valence, arousal, and pool meaningfulness all exhibited highly consistent positive effects (all $t(87)$'s ≥ 9.56, p's< 0.001). List meaningfulness also exhibited a reliably positive distribution of β coefficients, $t(87) = 2.34$, $p < 0.05$.

Next, when we removed pool meaningfulness from the model, in addition to the significant variables from the models above, we also observed the highly consistent positive influence of concreteness on recall performance. In other words, distribution of β coefficients of concreteness, word frequency, valence, arousal, and list meaningfulness were all highly consistent and positive (all $t(87)$'s ≥ 7.22, p's< 0.001).

Our reduced word-level recall model without valence demonstrated the highly consistent positive influence of word frequency, arousal, and pool meaningfulness on recall, (all $t(87)$'s ≥ 9.74, p's< 0.001), while the reduced model without arousal demonstrated the reliable negative influence of concreteness, $t(87) = -3.46$, $p < 0.01$, and highly consistent positive influence of word frequency, valence, and pool meaningfulness, (all $t(87)$'s ≥ 5.48, p's< 0.001). List meaningfulness also had reliable positive influence on these two reduced models, (both $t(87)$'s ≥ 2.19, p's< 0.05).

In each of the multivariate models of word-level recall, we observed strong positive effects of word-frequency, valence, arousal, and pool meaningfulness on recall performance. List meaningfulness also predicted successful word-level recall, but its effect was weaker than that of the other predictor variables. In our reduced model that excluded pool meaningfulness we observed a significant positive effect of concreteness, as has been reported in earlier studies (Dukes & Bastian, 1966; Paivio, 1967; Hamilton & Rajaram, 2001). Our results did not highlight any reliable effects of word length on memory.

The preceding analyses of word-level predictors of recall ignore the fact that recalling a list is a dynamic, path-dependent, process. Because we often evaluate memory using measures of list-level performance, both in pure and applied settings, it is critically important to be able to model variability in memory performance at the level of lists. Furthermore, a comparison of list-level and word-level models may uncover organizational principles supporting recall of entire lists that would be missed using an item-level analysis. As such, we next sought to evaluate a multivariate model of list-level recall using the same predictor variables as in the word-level recall model.

Figure 3 and Table 4 demonstrate that in the full list recall model, word frequency and list meaningfulness exhibited highly consistent positive influence (both $t(87)$'s ≥ 3.74, p's< 0.001), and emotional valence demonstrated also reliable but weaker positive influence $t(87) = 3.58$, $p < 0.01$. However, pool meaningfulness β coefficients exhibited consistent negative influence ($t(87) = -3.59$, $p < 0.01$).

Similar to the approach used in word recall model, we next fit our data to reduced list recall models. More specifically, we first excluded pool meaningfulness and list meaningfulness, one at a time. When we removed pool meaningfulness from the list recall model, word frequency, valence, and list meaningfulness exhibited consistent positive influence (all $t(87)$'s ≥ 3.58, p's< 0.001). Similarly, when we removed list meaningfulness from the list recall model, word frequency and valence exhibited reliable positive influence (both $t(87)$'s ≥ 3.40, p's< 0.01).

Our reduced list-level recall model without valence demonstrated highly consistent positive influence of word frequency and list meaningfulness (both $t(87)$'s ≥ 4.28, p's< 0.001) along with the reliable negative influence of pool meaningfulness, $t(87) = -3.40$, $p < 0.01$. Finally, the reduced list-level recall model without arousal demonstrated highly consistent positive influence of word frequency, valence, and list meaningfulness (all $t(87)$'s ≥ 3.63, p's< 0.001) and highly consistent negative influence of pool meaningfulness, $t(87) = -3.67$, $p < 0.001$.

Our analysis of list-level recall performance revealed strong positive effects of average word-frequency, valence, and list meaningfulness as well as strong negative effects of pool meaningfulness. This overall pattern appeared consistently in both the full and the reduced models. Our results did not demonstrate any reliable influence of average concreteness, word length, or arousal. Lastly, if we look at the two additional variables of the list recall model, we see that while trial number exhibited highly negative influence in all models (all $t(87)$'s $\leq$ -15.30, p's$<$ 0.001), session number did not have any consistent effect on overall list recall.

The foregoing analyses demonstrate that we can reliably account for both item-level and list-level recall, however the specific weighting of each predictor variable differed across the two sets of models. Whereas word frequency and emotional valence exhibited consistent positive correlations with both item and list-level recall, pool meaningfulness exhibited opposite effects, being positively correlated with item-level recall and negatively correlated with list-level recall. List meaningfulness, which was one of the strongest predictors of variability in list-level recall, was only weakly associated with item-level recall.

General Discussion

We asked how the properties of words influence their memorability. To answer this question we developed multivariate regression models to account for variability in both item and list-level recall. Our model included six standard predictor variables (concreteness, frequency, word length, emotional valence, and arousal) as well as two new "meaningfulness" measures designed capture a word's semantic relatedness to other words in the lexicon (pool meaningfulness) or to other words in the target list (list meaningfulness). In the item-level model, we assessed how these predictor variables accounted for variability in recall probability across the 576 items seen by each subject in our study. In the list-level model, we evaluated how the same predictor variables, averaged across all items in a given list, accounted for that list's average recall probability. We

applied this model to all 48,576 lists seen by the 88 subjects in our study and included covariates to account for practice and interference effects. Results from the item-level model replicated previous findings regarding the positive influence of word frequency, emotional valence and emotional arousal on recall (Hall, 1954; Kensinger & Corkin, 2003; Dolcos et al., 2004; LaBar & Cabeza, 2006). The present study extended these previous results by showing that frequency and valence also predict variability in list-level recall.

The primary goal of the present study was to determine the influence of semantic structure on the recallability of items and lists. Using the Word2Vec algorithm (Mikolov et al., 2013b) we were able to assess an item's semantic relatedness both to its neighbors on a given list (list meaningfulness) and to the larger pool of 576 items used in our study (pool meaningfulness). This latter measure should be reflective of an item's overall similarity the words in the lexicon.

Our analyses revealed that subjects tend to recall words that have high meaningfulness to the entire word pool (and presumably to all other words in the lexicon). It is easy to imagine how pool meaningfulness may facilitate word-level recall through semantic elaboration during encoding. If, during study of a given word (e.g. "flower"), subjects elaborate the word's meaning in relation to other related words (e.g. "rabbit", "bouquet", "tree") there will be a large set of items that could potentially help cue that particular word during recall. Because high pool-meaningfulness words are densely interconnected in semantic similarity space other words within a list are likely to facilitate their retrieval, even if only indirectly. One would similarly expect that list meaningfulness should also improve recall of individual items, as they will tend to be cued by other items on the target list, which is exactly what we observed.

Whereas earlier studies also identified a positive correlation between concreteness (or, analogously, imageability) and word recall (Dukes & Bastian, 1966; Paivio, 1967; Hamilton & Rajaram, 2001), this was not observed in our full multivariate model. We can see this apparent discrepancy is resolved when we consider that concrete words also tend to exhibit

high pool meaningfulness (e.g., the correlation between these measures was 0.47 in our sample, $p < 0.001$). When we evaluated the reduced model excluding pool meaningfulness (see Table 3) we recovered a significant positive influence of concreteness on item level recall. However, we found a superior overall fit for the reduced model that excluded concreteness as compared with the reduced model that excluded pool meaningfulness.

In a contemporaneous report, Lau, Goh, and Yap (2018) also examined predictors of item-level recall performance. As in the present study, they found recall probability to be significantly positively correlated with word frequency, arousal, and a measure of semantic density that is nearly identical to our measure of pool meaningfulness. Both studies found consistent effects despite several major methodological differences, such as the use of naive versus practiced subjects, the immediate vs. delayed nature of recall, and the statistics of item characteristics in the pools being used.

Whereas both the present work and earlier studies considered how word properties predict recall of individual items, here we also considered how these predictor variables could account for overall levels of list recall. Given that we usually assess subject's performance at the list-level, understanding the predictors of list-level recall can have important practical value in both designing experiments and optimizing neuropsychological measures of memory function.

Our multivariate analysis of list-level recall revealed significant positive contributions of average word frequency and average valence (as in the word recall model). In addition, we found a significant positive contribution of average list meaningfulness but a significant negative contribution of average pool meaningfulness. Because each recalled item serves as a cue for subsequent recalls, lists with semantically related items will be easier to recall, consistent with our list meaningfulness findings. In contrast, lists whose items are semantically related to many non-target items will have higher levels of proactive interference, as it is known that similarity across lists tends to evoke intrusions (Zaromb et al., 2006; Lohnas, Polyn, & Kahana, 2015b). In contrast to the item recall model, where

several predictor variables each accounted for a comparable share of the variance, in the list recall recall, only list meaningfulness strongly predicted recall. Our findings align with previous work by Nelson and colleagues (Nelson et al., 2003) and by Anderson (Anderson, 1974) showing how semantic associative networks can be a source of both positive and negative transfer in cued-recall tasks. Here we show how similar ideas can help to explain variability in the recallability of items and lists in the setting of delayed free recall.

Conclusion

For over a century, students of memory have turned to common words as the memoranda of choice in their experiments. As such, understanding how word properties relate to their memorability has attracted considerable attention (Schlosberg & Woodworth, 1954; Rubin & Friendly, 1986). In this paper we exploit powerful new methods for measuring meaningfulness, both at the level of individual words and entire lists, and show the importance of both metrics. At the level of items, both list and pool meaningfulness positively predict recall, whereas at the level of lists, these two factors have opposite effects, with list meaningfulness positively predicting recall and pool meaningfulness negatively predicting recall. We can understand both of these effects as reflecting the powerful role of semantic similarity in providing a cue for transitions between target items and also as a source of interference when list items trigger memories of non-target items. In extending the analysis of linguistic and semantic factors in predicting recall from the item-level to the list-level, our study provides valuable new experimental design tools. Specifically, these models can help to optimize experimental design to better control variability in list-level recall performance, both for more accurate assessment of individual differences and influence of experimental variables on recall.

