Journal of Experimental Psychology: Learning, Memory, and Cognition

Semantic Processes in Preferential Decision Making

Sudeep Bhatia

Online First Publication, July 19, 2018. http://dx.doi.org/10.1037/xlm0000618

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/18/$12.00

http://dx.doi.org/10.1037/xlm0000618

Semantic Processes in Preferential Decision Making

Sudeep Bhatia
University of Pennsylvania

This article examines how semantic memory processes influence the items that are considered by decision makers in memory-based preferential choice. Experiments 1A through 1C ask participants to list the choice items that come to their minds while deliberating in a variety of everyday choice settings. These experiments use semantic space models to quantify the semantic relatedness between pairs of retrieved items and find that choice item retrieval displays robust semantic clustering effects, with retrieved items increasing the retrieval probabilities of related items. Semantic clustering can be disassociated from the effect of item desirability and can lead to inefficiencies such as the consideration and evaluation of undesirable items early on in the decision. Experiments 2A through 2C use a similar approach to study the effects of contextual cues on item retrieval and find that decision makers are biased toward retrieving choice items that are semantically related to the choice context. This effect is usually strongest early on in deliberation and weakens as additional items are retrieved. Overall, the results highlight the role of semantic memory processes in guiding the generation of memory-based choice sets, and illustrate the value of semantic space models for studying preferential decision making.

Keywords: decision making, semantic memory, semantic clustering, preregistration, open science

Preferential decision making involves the selection of a favored item from a set of feasible choice items. Most experiments on preferential decision processes explicitly present a choice set to participants. Correspondingly, psychological theories of decision making have been concerned primarily with how decision makers choose between such an exogenously determined set of items, that is, the decision rules they use to evaluate these items, as well as the effects of the composition of the choice set, and other related contextual factors, on their choices (Busemeyer & Rieskamp, 2014; Oppenheimer & Kelso, 2015).

However, many common decision scenarios do not involve a fixed, exogenous set of choice items. Rather, decision makers must construct such choice sets by themselves, typically through the use of memory processes (see Alba & Hutchinson, 1987; Lynch & Srull, 1982 for early discussions). Consider, for example, the task of planning what to eat, buying a gift for a friend or family member, or deciding on a vacation destination. In such settings, the set of items that decision makers choose between is determined by the set of items that comes to their mind. Decision makers may be exposed to various external information sources, but in the absence of such information, the items that do come to mind, and thus form the choice set, are the items that are successfully retrieved from memory.

The key role of memory in generating choice sets in common choice tasks raises a number of important questions at the intersection of memory and decision making research. From a theoretical perspective: What are the mechanisms that determine the items that are retrieved by decision makers when exogenous choice sets are not provided? How do these mechanisms relate to core memory processes known to play a role in nonpreferential choice tasks, and do these memory processes facilitate or hinder efficient memory retrieval for decision making? Practically, can the mechanisms at play in memory-based decision making be tested? The set of retrieved choice items in everyday decision making tasks is completely unconstrained—any choice item can come to mind, and the items that do come to mind often lack a clear category structure. So how can the relationship between the various retrieved items, and between these items and other relevant variables (such as choice context), be quantified?

In this article, I attempt to address these questions using existing insights on human memory. The task of retrieving a feasible set of choice items from memory has similarities to well-studied memory tasks such as free recall and free association. Thus it is likely that both the mechanisms that guide retrieval in these tasks, as well as the effects generated by these mechanisms, carry over to the domain of preferential decision making. For example, as with free recall and free association, the generation of memory-based choice sets may involve associative activation processes (Anderson, Bothell, Lebiere, & Matessa, 1998; Atkinson & Shiffrin, 1968; Hintzman, 1984; Polyn, Norman, & Kahana, 2009). This would cause memory-based choice sets to display semantic clustering, with retrieved items increasing the retrieval probability of other semantically related items (Bousfield & Sedgewick, 1944; Gruenewald & Lockhead, 1980; Howard & Kahana, 2002; Romney, Brewer, & Batchelder, 1993). For this reason, one would also expect retrieved items to depend on contextual cues, such as choice context, with items that are semantically related to these cues being more likely to be retrieved (Hare, Jones, Thomson, Kelly, & McRae, 2009;

Funding for Sudeep Bhatia was received from the National Science Foundation grant SES-1626825.

Correspondence concerning this article should be addressed to Sudeep Bhatia, Department of Psychology, University of Pennsylvania, D22 Solomon Labs, 3720 Walnut Street, Philadelphia, PA 19104. E-mail: bhatiasu@sas.upenn.edu

Moss, Ostrin, Tyler, & Marslen-Wilson, 1995; Nelson, McEvoy, & Schreiber, 2004).