References

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*(4), 451–474. doi: 10.1016/0010-0285(74)90021-8

Baayen, R., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (Release 2). *Linguistic Data Consortium. Philadelphia (PA): University of Pennsylvania.*.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. doi: 10.1016/S0022-5371(75)80045-4

Bhatia, S. (2016). Vector Space Semantic Models Predict Subjective Probability Judgments for Real-World Events. In *Proceedings of the 38th annual conference of the cognitive science society* (pp. 1937–1942).

Bian, J., Gao, B., & Liu, T. Y. (2014). Knowledge-powered deep learning for word embedding. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8724 LNAI*(PART 1), 132–148. doi: 10.1007/978-3-662-44848-9_9

Bruza, P., Kitto, K., Nelson, D., & McEvoy, C. (2009). Extracting spooky-activation-at-a-distance from considerations of entanglement. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 5494, pp. 71–83). doi: 10.1007/978-3-642-00834-4_8

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. doi: 10.3758/s13428-013-0403-5

Christian, J., Bickley, W., Tarka, M., & Clayton, K. (1978). Measures of free recall of 900 English nouns: Correlations with imagery, concreteness, meaningfulness, and frequency. *Memory & Cognition*, *6*(4), 379–390. doi: 10.3758/BF03197470

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the amygdala and the

medial temporal lobe memory system predicts better memory for emotional events. *Neuron*, *42*(5), 855–863. doi: 10.1016/S0896-6273(04)00289-2

Dukes, W. F., & Bastian, J. (1966). Recall of abstract and concrete words equated for meaningfulness. *Journal of Verbal Learning and Verbal Behavior*, *5*(5), 455–458. doi: 10.1016/S0022-5371(66)80060-9

Epstein, W., Rock, I., & Zuckerman, C. B. (1960). Meaning and familiarity in associative learning. *Psychological Monographs: General and Applied*, *74*(4, Whole No. 491), 22.

Glaze, J. A. (1928). The Association Value of Non-Sense Syllables. *Pedagogical Seminary and Journal of Genetic Psychology*, *35*(2), 255–269. doi: 10.1080/08856559.1928.10532156

Gorman, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology*, *61*(1), 23–29. doi: 10.1037/h0040561

Hall, J. F. (1954). Learning as a function of word-frequency. *The American journal of psychology*, *67*(1), 138–140. doi: 10.2307/1418830

Hamilton, M., & Rajaram, S. (2001). The Concreteness Effect in Implicit and Explicit Memory Tests. *Journal of Memory and Language*, *44*(1), 96–117. doi: 10.1006/jmla.2000.2749

Healey, M. K., Crutchley, P., & Kahana, M. J. (2014). Individual differences in memory search and their relation to intelligence. *Journal of Experimental Psychology: General*, *143*(4), 1553–1569. doi: 10.1037/a0036306

Healey, M. K., & Kahana, M. J. (2014). Is memory search governed by universal principles or idiosyncratic strategies? *Journal of Experimental Psychology: General*, *143*(2), 575–596. doi: 10.1037/a0033715

Healey, M. K., & Kahana, M. J. (2016). A four-component model of age-related memory change. *Psychological Review*, *123*(1), 23–69. doi: 10.1037/rev0000015

Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic

retrieval? *Journal of Memory and Language*, *46*(1), 85–98. doi: 10.1006/jmla.2001.2798

Hull, C. (1933). The Meaningfulness of 320 Selected Nonsense Syllables. *The American journal of psychology*, *45*(4), 730–734.

Hulme, C., Suprenant, A. M., Bireta, T. J., Stuart, G., & Neath, I. (2004). Abolishing the word-length effect. *Journal of experimental psychology. Learning, memory, and cognition*, *30*(1), 98–106. doi: 10.1037/0278-7393.30.1.98

Hunt, C., Trammel, J., & Krumrei-Mancuso, E. J. (2015). The Interaction of Semantic Relatedness and Emotion: Why Emotion May Not Lead to Better Memory. *Psi Chi Journal of Psychological Research*, *20*(1), 45–51.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103–109. doi: 10.3758/BF03197276

Kahana, M. J., Aggarwal, E. V., & Phan, T. D. (2018). The Variability Puzzle in Human Memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*.

Katkov, M., Romani, S., & Tsodyks, M. (2014). Word length effect in free recall of randomly assembled word lists. *Frontiers in computational neuroscience*, *8*(October), 129. doi: 10.3389/fncom.2014.00129

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. doi: 10.3758/BF03195800

Klein, K. A., Addis, K. M., & Kahana, M. J. (2005). A comparative analysis of serial and free recall. *Memory & Cognition*, *33*(5), 833–839. doi: 10.3758/BF03193078

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*(1), 54–64. doi: 10.1038/nrn1825

Lau, M. C., Goh, W. D., & Yap, M. J. (2018). An item-level analysis of lexical-semantic effects in free recall and recognition memory using the megastudy approach. *Quarterly Journal of Experimental Psychology*, *0*(0), 1747021817739834. doi:

10.1177/1747021817739834

Lohnas, L. J., & Kahana, M. J. (2013). Parametric effects of word frequency in memory for mixed frequency lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1943–1946. doi: 10.1037/a0033669

Lohnas, L. J., & Kahana, M. J. (2014). Compound cuing in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 12–24. doi: 10.1037/a0033698

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015a). Expanding the scope of memory search: Intralist and interlist effects in free recall. *Psychological Review*, *122*(2), 337–363. doi: 10.1037/a0039036

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015b). Expanding the scope of memory search: Modeling intralist and interlist effects in free recall. *Psychological Review*, *122*(2), 337–363. doi: 10.1037/a0039036

Long, N. M., Danoff, M. S., & Kahana, M. J. (2015). Recall dynamics reveal the retrieval of emotional context. *Psychonomic Bulletin and Review*, *22*(5), 1328–1333. doi: 10.3758/s13423-014-0791-2

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013a). Distributed Representations ofWords and Phrases and their Compositionality. *CrossRef Listing of Deleted DOIs*, *1*, 1–9. doi: 10.1162/jmlr.2003.3.4-5.951

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013b). Efficient Estimation of Word Representations in Vector Space. , 1–12. doi: 10.1162/153244303322533223

Nelson, D. L., & McEvoy, C. L. (1979). Encoding context and set size. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(3), 292–314. doi: 10.1037/0278-7393.5.3.292

Nelson, D. L., McEvoy, C. L., & Pointer, L. (2003). Spreading Activation or Spooky Action at a Distance? *Journal of Experimental Psychology: Learning Memory and Cognition*, *29*(1), 42–52. doi: 10.1037//0278-7393.29.1.42

Nelson, D. L., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review*, *99*(2), 322–348. doi: 10.1037/0033-295X.99.2.322

Noble, C. E. (1952). An Analysis of Meaning. *Psychological Review*, *59*(6), 421–430. doi: 10.1037/h0054087

Paivio, A. (1967). Paired-Associate Learning and Free Recall of Nouns As a Function of Concreteness, Specificity, Imagery, and Meaningfulness. *Psychological Reports*, *20*, 239–245.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1, Pt.2), 1–25. doi: 10.1037/h0025327

Rubin, D. C., & Friendly, M. (1986). Predicting which words get recalled: measures of free recall, availability, goodness, emotionality, and pronunciability for 925 nouns. *Memory & cognition*, *14*(1), 79–94. doi: 10.3758/BF03209231

Schlosberg, H., & Woodworth, R. S. (1954). *Experimental Psychology.* Holt.

Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition*, *32*(5), 742–751. doi: 10.3758/BF03195864

Tehan, G., & Tolan, G. A. (2007). Word length effects in long-term memory. *Journal of Memory and Language*, *56*(1), 35–48. doi: 10.1016/j.jml.2006.08.015

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* (vii, 152 ed.). Oxford, England: Lawrence Erlbaum.

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1256–1271. doi: 10.1037/0278-7393.25.5.1256

Zaromb, F. M., Howard, M. W., Dolan, E. D., Sirotin, Y. B., Tully, M., Wingfield, A., &

Kahana, M. J. (2006). *Temporal associations and prior-list intrusions in free recall* (Vol. 32) (No. 4). doi: 10.1037/0278-7393.32.4.792

Appendix

Word	Recall Prob	Concreteness	Frequency	Valence	Arousal	M List	M Pool
QUEEN	0.687	4.45	889	5.77	5.47	0.128	0.123
BOYFRIEND	0.686	4.59	80	6.3	5.77	0.127	0.133
WIFE	0.678	4.13	3777	6.64	5.05	0.125	0.117
PRINCESS	0.675	4.72	217	6.48	5.32	0.128	0.138
SISTER	0.673	4	1464	6.52	4.43	0.127	0.100
GIRL	0.669	4.85	4944	6.77	5.29	0.130	0.149
BROTHER	0.656	4.43	1531	6.41	4.88	0.126	0.104
OX	0.653	4.86	79	5.15	4.81	0.127	0.129
HUSBAND	0.651	4.11	2405	6.54	4.89	0.130	0.127
ASIA	0.650	–	475	5.96	5.44	0.121	0.049
WOMAN	0.647	4.46	6072	6.94	5.45	0.131	0.146
DAUGHTER	0.646	4.79	1797	6.8	4.71	0.128	0.127
LONDON	0.631	–	5368	6.43	5.7	0.120	0.054
JAPAN	0.628	4.82	541	6.09	5.53	0.121	0.063
LOVER	0.623	3.68	430	7.74	6.47	0.127	0.139
TEACHER	0.618	4.52	1490	6.11	5.34	0.124	0.100
STUDENT	0.617	4.92	3587	5.87	5.19	0.124	0.084
WAITRESS	0.616	4.56	138	5.95	5.02	0.130	0.140
ACTRESS	0.614	4.54	254	6.09	5.45	0.124	0.093
INFANT	0.613	4.93	372	6.27	4.87	0.130	0.119
BABY	0.610	5	3281	6.69	4.84	0.129	0.151
POLICE	0.608	4.79	3694	4.68	7.09	0.123	0.090
UNCLE	0.608	4.24	1061	6.22	4.45	0.126	0.123
LADY	0.607	4.33	1219	6.65	5.15	0.129	0.139
COD	0.605	4.61	87	5.01	4.44	0.126	0.099
PRINCE	0.604	4.44	592	5.85	5.51	0.126	0.116
SIBLING	0.598	4.37	26	6.33	4.97	0.126	0.083
PARENT	0.598	4.56	1266	6.42	4.71	0.124	0.055
FRIAR	0.595	3.88	19	5.14	4.27	0.126	0.111
EUROPE	0.594	–	2784	6.43	5.48	0.120	0.040
EGYPT	0.593	–	584	5.62	5.84	0.125	0.079
PUPPY	0.590	4.78	87	7.78	4.99	0.132	0.156
FRANCE	0.589	–	1779	6.15	5.49	0.118	0.055

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
SPOUSE	0.588	3.85	50	6.82	5.02	0.127	0.096
COLLEGE	0.588	4.62	1380	6.21	5.64	0.121	0.076
BISON	0.588	4.68	26	5.55	5.03	0.127	0.105
CHILD	0.587	4.78	7645	6.21	5.02	0.129	0.133
NURSE	0.587	4.39	556	6.14	5.2	0.126	0.109
ARM	0.586	4.96	1860	5.44	4.61	0.125	0.096
MEAT	0.583	4.9	1280	5.9	5.4	0.129	0.134
SHERIFF	0.580	4.5	49	4.73	6.53	0.123	0.087
ORANGE	0.580	4.66	258	6.65	4.75	0.129	0.122
COUSIN	0.579	3.7	400	6.21	4.59	0.128	0.121
LION	0.578	4.96	152	5.77	6.78	0.125	0.101
ANKLE	0.577	4.81	185	5.23	4.53	0.125	0.087
ARMY	0.577	4.7	1933	4.76	6.64	0.123	0.090
FEMALE	0.575	4.57	1044	6.91	5.48	0.126	0.083
FRIEND	0.574	3.07	3087	7.68	4.5	0.127	0.122
MAILMAN	0.574	4.57	3	5.66	4.92	0.127	0.125
CHAUFFEUR	0.573	4.43	83	5.5	4.68	0.126	0.116
HORSE	0.573	5	1518	6.29	4.88	0.130	0.129
SPHINX	0.573	4.83	15	5.83	5.28	0.129	0.125
LEOPARD	0.573	5	118	5.86	6.23	0.129	0.139
COLONEL	0.572	3.89	1610	4.99	5.61	0.124	0.098
DOLPHIN	0.571	4.96	30	6.75	4.34	0.128	0.131
APPLE	0.571	5	315	6.86	4.54	0.131	0.153
FIREMAN	0.570	4.8	20	6.35	6.35	0.129	0.136
PREACHER	0.570	4.7	58	4.91	4.88	0.124	0.107
LEG	0.570	4.83	1137	5.49	4.69	0.125	0.112
ZEBRA	0.569	4.86	12	6.27	4.82	0.127	0.132
LILY	0.568	4.69	120	6.85	3.38	0.128	0.145
OYSTER	0.568	4.85	64	4.7	4.46	0.129	0.127
AIRPORT	0.568	4.87	932	5.37	6.26	0.125	0.080
STALLION	0.568	4.72	24	6.18	5.8	0.123	0.090
BEEF	0.567	4.74	273	5.93	5.32	0.126	0.087
BREAST	0.567	4.89	774	6.38	6	0.126	0.103
COWBOY	0.567	4.72	84	5.81	5.65	0.126	0.122
GAZELLE	0.566	4.72	13	6.06	4.61	0.127	0.114

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
SCALPEL	0.566	4.86	18	3.8	6.64	0.127	0.114
LAKE	0.565	4.88	718	6.85	3.38	0.127	0.126
GRIZZLY	0.565	3.17	3	4.22	6.88	0.126	0.119
THIEF	0.565	4.37	120	2.46	7.03	0.128	0.130
DENTIST	0.565	4.93	115	3.52	6.91	0.126	0.110
INMATE	0.564	4.19	106	2.71	6.82	0.124	0.109
CHAMPAGNE	0.562	4.82	285	6.33	5.41	0.125	0.098
RIFLE	0.561	4.85	286	3.93	7.06	0.128	0.121
KITCHEN	0.561	4.97	1890	6.18	4.54	0.132	0.156
FARMER	0.560	4.54	1017	6.14	4.2	0.127	0.113
CREVICE	0.559	4.43	33	4.31	5.16	0.127	0.135
WRIST	0.559	4.93	358	5.23	4.61	0.126	0.094
KNIFE	0.559	4.9	635	4.18	6.64	0.130	0.148
SKIRT	0.558	4.82	351	6.18	5.41	0.128	0.115
PALACE	0.558	4.57	761	6.35	5.39	0.127	0.110
CRIB	0.557	4.86	19	5.66	4.1	0.130	0.132
ACTOR	0.557	4.57	785	5.76	5.6	0.127	0.096
YACHT	0.556	4.97	83	6.11	4.74	0.122	0.087
SAUSAGE	0.556	4.88	103	5.75	5.06	0.131	0.148
CONVICT	0.555	4.11	30	2.39	7.14	0.123	0.091
TORTOISE	0.555	4.87	71	5.95	3.52	0.128	0.139
VAGRANT	0.554	3.46	9	3.43	6.04	0.129	0.132
BACON	0.554	4.9	288	6.77	5.67	0.129	0.135
IGLOO	0.554	4.73	8	5.71	4.48	0.126	0.119
GARDEN	0.553	4.73	1954	6.94	3.22	0.130	0.147
PEACH	0.552	4.9	47	6.81	4.35	0.131	0.155
KETCHUP	0.552	5	26	5.97	4.63	0.125	0.137
BUNNY	0.551	4.97	19	7.03	3.64	0.130	0.163
QUAIL	0.550	4.65	7	5.69	4.35	0.131	0.133
GANGSTER	0.550	3.93	50	2.84	7.06	0.128	0.103
BASEMENT	0.550	4.89	224	4.98	4.67	0.127	0.106
PLAYGROUND	0.550	4.77	211	6.72	5.34	0.123	0.106
AIRPLANE	0.550	4.96	73	5.68	6.15	0.128	0.125
SUMMIT	0.549	4.21	241	5.86	5.37	0.120	0.055
HIKER	0.548	4.53	12	5.8	5.4	0.125	0.120

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
TROUT	0.547	4.72	280	5.38	4.52	0.129	0.126
WIDOW	0.547	4.33	254	3.57	5.47	0.129	0.110
PENGUIN	0.547	5	69	6.63	4.7	0.130	0.146
PILOT	0.547	4.67	235	5.85	5.76	0.121	0.066
CROWN	0.547	4.81	412	5.93	5.42	0.124	0.081
FOX	0.546	4.97	177	5.84	5.66	0.129	0.156
PLAID	0.546	4.23	49	5.37	4.41	0.128	0.121
MUMMY	0.546	4.72	198	4.25	5.71	0.130	0.157
DORM	0.546	4.41	72	5.12	5.03	0.125	0.097
TOMBSTONE	0.546	4.71	29	3.12	5.65	0.128	0.149
PAIL	0.545	4.93	48	5.08	4.26	0.130	0.166
CHEMIST	0.544	4.24	238	5.49	5.49	0.121	0.084
SALMON	0.544	4.81	108	5.6	4.6	0.130	0.118
DRAGON	0.544	4.39	135	5.39	6.96	0.128	0.143
MARS	0.544	4.48	166	6.01	5.48	0.121	0.041
CHURCH	0.543	4.9	2844	4.86	4.14	0.126	0.100
DRESS	0.543	4.93	1332	6.09	4.66	0.129	0.131
BOX	0.543	4.9	704	5.17	4.57	0.125	0.099
FOOT	0.543	4.9	1753	5.05	4.58	0.125	0.104
BUTCHER	0.542	4.65	87	4.4	6.05	0.127	0.140
TYPIST	0.542	4.41	38	5.37	4.74	0.125	0.099
THRONE	0.541	4.64	175	5.57	5.64	0.125	0.085
CATTLE	0.541	4.64	568	5.35	4.6	0.128	0.105
FINGER	0.540	5	867	5.47	4.75	0.126	0.103
BEAVER	0.539	4.68	37	5.56	4.91	0.130	0.148
BLOUSE	0.539	4.96	156	5.89	4.64	0.129	0.138
LETTUCE	0.539	4.97	115	5.81	4.16	0.130	0.147
BELLY	0.539	4.8	305	4.83	4.9	0.130	0.147
SEAFOOD	0.539	4.83	9	5.74	5.02	0.126	0.107
LACE	0.539	4.85	186	6.3	4.26	0.128	0.120
SHORTS	0.539	4.82	195	5.77	4.57	0.125	0.085
SKILLET	0.539	4.73	2	5.63	4.95	0.128	0.133
SHARK	0.538	4.93	246	4.11	7.26	0.125	0.123
BANKER	0.537	4.43	129	4.54	5.44	0.123	0.086
FOREST	0.537	4.76	1219	6.65	3.81	0.125	0.101