Moreover, it may also be possible to apply the methodological tools developed to study memory processes in free recall and free association tasks to the domain of preferential choice. For example, memory researchers have previously used semantic space models (e.g., Jones & Mewhort, 2007; Landauer & Dumais, 1997) to evaluate the semantic relatedness between items, and between items and contextual cues (e.g., Hills, Jones, & Todd, 2012; Howard & Kahana, 2002). The power and generalizability of these models implies that they could also be used to describe the semantic relationships at play when decision makers are asked to generate choice sets from memory.

Of course, it may also be possible that the types of memory processes outlined in the preceding text play only a small role, or no role whatsoever, in preferential choice. After all, making preferential choices requires the selection of items based on their desirability, and decision makers may be adept at retrieving only the most desirable items from their memories, ignoring semantic structure. It may even be possible that desirability-based retrieval leads to the appearance of semantic clustering, implying that the effect of desirability must be carefully controlled for when analyzing retrieval dynamics.

In either case, the study of semantic memory processes in memory-based choice is necessary for building richer, more comprehensive, and more generalizable theories of preferential choice. It can also facilitate the deeper integration of research on memory and research on decision making, two important but separate areas of investigation in psychology. Finally, the study of memory-based choice can be used to determine whether or not memory processes lead to efficient decision making, and in turn, can generate practical insights about how to improve the choices of individuals.

The goal of this article is to perform such a study. For this purpose, it presents the results of six preregistered experiments investigating the effect of semantic structure when decision makers are asked to deliberate about preferential choice items in the absence of exogenously provided choice sets. Experiments 1A through 1C test for semantic clustering in such memory-based choice sets. These experiments also examine the relationship between retrieval dynamics and choice item desirability. Experiments 2A through 2C expand on this work by considering the impact of choice context on the generation of memory-based choice sets. Both sets of experiments use semantic space models to quantify the semantic relatedness between pairs of items and between items and the choice context, thereby allowing for a rigorous analysis of the effects of semantic structure in an otherwise unconstrained choice item domain.

Retrieving Items From Memory

The study of how individuals retrieve items from memory has a long history in psychological research. Memory retrieval has been examined using a wide variety of different tasks, including free recall from lists, free recall of natural categories, and free association. In the first task individuals are presented with a list of items (usually common words) and are asked to recall all the items on the list that they are able to. In the second task, individuals are not given a list, but are rather ask to list all exemplars of a given category (e.g., “animals”) that they can think of. In the third task, individuals are merely asked to list all the items that come to their mind as they are presented with a specific prompt (e.g., “dog”; see Kahana, 2012 for a comprehensive review).

These three tasks shed light on the organization of memory, and suggest a major role for semantic memory processes in item retrieval. For example, free recall and free association tasks are vulnerable to semantic clustering effects according to which items that are similar to each other are retrieved alongside each other (Bousfield & Sedgewick, 1944; Gruenewald & Lockhead, 1980; Howard & Kahana, 2002; Howard, Jing, Addis, & Kahana, 2007; Romney et al., 1993; Wixted & Rohrer, 1994). Additionally, varying contextual cues, such as prompts in free association tasks, can bias retrieval in favor of items that are semantically similar to the prompt (Hare et al., 2009; Moss et al., 1995; Nelson, McEvoy, & Dennis, 2000; Roediger, Watson, McDermott, & Gallo, 2001). These effects are often modeled using theories of associative memory (Anderson et al., 1998; Atkinson & Shiffrin, 1968; Hintzman, 1984; Polyn et al., 2009). According to these theories memory is probed with cues that activate closely associated items. As items are retrieved, they themselves cue subsequent retrieval, leading to semantic clustering and other complex dynamics.

The types of tasks used in the aforementioned work closely mimic common decision scenarios in which decision makers have to choose between items in the absence of exogenous choice sets (Alba & Hutchinson, 1987; Lynch & Srull, 1982). If similar memory mechanisms are at play in both settings, then many of the semantic effects observed in free recall and free association should carry over to the preferential choice domain. To test for these effects, I consider memory-based decisions in which individuals are asked to list the items that come to their mind while making choices in various everyday scenarios, such as deciding what to eat, what to buy as a gift, and where to go for a vacation. I expect to observe both semantic clustering, with items that are semantically related to the previously retrieved item being more likely to be retrieved, and context dependence, with items that are semantically related to the choice context being most likely to be retrieved.