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
HIGHWAY	0.537	4.72	285	5.15	5.68	0.123	0.094
COOK	0.536	4.32	266	6.32	4.89	0.129	0.128
JUDGE	0.536	3.75	732	4.75	6.17	0.124	0.076
TUTU	0.536	4.68	0	5.53	4.92	0.129	0.141
ELBOW	0.536	5	280	5.13	4.6	0.126	0.102
WAITER	0.536	4.67	299	5.7	4.82	0.127	0.145
LIPSTICK	0.536	4.9	122	5.71	5.21	0.128	0.123
LIZARD	0.536	4.68	47	4.81	5.59	0.132	0.175
HOUND	0.536	4.48	33	5.35	5.28	0.126	0.101
SHIRT	0.536	4.94	812	5.78	4.41	0.128	0.132
OMELET	0.535	4.93	1	6.37	4.61	0.130	0.142
WINDOW	0.535	4.86	2372	5.92	4.23	0.125	0.121
TRANSPLANT	0.535	3.77	49	4.73	6.03	0.125	0.083
PIZZA	0.535	5	28	7.37	5.55	0.131	0.132
SIDEWALK	0.535	4.96	104	5.44	4.35	0.126	0.122
MATTRESS	0.535	5	170	6.05	3.33	0.129	0.135
NIGHTGOWN	0.534	4.9	45	5.81	4.31	0.133	0.169
VENUS	0.534	4.54	108	6.12	5.29	0.123	0.110
LAGOON	0.534	4.5	56	6.2	3.8	0.128	0.107
CABIN	0.534	4.92	485	6.26	3.52	0.127	0.119
SWIMMER	0.534	4.77	41	6.01	4.96	0.124	0.099
CABBAGE	0.534	4.75	144	4.95	4.33	0.131	0.151
ISLAND	0.533	4.96	1209	6.89	3.77	0.124	0.097
PISTOL	0.533	4.89	254	3.81	7.26	0.127	0.127
SLIME	0.533	4.48	35	2.97	5.91	0.127	0.118
RABBIT	0.533	4.93	189	6.52	4.41	0.133	0.179
CASKET	0.532	4.86	35	2.73	6.22	0.127	0.122
FRUIT	0.532	4.81	513	6.99	4.49	0.129	0.134
JEANS	0.531	5	229	6.15	4.52	0.129	0.119
RIVER	0.531	4.89	1938	6.62	3.9	0.126	0.114
CARRIAGE	0.531	4.86	231	5.67	4.41	0.127	0.110
ONION	0.530	4.86	170	5.21	4.98	0.130	0.139
SNAKE	0.530	5	251	3.64	6.89	0.132	0.178
MUSTARD	0.530	4.93	84	5.17	4.66	0.126	0.113
PICNIC	0.530	4.83	184	7.03	3.63	0.127	0.120

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
DONKEY	0.529	5	162	5.29	4.94	0.128	0.147
HAND	0.529	4.72	7889	5.58	4.52	0.126	0.104
CEILING	0.529	4.85	463	5.3	4.2	0.128	0.098
SILK	0.529	4.7	440	6.72	3.77	0.128	0.126
COTTAGE	0.529	4.85	555	6.29	3.42	0.127	0.124
BANQUET	0.529	4	67	6.3	5.33	0.125	0.080
JURY	0.528	4.64	521	4.53	6.1	0.123	0.070
BLADE	0.527	4.93	205	4.42	6.53	0.130	0.136
MILK	0.527	4.92	1792	6.04	3.99	0.128	0.115
MISSILE	0.527	4.83	465	3.22	7.07	0.126	0.076
SHOULDER	0.526	4.93	1196	5.38	4.57	0.125	0.094
ROD	0.526	4.43	503	4.91	5.13	0.128	0.142
SCALLOP	0.526	4.61	17	5.34	4.59	0.128	0.122
OUTLAW	0.526	3.61	8	3.49	6.91	0.122	0.064
CLAM	0.526	4.89	13	4.95	4.23	0.128	0.132
FIREPLACE	0.525	4.68	148	6.55	3.95	0.128	0.142
DANCER	0.525	4.75	146	6.31	5.73	0.127	0.106
HUMAN	0.525	4.93	5113	6.41	5.21	0.123	0.085
POISON	0.524	4.27	204	2.27	7.2	0.127	0.108
DUST	0.524	4.4	762	3.55	4.72	0.127	0.103
DRIVER	0.524	4.71	746	5.67	5.54	0.123	0.091
CARROT	0.524	5	99	5.97	4.31	0.129	0.117
SLUG	0.524	4.64	38	3.45	5.06	0.126	0.105
APRON	0.524	4.87	131	5.29	4.07	0.130	0.147
BURGLAR	0.523	4.44	44	2.38	7.31	0.127	0.125
COOKBOOK	0.523	4.9	8	6.22	4.24	0.128	0.101
ORCHID	0.523	4.92	44	6.88	3.58	0.129	0.152
VIKING	0.523	–	17	5.1	6.11	0.126	0.108
VELVET	0.523	4.44	206	6.21	3.7	0.129	0.141
MARINE	0.522	4.25	183	5.78	5.61	0.122	0.072
CHIPMUNK	0.522	4.97	2	6.2	4.65	0.129	0.154
HOSTESS	0.522	4.12	164	5.9	5.06	0.125	0.100
JELLY	0.522	4.93	172	5.98	4.56	0.130	0.148
LIGHTNING	0.521	4.59	253	4.92	6.81	0.124	0.076
ROBBER	0.521	4.31	42	2.31	7.13	0.128	0.123

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
DENIM	0.521	4.77	50	5.85	4.33	0.125	0.095
SWORD	0.521	4.93	237	5.07	6.58	0.131	0.141
PARSLEY	0.519	4.77	118	5.79	4.04	0.131	0.149
COUCH	0.519	4.71	161	6.05	3.31	0.129	0.140
TURTLE	0.519	5	20	6.23	3.62	0.132	0.162
COBRA	0.519	5	35	3.45	7.33	0.132	0.138
RUG	0.519	4.79	321	5.48	3.94	0.126	0.122
PORK	0.518	4.79	150	5.37	5.03	0.128	0.110
DAISY	0.518	5	531	7.01	3.36	0.130	0.131
SERVER	0.518	4.55	12	5.47	5.05	0.126	0.086
BUG	0.517	5	36	3.66	5.82	0.127	0.109
PANTHER	0.517	4.93	50	5.52	6.78	0.129	0.124
TAXI	0.517	4.93	509	5.17	5.33	0.124	0.104
DINER	0.517	4.82	47	6.26	4.72	0.128	0.143
PUPIL	0.516	4.55	617	5.43	4.97	0.123	0.082
GLASS	0.516	4.82	2246	5.74	4.69	0.129	0.136
HAWK	0.516	4.93	67	5.81	5.87	0.125	0.093
CRAB	0.516	4.9	80	5.24	5.15	0.131	0.151
PARTNER	0.516	3.53	459	6.67	4.87	0.120	0.040
HONEY	0.516	4.88	371	6.91	4.1	0.130	0.155
PLIERS	0.516	4.93	26	4.79	5.23	0.127	0.124
BRANDY	0.515	4.81	291	5.54	4.76	0.128	0.140
SANDWICH	0.514	4.9	184	6.82	4.63	0.129	0.157
GRAVEL	0.514	5	185	4.68	4.63	0.125	0.113
TURKEY	0.514	4.89	173	6.05	4.66	0.132	0.153
FLOWER	0.514	5	476	7.17	3.28	0.132	0.165
NOMAD	0.514	4.1	46	5.2	5.14	0.125	0.084
SPARROW	0.513	4.85	47	6.24	4.26	0.128	0.137
PIGEON	0.513	4.71	38	4.9	4.58	0.129	0.148
MARSH	0.513	4.85	213	4.77	4.23	0.125	0.126
CITRUS	0.512	4.21	1	6.56	4.93	0.129	0.109
CELLAR	0.512	4.68	183	4.71	4.73	0.125	0.103
CANYON	0.512	4.81	149	5.98	4.64	0.125	0.110
GYMNAST	0.512	4.85	10	5.99	5.71	0.125	0.081
GRAVE	0.512	4.56	270	2.8	5.99	0.126	0.091

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
CHEST	0.511	4.93	778	5.65	5.03	0.129	0.130
BULLET	0.511	4.83	229	3.31	7	0.127	0.116
OFFICE	0.511	4.93	4460	4.94	4.95	0.123	0.063
SEAL	0.511	4.63	88	6.04	4.41	0.123	0.086
SALT	0.510	4.89	629	5.69	4.77	0.126	0.106
GHETTO	0.510	3.82	205	2.64	6.53	0.125	0.093
SUNSET	0.510	4.54	165	7.64	2.75	0.126	0.094
ICEBERG	0.509	4.96	41	4.73	5.33	0.122	0.066
EAGLE	0.509	5	129	6.45	5.39	0.129	0.122
GRAPE	0.508	5	37	6.58	4.59	0.128	0.133
PLUTO	0.508	–	20	5.83	4.97	0.120	0.074
SALAD	0.508	4.97	281	6.17	4.19	0.131	0.157
MONSTER	0.508	3.72	262	3.1	7.25	0.129	0.115
LAVA	0.508	4.82	63	4.32	6.61	0.127	0.111
TUNNEL	0.507	4.82	237	5.14	5.32	0.126	0.084
PERCH	0.506	4.1	34	5.46	4.39	0.126	0.111
SUSPECT	0.506	2.59	133	3.17	6.63	0.124	0.100
BLACKBOARD	0.505	4.72	96	5.16	4.64	0.125	0.110
STREAM	0.505	4.5	661	6.52	3.25	0.122	0.058
GUARD	0.505	4.04	658	5.17	5.81	0.124	0.070
DIVER	0.505	4.69	16	5.61	5.5	0.125	0.101
BUGGY	0.505	4.18	15	4.9	4.64	0.125	0.113
BODY	0.504	4.79	5243	5.66	5.56	0.126	0.118
PEBBLE	0.504	4.86	85	5.39	4.07	0.130	0.136
SWAMP	0.504	4.96	76	3.95	4.99	0.128	0.112
BEAST	0.504	4.63	298	3.79	6.69	0.129	0.121
TRUCK	0.503	4.84	444	5.54	5.34	0.130	0.129
VALLEY	0.503	4.72	885	6.15	3.59	0.123	0.085
LAPEL	0.503	4.56	50	5.28	4.45	0.126	0.103
TROMBONE	0.503	4.9	11	5.35	5.2	0.124	0.111
SEAGULL	0.503	5	18	5.19	4.34	0.129	0.145
LIVER	0.502	4.68	242	4.16	5.08	0.128	0.115
BANK	0.502	4.78	2333	5.22	5.48	0.125	0.082
FRAGRANCE	0.501	4.03	46	6.29	5.02	0.124	0.090
FOOTBALL	0.501	4.73	573	5.81	5.91	0.122	0.081

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
CHAPEL	0.501	4.6	372	5.59	3.83	0.127	0.103
TABLE	0.501	4.9	3645	5.58	4.21	0.124	0.096
PUDDING	0.501	4.9	198	6.78	4.4	0.130	0.148
TEMPLE	0.500	4.53	416	5.64	4.24	0.127	0.117
TOURIST	0.500	4.59	333	5.4	5.24	0.125	0.105
JUNGLE	0.500	4.66	254	5.48	5.87	0.128	0.113
GOBLIN	0.500	4.38	6	3.63	6.46	0.127	0.134
ATOM	0.499	3.34	187	5.59	5.22	0.123	0.078
GAVEL	0.499	4.88	12	4.67	5.5	0.124	0.087
STAR	0.499	4.69	952	6.93	4.98	0.120	0.067
HERO	0.499	3.07	545	7.22	6.21	0.124	0.093
PLATE	0.498	4.77	656	5.49	4.4	0.127	0.112
COUNTRY	0.498	4.17	6036	6.12	4.37	0.121	0.066
SHORTCAKE	0.498	4.41	2	6.74	4.72	0.131	0.143
HATCHET	0.498	4.93	28	3.88	6.47	0.125	0.096
VICTIM	0.498	3.59	497	2.77	7.11	0.126	0.114
FOREHEAD	0.497	4.9	438	5.19	4.51	0.131	0.139
ROCKET	0.497	4.73	142	5.73	6.56	0.122	0.090
SPACE	0.497	3.54	2259	6.39	5.03	0.120	0.068
MONEY	0.497	4.54	7226	6.96	6.49	0.121	0.076
KIDNEY	0.497	4.96	84	4.82	5.16	0.126	0.111
JELLO	0.496	–	8	6.29	4.16	0.131	0.137
SNAIL	0.495	4.93	46	4.49	3.93	0.131	0.144
KNAPSACK	0.495	4.9	15	5.48	4.4	0.128	0.144
HEART	0.495	4.52	2597	6.6	5.22	0.126	0.092
GARLIC	0.494	4.89	115	5.61	5.08	0.131	0.147
LUGGAGE	0.494	4.83	201	5.32	4.84	0.127	0.116
STOMACH	0.494	4.89	705	4.82	5.09	0.128	0.119
POSSUM	0.494	4.73	20	4.26	4.99	0.130	0.154
CLIMBER	0.494	4.5	94	5.52	5.61	0.125	0.100
LADDER	0.493	5	238	5.31	4.9	0.124	0.087
UMPIRE	0.493	4.27	25	4.91	5.49	0.123	0.080
RELISH	0.492	3.3	99	5.39	4.52	0.123	0.077
XEROX	0.491	3.96	2	4.96	4.66	0.125	0.084
MOP	0.491	4.97	38	4.51	4.6	0.124	0.081

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
BATHTUB	0.491	4.92	31	6.05	3.22	0.131	0.152
MISTER	0.491	3.15	68	5.35	4.72	0.128	0.124
CAMEL	0.490	4.93	146	5.36	4.66	0.126	0.128
DRESSER	0.490	4.96	83	5.48	4.26	0.129	0.171
FLEET	0.490	3.81	240	5.24	5.35	0.119	0.049
COIN	0.490	4.89	132	6.02	5.02	0.126	0.119
TRACTOR	0.490	5	123	5.36	4.95	0.129	0.127
SUNRISE	0.489	4.69	51	7.56	3.5	0.125	0.098
BRIEFCASE	0.489	4.86	144	5.13	4.67	0.126	0.136
STOVE	0.489	4.96	294	5.62	5.03	0.131	0.151
AUTHOR	0.489	4.26	522	5.98	4.51	0.120	0.041
COFFEE	0.489	4.81	1644	6.7	5.61	0.126	0.122
OINTMENT	0.488	4.5	46	4.68	4.77	0.127	0.099
LUNCH	0.488	4.31	1381	6.72	4.62	0.124	0.112
TURNIP	0.488	4.79	49	4.57	4.12	0.129	0.144
JUGGLER	0.488	4.5	6	5.89	5.57	0.127	0.096
NECKLACE	0.487	4.96	41	6.3	5.05	0.127	0.151
TWIG	0.487	4.75	44	5.27	4.17	0.130	0.162
GREASE	0.487	4.61	106	3.88	5.3	0.126	0.118
CAFE	0.487	4.96	342	6.41	4.43	0.129	0.111
ANTLER	0.487	4.86	35	5.22	5.05	0.127	0.135
SPIDER	0.487	4.97	73	3.04	6.95	0.130	0.167
STONE	0.487	4.72	737	5.27	4.37	0.131	0.128
RAFT	0.487	5	57	5.59	4.9	0.122	0.077
DOUGHNUT	0.487	4.96	20	6.86	5.05	0.128	0.150
TILE	0.487	4.68	58	5.41	4.19	0.127	0.131
MAMMAL	0.486	4.59	116	6.16	4.85	0.128	0.140
BRICK	0.486	4.83	498	5.02	4.79	0.125	0.121
WARDROBE	0.486	4.67	160	5.8	4.62	0.128	0.117
FLASK	0.486	4.79	76	5.36	4.84	0.128	0.142
WHISKERS	0.483	4.89	41	5.77	4.43	0.129	0.129
MILDEW	0.483	4.57	34	2.61	5.73	0.128	0.108
SUBWAY	0.483	4.86	37	5.29	5.6	0.124	0.080
CLOTHES	0.482	4.76	2225	6.03	4.55	0.127	0.132
BUREAU	0.482	4.04	289	4.79	4.76	0.122	0.049

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
RUST	0.482	4.52	67	3.37	5.03	0.125	0.099
MAILBOX	0.482	5	24	5.43	4.72	0.125	0.120
PLAZA	0.481	4.44	61	5.66	4.97	0.123	0.085
CUSTARD	0.481	4.85	54	5.75	4.34	0.130	0.156
BOUQUET	0.481	4.74	75	6.55	4.14	0.126	0.112
MOTEL	0.481	4.93	125	5.2	4.86	0.127	0.108
CYCLONE	0.481	4.48	10	3.44	7.34	0.121	0.070
AWARD	0.481	4.14	169	6.92	6.1	0.119	0.029
TWEEZERS	0.481	4.96	29	4.78	5.22	0.128	0.129
ULCER	0.481	4.69	29	2.39	6.59	0.124	0.091
WORLD	0.480	4.36	13345	6.51	5.79	0.121	0.059
SERVANT	0.480	4.64	892	4.05	5.24	0.124	0.102
DASHBOARD	0.480	4.61	37	5.39	4.83	0.126	0.101
EXPERT	0.480	2.85	331	6.06	5.27	0.121	0.045
CUPBOARD	0.480	4.79	270	5.43	4.12	0.127	0.134
CASHEW	0.479	4.92	8	6.29	4.47	0.127	0.096
ICING	0.479	4.66	42	6.57	4.73	0.124	0.076
PIMPLE	0.479	4.77	11	2.74	5.87	0.130	0.129
HANDBAG	0.478	4.93	137	5.64	4.56	0.130	0.152
PEANUT	0.478	4.89	49	5.97	4.63	0.126	0.118
AGENT	0.478	3.61	766	4.9	5.54	0.122	0.056
RADISH	0.477	4.87	11	4.81	4.46	0.132	0.160
SCOTCH	0.477	4.55	153	5.66	4.99	0.127	0.102
PEPPER	0.477	4.59	117	5.59	5.4	0.127	0.116
WORKER	0.477	4.59	3020	5.52	5.44	0.125	0.093
HEADBAND	0.476	5	13	5.38	4.42	0.129	0.124
COCKTAIL	0.476	4.4	147	6.06	5.17	0.126	0.100
RAZOR	0.475	4.9	139	4.48	5.98	0.128	0.121
SCISSORS	0.475	4.85	79	5.12	5.38	0.128	0.129
TERMITE	0.475	4.7	42	2.47	6.36	0.126	0.101
BEAKER	0.475	4.72	21	5.26	4.86	0.126	0.138
TRENCH	0.475	4.46	94	4.16	5.4	0.127	0.109
DIME	0.475	4.85	184	5.75	4.95	0.125	0.094
LODGE	0.475	4	222	5.93	3.78	0.127	0.091
BARREL	0.474	4.86	253	5.12	4.51	0.125	0.106