Semantic Space Models

Memory-based choice tasks are relatively unconstrained, and decision makers can list almost any item (or combination of items) that they think of. The unconstrained nature of these tasks imposes the challenge of quantifying semantic relatedness: How can one measure semantic clustering or the effect of choice context if the items generated in the task are not predetermined by the experimenter?

Fortunately, researchers have already proposed a solution to this problem: semantic space models (Griffiths, Steyvers, & Tenenbaum, 2007; Jones & Mewhort, 2007; Kwantes, 2005; Landauer & Dumais, 1997; Lund & Burgess, 1996; Mikolov, Sutskever, Chen, Corrado, & Dean, 2013). Such models use word co-occurrence statistics in large natural language data sets to derive high-dimensional semantic vectors for a large set of words. Semantically related words have vectors that are close to each other, so that semantic relatedness between nearly any pair of words can be quantified by vector distance between their corresponding vectors.

Semantic space models have been shown to predict a wide range of psychological phenomena involving similarity judgment, categorization, text comprehension, and semantic priming (Griffiths et al., 2007; Jones & Mewhort, 2007; Landauer & Dumais, 1997; also see Bullinaria & Levy, 2007 or Jones, Willits, Dennis, & Jones, 2015 for a review). Critically, they have also been shown to be successful at specifying semantic relatedness, and, for this reason, they can predict the set of listed items (as well as the order in which these items are listed) in completely unconstrained free recall and free association tasks (e.g., Hills et al., 2012; Hills, Todd, & Jones, 2015). Finally, recent work has found that semantic space models are also able to describe high-level cognitive phenomena related to decision making such as factual judgment, probability judgment, forecasting, and social judgment (Bhatia, 2017a, 2017b), suggesting that such models may also be applicable to memory-based preferential choice.

In this article, I apply pretrained semantic vectors obtained though the Word2Vec methods of Mikolov et al. (2013). These vectors were generated by applying a recurrent neural network on a corpus of Google News articles with over 100 billion words. They have a vocabulary of 300 million words and phrases, with each word or phrase being defined on 300 dimensions. I specify semantic relatedness using cosine similarity. Thus for any pair of items x and y listed in the experiments, I measure the semantic relatedness of x and y by calculating $S(x, y) = x \cdot y/(\|x\| \cdot \|y\|)$, where x and y are the Word2Vec vectors corresponding to items x and y. $S(x, y)$ ranges from $+1$ for vectors that are in identical directions, to -1 for vectors that are in opposite directions ($S[x, y] = 0$ corresponds to orthogonal vectors). Note that, occasionally, I had to compute semantic relatedness when decision makers mention items composed of multiple words. In these settings, I merely average the vectors of the component words to obtain a single composite vector representation for the item (which is passed through the preceding cosine similarity function to compute semantic relatedness). I also use this approach to measure the semantic relatedness between any item and a given choice context (such as *dinner* or *breakfast* in a task asking decision maker to consider what to eat). Thus for an item x and a context cue c, I calculate the semantic relatedness of x and c using $S(x, c)$.

Desirability

The setting I am examining does have one important feature not usually at play in common memory tasks: the effect of item desirability. Decision makers need to retrieve items in order to find items to choose. Thus efficient decision making, which involves the selection of the most desirable item as quickly as possible, requires the retrieval of the most desirable items first. Some prior work on memory processes suggests that item desirability may play a role in item retrieval. For example, emotionality has been shown to influence retrieval probability in free-recall tasks, with strongly valenced or arousing items being more likely to be retrieved earlier in the task (e.g., Doerksen & Shimamura, 2001; Rubin & Friendly, 1986; but see Talmi & Moscovitch, 2004). However, it is not clear whether desirability has a similar effect in the decision making context, and whether this effect is so strong so as to wash out the semantic components of retrieval.

In fact, the potential effect of desirability on retrieval can lead to incorrect conclusions about underlying semantic processes, if people have similar preferences for similar items. Consider, for example, a decision maker who likes pasta dishes more than salad dishes. For this decision maker, the best possible meal choice is pasta Bolognese, followed by pasta Alfredo, and then Mediterranean salad and Caesar salad. If this decision maker retrieves food items purely in order of their desirability, one would still expect to observe some degree of semantic clustering; that is, similar items (which are similarity desirable) will be retrieved successively. However, this type of clustering would emerge even if semantic processes were not directly at play in the task.

Controlling for the effects of desirability on item retrieval, the present experiments involve evaluations of desirability for all items that were considered by the decision maker. I do expect that—all else equal—items rated as being more desirable will be retrieved first, but my primary analysis involves the additional effect of semantic relatedness. Thus, I test for whether semantic clustering and related effects emerge even when the effect of desirability is controlled for.