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
SPOON	0.474	4.96	201	5.64	4.21	0.133	0.181
CRATER	0.474	4.61	85	4.55	5.49	0.125	0.103
PATROL	0.473	3.86	161	4.73	5.98	0.120	0.079
ZIPPER	0.473	4.83	24	5.33	4.99	0.128	0.138
FUNGUS	0.473	4.59	94	2.97	5.91	0.128	0.130
ATLAS	0.473	4.79	40	5.89	4.82	0.123	0.099
CURTAIN	0.472	4.82	346	5.54	4.12	0.124	0.086
CASHIER	0.472	4.89	33	5.47	4.96	0.126	0.115
CREATURE	0.472	4.07	619	4.71	5.95	0.131	0.167
CLOUD	0.472	4.54	536	6.46	2.78	0.122	0.072
TART	0.471	3.27	39	5.59	5.37	0.128	0.123
BOARD	0.471	4.57	1525	5.12	4.37	0.122	0.037
ROBIN	0.471	4.61	207	6.43	4.02	0.129	0.145
DUSTPAN	0.471	5	8	4.45	4.51	0.127	0.128
PUDDLE	0.471	4.67	37	4.78	4.33	0.130	0.143
SHELF	0.471	4.96	246	5.33	4.49	0.124	0.089
PASTRY	0.470	4.97	129	7.03	4.59	0.131	0.152
SOFTBALL	0.470	4.89	8	5.59	5.01	0.125	0.079
YOLK	0.470	4.78	47	5.11	4.37	0.128	0.143
CROSS	0.469	4.44	400	5.26	4.81	0.120	0.057
OUTDOORS	0.469	4.61	68	6.84	4.12	0.125	0.096
SLOPE	0.469	4.07	330	5.14	4.71	0.127	0.106
LEADER	0.468	3.89	1227	5.81	5.78	0.120	0.038
HEDGE	0.468	4.54	178	5.4	4.15	0.124	0.074
COMET	0.467	4.67	26	5.96	6.2	0.126	0.116
CANAL	0.467	4.68	211	5.39	4.49	0.126	0.100
CANDY	0.467	4.83	113	7.14	5.35	0.126	0.131
PROTON	0.467	3	55	5.43	5	0.122	0.063
CONCERT	0.467	4.35	272	6.74	6.36	0.123	0.062
ROOSTER	0.466	4.75	10	5.33	5.4	0.131	0.168
SPOOL	0.465	4.62	56	5.18	4.21	0.127	0.130
WASP	0.465	4.96	43	2.63	6.89	0.130	0.154
PAPER	0.465	4.93	3128	5.75	4.29	0.125	0.098
IDOL	0.464	3.63	86	5.34	5.57	0.125	0.112
BLOCKADE	0.464	4.25	48	3.83	6.21	0.119	0.049

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
GRILL	0.464	4.86	49	6.04	5.24	0.130	0.131
PORTRAIT	0.463	4.9	335	5.97	4.26	0.125	0.106
NOTEBOOK	0.463	4.92	137	5.79	4.42	0.123	0.097
TWISTER	0.462	4.44	8	3.64	6.97	0.126	0.100
PATIENT	0.462	2.5	707	5.03	4.79	0.123	0.086
MACHINE	0.462	4.25	1469	5.36	5.39	0.125	0.113
FREEZER	0.462	4.87	61	5.34	4.82	0.129	0.140
OZONE	0.462	3.5	62	4.89	5.25	0.122	0.066
WRENCH	0.462	4.93	47	4.93	5.1	0.127	0.116
TRIBE	0.462	4.14	416	5.47	5.5	0.126	0.079
GROUND	0.462	4.77	3084	5.32	4.1	0.124	0.082
NEUTRON	0.461	2.69	168	5.41	5.06	0.121	0.053
PECAN	0.461	4.87	7	5.93	4.43	0.128	0.131
POET	0.460	4.36	305	5.95	3.82	0.123	0.098
DONOR	0.460	3.54	48	6.29	5.09	0.122	0.073
FLESH	0.460	4.59	930	4.81	5.72	0.130	0.130
BANJO	0.460	4.9	6	5.91	5.09	0.127	0.127
SADDLE	0.460	4.85	148	5.5	4.68	0.125	0.109
MARKET	0.459	4.7	2374	5.66	5.3	0.119	0.050
PASTA	0.459	4.86	36	6.76	4.61	0.128	0.140
PIANO	0.459	4.9	466	6.62	4.1	0.127	0.129
SCARF	0.459	4.97	142	5.74	4.2	0.130	0.148
WAGON	0.459	4.89	143	5.44	4.67	0.128	0.121
FRECKLE	0.458	4.56	1	5.43	4.59	0.131	0.146
SLIDE	0.458	4.48	153	6.24	4.92	0.122	0.070
SCULPTURE	0.458	4.79	373	6.09	4.31	0.129	0.129
SCARECROW	0.458	4.68	14	4.95	5.34	0.129	0.137
BIKE	0.458	5	149	6.24	5.21	0.128	0.134
COCOON	0.457	4.83	31	5.17	3.84	0.126	0.104
PACKAGE	0.457	4.72	281	5.97	5.32	0.121	0.067
GLOBE	0.457	4.59	184	6.14	4.84	0.119	0.046
OUTFIT	0.457	4.12	182	5.59	4.8	0.124	0.085
BALLOON	0.457	4.92	58	6.43	4.28	0.126	0.115
HORNET	0.455	4.96	10	2.81	7	0.129	0.129
CAPTIVE	0.455	3.03	81	2.9	7	0.120	0.057

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
CRAYON	0.455	4.87	13	6.23	4.27	0.129	0.148
POWDER	0.455	4.76	327	5.39	4.25	0.127	0.116
PISTON	0.454	4.81	56	4.97	5.4	0.125	0.094
BUCKET	0.454	4.96	237	5.1	4.46	0.128	0.137
BROOK	0.454	4.43	74	6.38	3.15	0.125	0.095
KLEENEX	0.454	4.92	42	5.16	4.27	0.130	0.160
BARLEY	0.454	4.59	174	5.34	4.12	0.127	0.099
LUMBER	0.453	4.56	25	5.36	4.61	0.124	0.094
JOURNAL	0.453	4.63	318	5.73	4.37	0.123	0.070
MUFFIN	0.452	4.78	6	6.72	4.24	0.133	0.161
MEDAL	0.452	4.89	99	6.15	5.54	0.122	0.092
GALLON	0.451	3.92	117	5.34	4.62	0.128	0.103
GLOVE	0.451	4.97	82	5.46	4.3	0.128	0.125
VIRUS	0.451	3.48	167	2.39	7.02	0.125	0.073
BACKBONE	0.448	4.19	78	5.47	5.13	0.121	0.050
HELMET	0.448	4.92	166	5.34	5.08	0.129	0.129
SNACK	0.448	4.36	58	6.68	4.95	0.127	0.132
MASK	0.447	4.96	237	4.85	5.43	0.130	0.121
POCKET	0.446	4.68	1003	5.5	4.45	0.128	0.114
PADDING	0.446	4.52	31	5.33	4.08	0.126	0.081
THREAD	0.445	4.83	182	5.43	4.15	0.125	0.083
EMPIRE	0.444	3	381	5.04	6.08	0.122	0.077
LOFT	0.444	4.32	46	5.98	4.08	0.127	0.110
HOOD	0.444	4.88	77	4.97	4.84	0.128	0.126
SUITCASE	0.444	4.97	220	5.36	4.93	0.129	0.151
KEEPER	0.444	3	96	5.64	4.74	0.125	0.095
TUBA	0.444	4.86	5	5.3	5.02	0.126	0.116
SMOG	0.444	4.14	32	2.7	5.7	0.124	0.079
PARROT	0.444	5	45	6.15	5.3	0.127	0.137
BLENDER	0.444	5	22	5.53	5.16	0.127	0.124
STAKE	0.443	4.21	231	4.59	5.41	0.120	0.024
BLUEPRINT	0.443	4.77	53	5.48	4.72	0.121	0.056
GUITAR	0.443	4.9	102	6.59	5.12	0.128	0.124
WITNESS	0.443	4.07	278	5.09	5.74	0.123	0.075
DRILL	0.442	4.4	109	4.74	6	0.123	0.059

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
PERFUME	0.442	4.66	116	6.09	5.3	0.127	0.118
DRAWING	0.442	4.6	485	6.21	4.36	0.118	0.048
YARN	0.442	4.93	24	5.7	3.76	0.125	0.109
CRITIC	0.442	3.55	786	3.8	6.08	0.121	0.049
SHOVEL	0.441	4.97	57	4.72	5.12	0.126	0.130
PRIMATE	0.441	4.5	76	5.34	5.52	0.126	0.101
SCRIBBLE	0.440	4.1	15	5.07	4.86	0.125	0.095
DRINK	0.440	4.76	1414	6.5	4.73	0.129	0.113
FRAME	0.440	4.3	442	5.3	4.56	0.123	0.092
EYELASH	0.439	5	52	5.65	4.61	0.124	0.079
REBEL	0.439	3.07	74	4.5	6.72	0.124	0.067
STATUE	0.439	4.93	273	5.7	4.43	0.128	0.137
BADGE	0.439	4.93	117	5.16	5.53	0.123	0.105
MARROW	0.439	4.48	50	4.72	5.14	0.124	0.113
PASSAGE	0.438	3.8	655	5.58	4.57	0.118	0.043
GIFT	0.438	4.56	560	7.41	5.72	0.126	0.105
LASER	0.438	4.5	31	5.61	6.17	0.122	0.076
QUARTER	0.436	4.43	911	6	4.93	0.118	0.035
RATTLE	0.436	4.07	47	4.67	5.82	0.126	0.075
ROBOT	0.435	4.65	67	5.74	5.7	0.129	0.117
FLUTE	0.435	5	44	6.15	4.27	0.128	0.136
HOOK	0.434	4.79	487	4.54	5.49	0.125	0.090
CIGAR	0.434	4.93	882	3.81	5.06	0.130	0.135
MESSAGE	0.434	3.97	1214	5.67	5.21	0.121	0.050
CATCHER	0.434	4.44	21	5.39	5.02	0.126	0.092
PADDLE	0.434	4.8	11	5.05	5.05	0.128	0.129
BUBBLE	0.434	4.6	67	6.41	3.92	0.123	0.077
NOVEL	0.433	4.21	495	6.45	4.39	0.121	0.058
COSTUME	0.432	4.57	367	5.87	5.28	0.129	0.133
DANDRUFF	0.432	4.79	6	2.81	5.47	0.127	0.103
HAMPER	0.432	4.21	29	4.78	4.28	0.121	0.049
FLASHLIGHT	0.432	5	77	5.71	4.82	0.128	0.138
CORAL	0.431	4.4	39	6.21	4.04	0.128	0.121
BRACES	0.430	5	33	3.97	5.15	0.124	0.087
CRICKET	0.429	4.77	520	4.77	4.44	0.125	0.077

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
FIDDLE	0.429	4.81	28	5.87	4.85	0.123	0.085
SUPPER	0.429	4.63	480	6.61	4.67	0.127	0.127
PLASTER	0.427	4.59	209	4.82	4.56	0.128	0.130
GARBAGE	0.426	4.69	166	2.74	5.52	0.125	0.109
POSTAGE	0.426	4.37	32	4.91	4.53	0.121	0.067
DRUG	0.426	4.48	344	3.69	6.2	0.120	0.068
SHRUB	0.425	4.92	66	5.47	4.08	0.129	0.142
FLIPPER	0.424	4.26	11	5.53	4.81	0.128	0.127
FILM	0.424	4.71	1546	6.33	5.14	0.125	0.087
PICTURE	0.424	4.52	1905	6.26	4.54	0.125	0.093
BRANCH	0.422	4.9	961	5.57	4.37	0.122	0.068
MAPLE	0.422	4.46	56	6.43	4.03	0.127	0.134
SHELTER	0.422	4.64	358	6.05	4.04	0.122	0.082
TOASTER	0.421	4.9	11	5.75	4.63	0.130	0.151
COURSE	0.420	3.82	11914	5.31	4.84	0.120	0.063
CRUTCH	0.420	4.5	50	3.96	5.43	0.127	0.098
RECEIPT	0.420	4.86	122	5.06	4.84	0.121	0.064
PEDAL	0.419	4.44	22	5.39	4.82	0.125	0.097
REPORT	0.419	3.92	1860	4.97	5.11	0.119	0.030
BINDER	0.418	4.89	16	5.01	4.51	0.124	0.122
COUNTY	0.418	4.04	785	5.48	4.46	0.122	0.068
POPCORN	0.416	5	14	6.67	4.75	0.129	0.136
MARKER	0.415	4.62	37	5.44	4.63	0.125	0.092
SODA	0.415	4.97	148	6.21	5.1	0.127	0.125
TREAT	0.414	3.79	145	7.06	5.39	0.125	0.065
NAPKIN	0.413	4.93	84	5.55	4.14	0.127	0.145
WASHER	0.412	4.7	30	5.54	4.77	0.128	0.130
CONTRACT	0.411	4.15	538	4.77	5.38	0.121	0.028
CLAMP	0.410	4.53	19	4.51	5.23	0.124	0.062
SKETCH	0.410	4.56	122	5.81	4.53	0.125	0.107
PIGMENT	0.408	4.4	31	5.38	4.72	0.125	0.098
VALVE	0.408	4.83	87	5.21	4.7	0.126	0.093
FLAG	0.408	4.79	168	5.94	5.25	0.124	0.097
MOTOR	0.406	4.84	465	5.33	5.41	0.125	0.082
MATCH	0.403	4.14	775	5.3	5.23	0.121	0.066

Word	Recall Probability	Concreteness	Frequency	Valence	Arousal	M List	M Pool
HARP	0.401	4.85	43	6.32	3.55	0.127	0.107
STAMP	0.399	4.7	183	5.33	4.53	0.124	0.083
ITEM	0.398	4.41	350	5.15	4.58	0.126	0.101
SPONGE	0.396	5	107	5.19	4.13	0.129	0.156
TABLET	0.389	4.82	48	5.89	4.71	0.124	0.097
STAPLE	0.389	4.34	73	5.07	4.8	0.122	0.067
COMPASS	0.387	4.66	82	5.62	4.8	0.126	0.103
LAMP	0.387	4.97	381	5.65	4.2	0.128	0.130
PERMIT	0.382	3.43	70	5.09	4.95	0.122	0.053
RACKET	0.381	4.26	139	4.62	5.67	0.125	0.104
BRAKE	0.375	4.44	65	5.04	5.3	0.121	0.095
SWITCH	0.372	4.07	315	5.05	4.98	0.121	0.047
STICKER	0.367	4.67	35	5.6	4.61	0.127	0.106
WHISTLE	0.366	4.42	140	5.22	5.73	0.125	0.087
CLOCK	0.364	5	637	5.52	4.6	0.123	0.075
LABEL	0.364	4.46	223	5.09	4.46	0.123	0.065
RECORD	0.359	4.15	1077	5.68	4.85	0.119	0.036
SURVEY	0.351	4.08	543	5.52	5.02	0.120	0.033

Table 2

Correlation Matrix of Predictors of Interest for Word and List Recall Models

Word-Level Recall Predictors

	Concreteness	Frequency	Length	Valence	Arousal	M List
Frequency	-0.05	*	*	*	*	*
Length	-0.03	-0.12	*	*	*	*
Valence	0.14	0.17	-0.02	*	*	*
Arousal	-0.24	0.00	0.04	-0.55	*	*
M List	0.14	-0.05	-0.01	0.02	-0.05	*
M Pool	0.47	-0.15	0.01	0.08	-0.19	0.30

List-Level Recall Predictors

	Concreteness	Frequency	Length	Valence	Arousal	M List
Frequency	-0.07	*	*	*	*	*
Length	-0.04	-0.12	*	*	*	*
Valence	0.13	0.19	-0.02	*	*	*
Arousal	-0.27	0.01	0.04	-0.56	*	*
M List	0.23	-0.11	0.00	0.04	-0.10	*
M Pool	0.20	-0.12	0.02	0.05	-0.10	0.61

The predictors appear to be weakly correlated except for Concreteness & Pool Meaningfulness in Word Recall Model, List Meaningfulness & Pool Meaningfulness in List Recall Model, and Valence & Arousal in both models.

Table 3

Regression Analysis for Variables Predicting Probability of Word Recall

	Full Model		*Model w/out Conc.*		*Model w/out M Pool*		*Model w/out Valence*		*Model w/out Arousal*	
	M β	*SE* β	*M* β	*SE* β	*M* β	*SE* β	*M* β	*SE* β	*M* β	*SE* β
Concreteness	-0.004	0.006	—	—	0.048***	0.006	-0.004	0.006	-0.023**	0.007
Word Frequency	0.061***	0.006	0.061***	0.006	0.047***	0.006	0.081***	0.007	0.072***	0.007
Word Length	-0.006	0.005	-0.006	0.005	-0.004	0.005	-0.004	0.005	-0.002	0.005
Valence	0.108***	0.009	0.108***	0.009	0.111***	0.009	—	—	0.039***	0.007
Arousal	0.13***	0.009	0.131***	0.009	0.124***	0.009	0.071***	0.007	—	—
List Meaningfulness	0.01*	0.005	0.011*	0.005	0.042***	0.005	0.01*	0.005	0.01*	0.005
Pool Meaningfulness	0.132***	0.008	0.129***	0.007	—	—	0.133***	0.008	0.126***	0.008

Note. Multiple Regression Analyses were conducted word-level. Each subject's data for all variables were regressed against probability of word recall to create a significant prediction model. $^{*}p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (all p's are FDR corrected).

Table 4

Regression Analysis for Variables Predicting Probability of List Recall

	Full Model		*Model w/out M Pool*		*Model w/out M List*		*Model w/out Valence*		*Model w/out Arousal*	
	M β	*SE* β	*M* β	*SE* β	*M* β	*SE* β	*M* β	*SE* β	*M* β	*SE* β
Concreteness	-0.008	0.005	-0.009	0.005	-0.001	0.005	-0.008	0.005	-0.008	0.005
Word Frequency	0.018***	0.005	0.019***	0.005	0.016**	0.005	0.021***	0.005	0.018***	0.005
Word Length	-0.005	0.004	-0.005	0.004	-0.005	0.004	-0.005	0.004	-0.005	0.004
Valence	0.019**	0.005	0.019***	0.005	0.019**	0.005	—	—	0.018***	0.004
Arousal	0.004	0.006	0.005	0.006	0.003	0.006	-0.007	0.005	—	—
List Meaningfulness	0.051***	0.005	0.04***	0.004	—	—	0.051***	0.005	0.051***	0.005
Pool Meaningfulness	-0.02**	0.006	—	—	0.008	0.005	-0.019**	0.006	-0.02***	0.006
Trial Number	-0.187***	0.012	-0.187***	0.012	-0.187***	0.012	-0.187***	0.012	-0.187***	0.012
Session Number	-0.027	0.026	-0.027	0.026	-0.027	0.026	-0.027	0.026	-0.027	0.026

Note. Multiple Regression Analyses were conducted list-level. Each subject's data for all variables were regressed against probability of recall to create a significant prediction model. $^{*}p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (all p's are FDR corrected).