Note that the emergence of a semantic clustering effect, in addition to a desirability effect, would imply a type of inefficiency in the decision process. If decision makers cluster items by semantic proximity, they may occasionally retrieve undesirable items that are semantically related to previously retrieved items before retrieving highly desirable but semantically unrelated items. This can increase the time necessary to make the optimal choice (or conversely increase the likelihood of making a suboptimal choice).

Memory and Decision Making

There has already been some research examining free recall in a marketing context. This work is concerned primarily with the recall of brands, and the effects of various marketing-related variables (e.g., order of entry into the market or exposure to advertisements) on the ability of consumers to successfully retrieve brands (Hutchinson, Raman, & Mantrala, 1994; Nedungadi, 1990; Shapiro, MacInnis, & Heckler, 1997; see also Roberts & Lattin, 1997 for a review). Some of this research suggests that brand recall dynamics display semantic clustering, with brands within the same product category being more likely to be retrieved together (Hutchinson, 1983; Lattin & Roberts, 1992). The setting that I am concerned with is more general than brand recall (participants can list any items that come to their mind as they deliberate about an everyday decision), and additionally examines the effect of contextual cues on item listing. For this reason, this analysis of participant behavior also requires the use of semantic space models to specify semantic relatedness, a technique that has not been used in the study of brand recall. However, despite these theoretical and methodological differences, the existing research does suggest that the semantic processes at play in free recall settings may also extend to preferential decision making and that one could expect to observe effects such as semantic clustering in these experiments.

Another set of research related to these tests involves theoretical models of decision making based on memory processes. There are a number of such models that allow for contextual cues and related probes to influence the retrieval of decision-relevant information (Dougherty, Gettys, & Ogden, 1999; Glöckner, Hilbig, & Jekel, 2014; Johnson, Häubl, & Keinan, 2007; Marewski & Schooler, 2011). However, the model that is most relevant to these hypotheses is the associative accumulation model (Bhatia, 2013), which allows choice items and contextual cues to bias attribute activation and in turn bias the activation of other choice items, in a manner that is very similar to standard associative memory models (such as, e.g., Polyn et al., 2009). For this reason, the Associative Accumulation Model can also be seen to predict the emergence of semantic clustering and context dependence in decisions from memory.

Experiments 1A, 1B, and 1C: Semantic Clustering

Experiments 1A, 1B, and 1C test for semantic clustering effects in settings in which decision makers have to retrieve choice items from memory. These effects involve the influence of retrieved items on successive retrieval and predict that semantically related items should be retrieved one after the other. All three of these experiments involve identical procedures and vary only in terms of the decision domain, and thus I present their methods and results simultaneously. Note that all three experiments have been preregistered and approved by the University of Pennsylvania Institutional Review Board (IRB).[1]

Method

Participants. Participants (N = 50; mean age = 29; 42% female) in Experiment 1A, participants in Experiment 1B (N = 50; mean age = 30; 30% female), and participants in Experiment 1C (N = 50; mean age = 30; 44% female), recruited from Prolific Academic, performed the experiment online. All participants were residents of the United States. They were compensated at a rate of approximately $US7.50/hr.

Procedures. In all three experiments participants were shown a description of a decision setting and were asked to list 20 items that came to their mind as they considered making their decisions. The instructions used were as follows:

> Experiment 1A (foods): Imagine that you could eat anything that you wanted for a meal tomorrow. In the boxes below please list 20 food items that come to your mind as you consider what to eat. Please make sure to list all foods that you think of, regardless of whether you would eventually want to eat them. Please list these foods in the order that they come to your mind (i.e., the first food that comes to your mind listed first, the second listed second, etc.)

> Experiment 1B (vacations): Imagine that you could go on any vacation that you wanted. In the boxes below please list 20 vacation destinations that come to your mind as you consider where to go. Please make sure to list all destinations that you think of, regardless of whether you would eventually want to go the destination. Please list these vacation destinations in the order that they come to your mind (i.e., the first destination that comes to your mind listed first, the second listed second, etc.)

> Experiment 1C (gifts): Imagine that you have to purchase a gift for a close friend or family member. In the boxes below please list 20 items that come to your mind as you consider what gift to purchase. Please make sure to list all items that you think of, regardless of whether you would eventually want to purchase them. Please list these items in the order that they come to your mind (i.e., the first item that comes to your mind listed first, the second listed second, etc.)