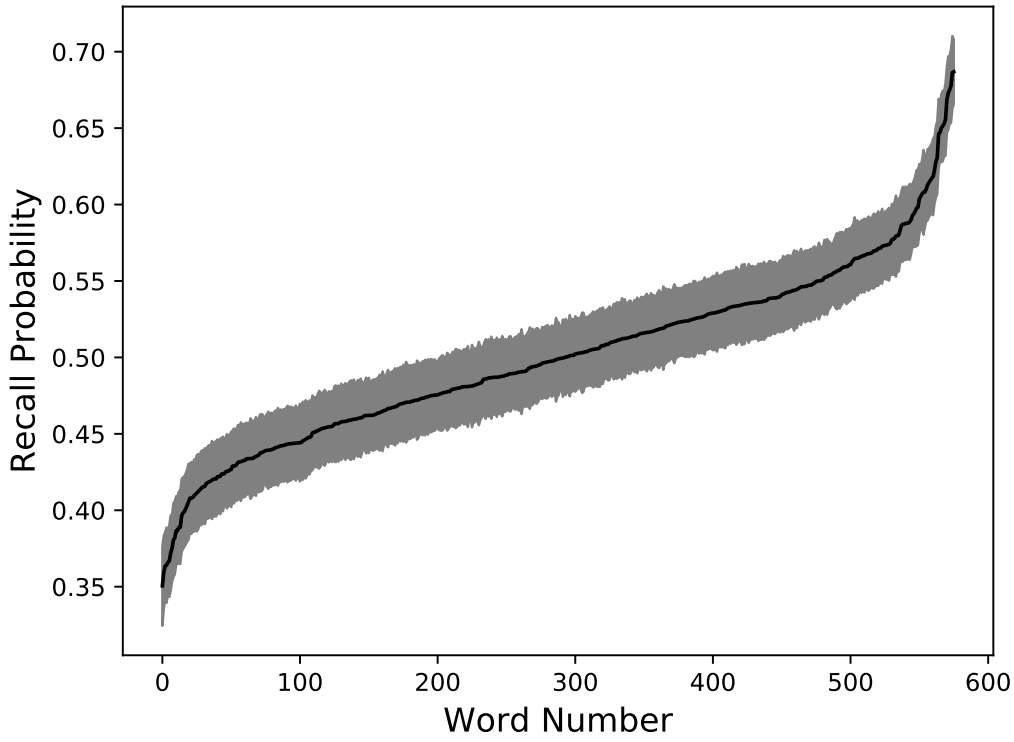

Figure 1. **Variability in free recall of words** Average recall probability of each word. Average recall probabilities are sorted from lowest to highest and plotted along with the standard error confidence band around the average values.

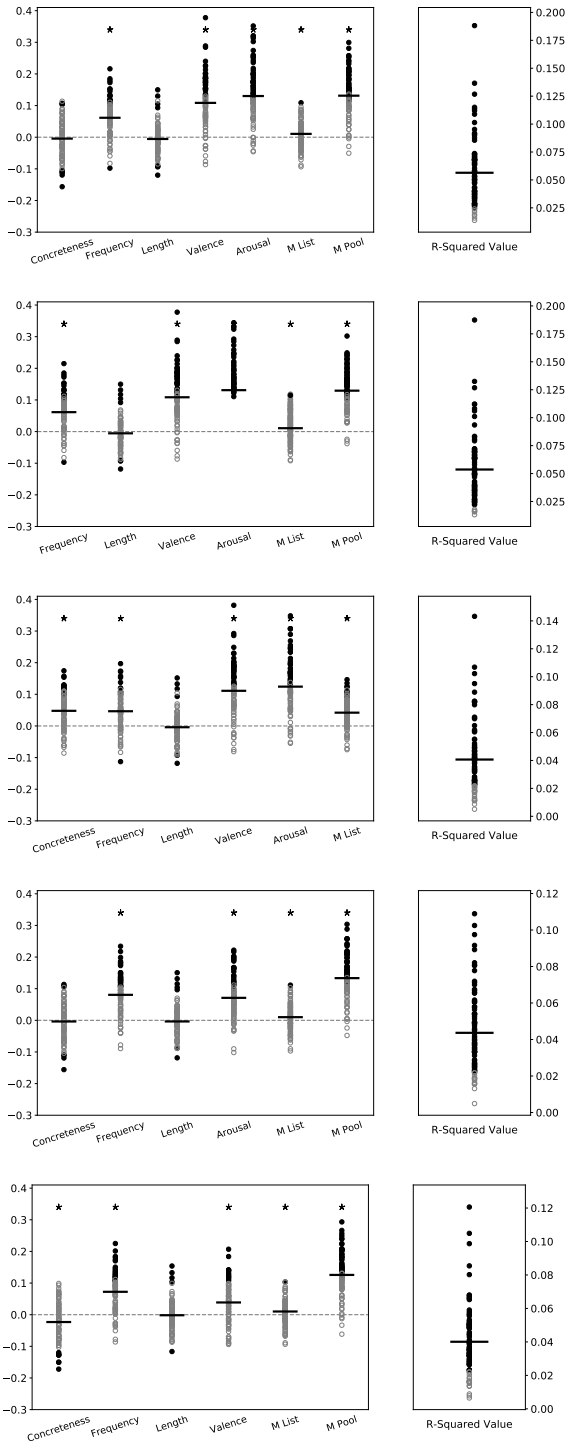

Figure 2. **Distributions of β values for each predictor variable in the word recall models and model R^2 values.** Each circle on the left plots denotes the normalized regression coefficient for a single subject, with filled circles indicating coefficients that met an FDR- correct $p < 0.05$ significance criterion. Each dot on the right plots denote the R^2 value for a given subject's model. Models that met our significance threshold ($p < 0.05$) are shaded in black.

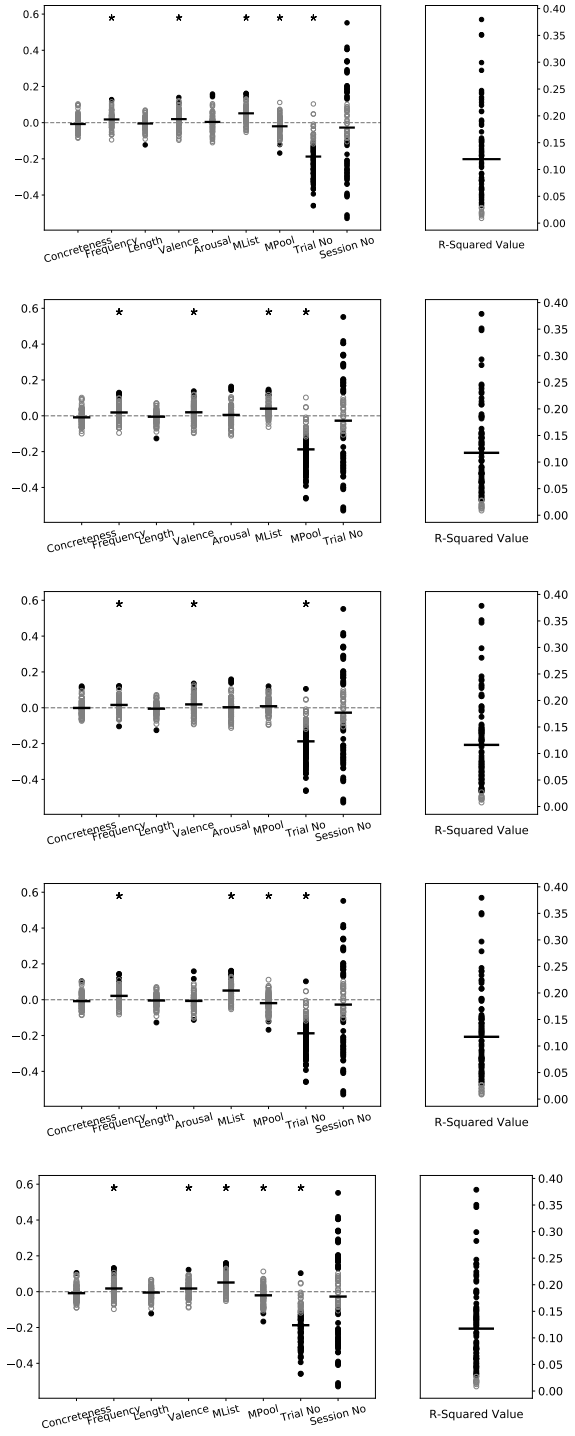

Figure 3. **Distributions of β values for each predictor variable in the list recall models and model R^2 values.** Each circle on the left plots denotes the normalized regression coefficient for a single subject, with filled circles indicating coefficients that met an FDR-correct $p < 0.05$ significance criterion. Each dot on the right plots denote the R^2 value for a given subject's model. Models that met our significance threshold ($p < 0.05$) are shaded in black.