Participants listed all the items on a single screen, in successive free entry boxes. After these items were listed participants were taken to a second screen on which they rated each of their 20 listed items in terms of desirability, on a scale from −3 to + 3 (−3 corresponding to extremely undesirable, 0 to neither desirable nor undesirable, and + 3 to extremely desirable). Thus participants in Experiment 1A rated their 20 items based on how much they would like to eat each of these items, participants in Experiment 1B rated their items based on how much they would like to go there for a vacation, and participants in Experiment 1C rated their items based on how much they would like to give the item as a gift.

Results

Conditional response probabilities. To test whether the items listed by participants were clustered semantically, I first analyzed the data using the approach suggested by Howard and Kahana (2002). This approach tests the relationship between the retrieval order and the similarity of the items retrieved, as measured by latent semantic analysis (LSA; Landauer & Dumais, 1997). More specifically it defines a measure: the LSA conditional response probability (LSA-CRP), which specifies the probability of retrieving one item given the previously retrieved item as a function of the LSA-based cosine similarities of the two items. This involves first dividing the pairwise similarities between all the items into a small number of equally sized bins, based on the magnitude of similarity (with the first bin corresponding to the smallest pairwise cosine similarity values, and the last bin corresponding to the largest pairwise cosine similarity values). Once these bins have been generated, LSA-CRP computes the relative frequencies of the similarities of successively retrieved items falling into each of these bins (ensuring that the measure only considers items that are available to be retrieved at each point in time). If there are no systematic semantic clustering effects, one would expect the conditional response probabilities for each bin to be roughly identical; that is, items that are semantically related to the previously retrieved item (i.e., items whose similarities with the retrieved item are in bins with large cosine similarity values) should have a similar probability of being retrieved as items that are semantically unrelated to the previously retrieved item (i.e., items whose similarities with the retrieved item are in bins with small cosine similarity values). In contrast, in the presence of semantic clustering, the conditional response probabilities should be highest for items whose similarities fall into the highest bins. Howard and Kahana did find that CRPs correlate positively with bin magnitudes, rigorously demonstrating semantic clustering effects in free recall.

The analysis I performed in Experiments 1A through 1C was identical to that performed by Howard and Kahana (2002), except for one minor difference. Howard and Kahana applied their approach to free recall from lists, and used the pairwise similarities between every word in their word pool to determine the pairwise cosine similarity bins. In contrast, these experiments allowed decision makers to list any items. Thus, I was unable to precompute pairwise similarity bins. Instead, for each participant, I took the 20 items listed by that participant and computed the pairwise similarities between these items, generating 190 distinct pairwise similarities. I then used these 190 pairwise similarities to generate 10 separate bins, with the first bin containing the smallest pairwise similarity values and the last bin containing the largest pairwise

[1] Preregistration materials are available at: https://osf.io/zudp7/

similarity values (see Hills et al., 2012 and 2015 for a similar approach to studying semantic clustering in unconstrained tasks).

Overall, I found a strong positive relationship between CRP and the Word2Vec-based similarity as measured by the approach outlined in the preceding text. This is shown in Figures 1A through 1C which plot the average CRPs across participants in each experiment, for each of the 10 bins. As can be seen in these figures, the average CRPs for the highest bins (corresponding to the largest pairwise similarity values) were substantially higher than the average CRPs for the remaining bins, for all three experiments. This indicates that after retrieving an item, decision makers were especially likely to retrieve other semantically related items (items whose pairwise similarities with the retrieved item are in the highest bin). Overall, I observed a correlation of 0.85 between bin number (1, 2,. . . 10) and the average CRP for Experiment 1A (foods), a correlation of 0.69 for Experiment 1B (vacations), and a correlation of 0.73 for Experiment 1C (gifts).

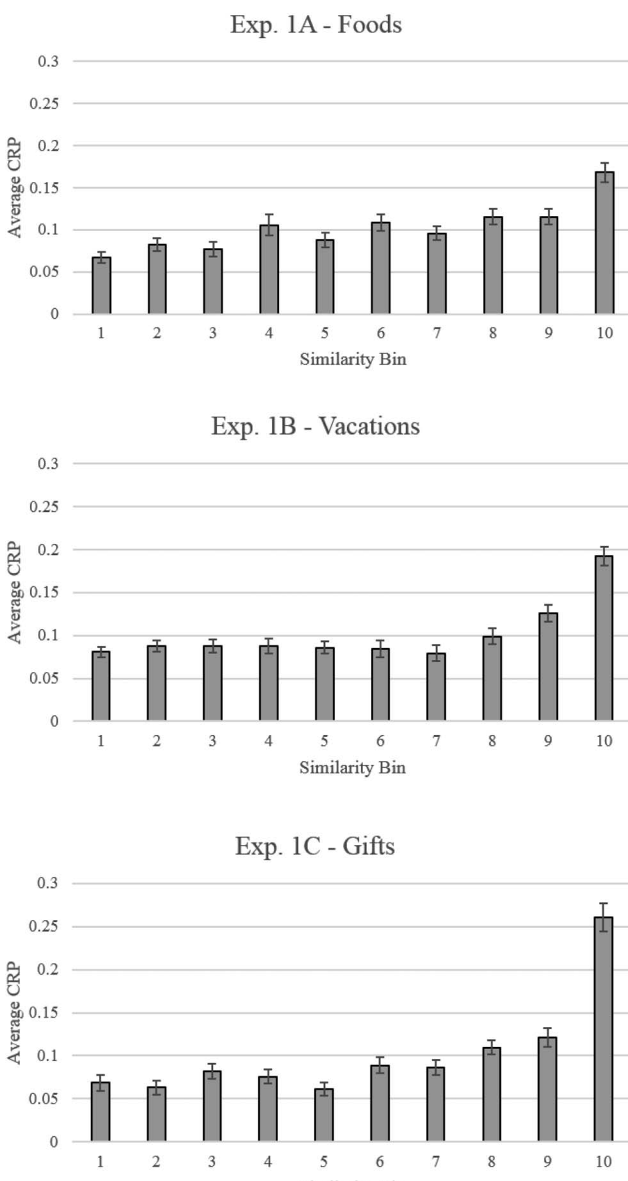

Figure 1. A (top), B (middle), and C (bottom): The average conditional response probability (CRP) for each of the 10 pairwise similarity bins for Experiments 1A through 1C. Errors bars display +/–1 standard error.

The preceding analysis aggregates CRPs for participants for each bin, however, a more rigorous test should also consider individual heterogeneity in the CRP-bin correlations. For this purpose, I ran a series of regressions (one for each experiment), in which the dependent variables were the calculated CRP values and the independent variables were the bin numbers. These regressions considered each CRP for each participant's bin to be a separate observation, and also permitted random effects on the participant level. These regressions found a very strong positive effect for all three experiments ($\beta = 0.008$, $z = 7.52$, $p < 0.001$, 95% CI = [0.006, 0.010] for Experiment 1A; $\beta = 0.008$, $z = 7.78$, $p < 0.001$, 95% CI = [0.006, 0.010] for Experiment 1B; $\beta = 0.014$, $z = 11.33$, $p < 0.001$, 95% CI = [0.012, 0.017] for Experiment 1C), again providing evidence for semantic clustering in the generation of memory-based choice sets.

Path analysis. Yet another approach to measuring semantic clustering has been proposed by Romney et al. (1993). Romney et al. (1993) used multidimensional scaling to derive pairwise semantic distance measures for all items in a free recall from lists task. They then computed the total distance traveled in a participant's retrieved item path. This distance is the sum of the distances between the first and second retrieved item, the second and the third retrieved item, the third and fourth retrieved item, and so on. To quantify the extent of semantic clustering, they compared the participant's actual distance with the distance that would be expected if the participant retrieved items in a random order. The latter can be calculated through Monte Carlo methods.

To ensure the robustness of the CRP-based analysis performed previously, I applied Romney et al.'s (1993) techniques to the present dataset. Again this required two minor modifications. First, instead of using multidimensional scaling-based distances, I used a measure $D(x,y) = 1 - S(x,y)$, with vectors x and y derived from the preexisting Word2Vec representations, and $S(x,y)$ corresponding to the cosine similarity between these vectors. D is bounded below by 0 and is thus a suitable measure of semantic distance (as assessed on the Word2Vec space). Second, as the set of items listed by participants was unconstrained (unlike Romney et al.'s set, which was determined by a list of items presented to participants prior to the recall task), I computed random path distance only on the set of items listed by each participant. This implies that these tests compared the participant's actual path over that participant's retrieved items, with a random path over the same set of retrieved items. In order to calculate these random paths, I conducted 100,000 simulations, with each simulation generating a random ordering over the set of items retrieved by the participant.

Our path analysis again showed strong effects of semantic clustering. This is illustrated in Figures 2A through 2C which plot each participant's actual distance against the mean distance of a random path on the participant's set of listed items. As is shown in these figures, the actual distances for participants were almost always shorter than the distances of the corresponding random paths. Specifically, 88% of participants in Experiment 1A, 96% of participants in Experiment 1B, and 94% of participants in Experiment 1C had shorter actual distances than mean random distances ($p < .001$ for all three experiments according to a binomial test).

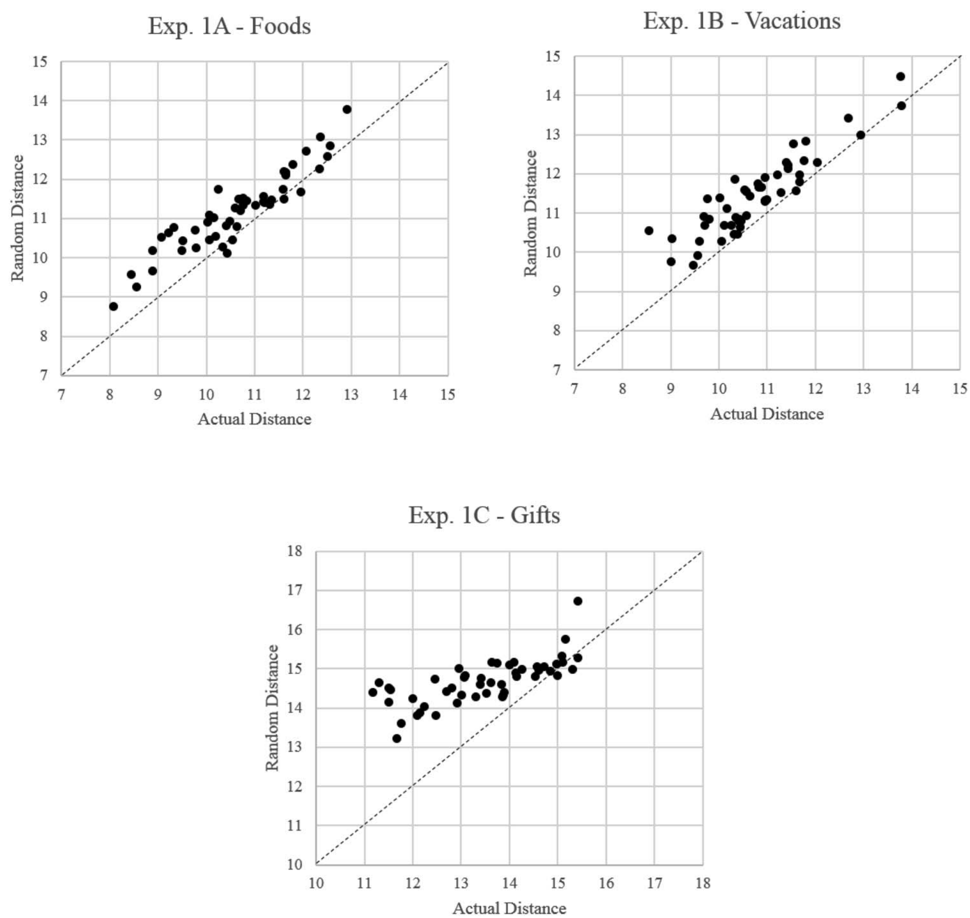

Figure 2. A (top left), B (top right), and C (bottom): Scatter plots of distances of participants' sequence of retrieved items versus average distances if sequence order was randomized, for Experiments 1A through 1C. Each point corresponds to a single participant in each of the three experiments.

Using a paired *t* test to compare the actual distances with the mean random distances, across participants, also showed that on aggregate these distances were significantly shorter, $t(49) = 8.84$, $p < 0.001$ for Experiment 1A; $t(49) = 10.09$, $p < 0.001$ for Experiment 1B; $t(49) = 9.54$, $p < 0.001$ for Experiment 1C.[2]

Desirability. A crucial variable missing from the preceding tests is item desirability. As discussed in the introductory sections of this article, efficient decision making involves the retrieval of the most desirable items first, and it is possible that a desirability-based retrieval strategy could indirectly generate the appearance of semantic clustering (without any semantic processes directly at play).

A simple analysis examining the relationship between item desirability and retrieval order showed that desirability does play a role in retrieval. This is illustrated in Figures 3A through 3C, which plots the desirabilities of items retrieved in each retrieval position, averaged across participants. As can be seen here, the items retrieved first had a high average desirability and the items retrieved last had a low average desirability (though these items were still rated to be somewhat desirable, rather than undesirable). Overall there was a correlation of −0.74 between item order (1, 2,. . . 20) and average item desirability in Experiment 1A (foods), a correlation of −0.93 in Experiment 1B (vacations), and a correlation of −0.76 in Experiment 1C (gifts). I also tested the desirability effect by performing a regression analysis over all the data, in which the dependent variable was retrieval order (1, 2, . . . 20), and the independent variable was the participant's rated item desirability, and each observation corresponded to a single retrieved item for a single participant. This analysis controlled for participant-level heterogeneity with random effects, and found a significant effect of desirability on retrieval order for all three experiments ($\beta = -0.60$, $z = -4.48$, $p < 0.001$, 95% CI [−0.86, −0.34] for Experiment 1A; $\beta = -1.30$, $z = -8.53$, $p < 0.001$, 95% CI [−1.60, −1.00] for Experiment 1B; $\beta = -0.51$, $z = -4.28$, $p < 0.001$, 95% CI [−0.75, −0.28] for Experiment 1C).

One way to control for the effect of desirability on retrieval, in order to examine semantic clustering effects independent of desir-

[2] Note that the preregistration analysis plan had specified the use of *z* tests for each participant, however a paired *t* test is better suited to evaluating differences across participants in the experiment. For this reason, I report only the paired *t* test here.

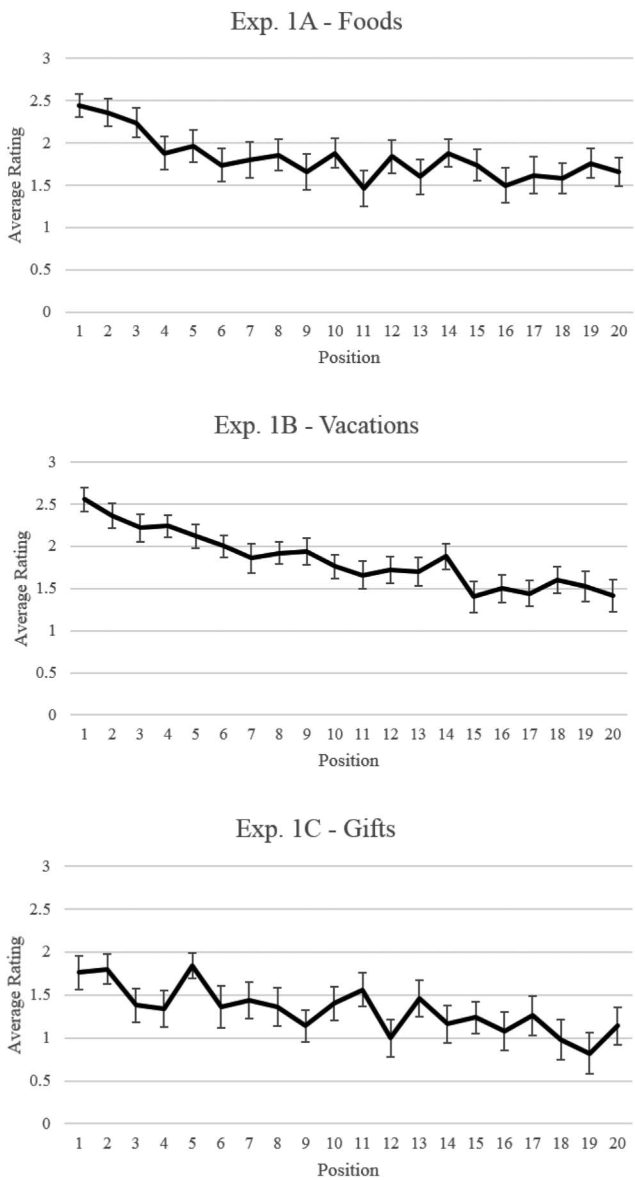

Figure 3. A (top), B (middle), and C (bottom): The average desirability ratings for items listed in different positions (1 for listed first to 20 for listed last), for Experiments 1A, 1B, and 1C. A rating of +3 corresponds to extremely desirable and a rating of 0 corresponds to neither desirable nor undesirable. Errors bars display +/–1 standard error.

ability, is to perform a variant of the path analysis presented in the previous section. Here, instead of comparing the participant's actual semantic distance with a random path distance, one can compare the participants' actual distance with the distance of a retrieval path that lists items in order of their desirability (with the most desirable item listed first, the second most desirable item listed second, and so on). If the participant's path distance is shorter than the desirability path distance, one can conclude that semantic clustering effects emerge even when controlling for the type of artificial semantic clustering that can be generated by a purely desirability-based retrieval strategy. In fact, one can also compare the desirability path distance with the random path distance to quantify the extent of this type of indirect desirability-based semantic clustering.

Note that the desirability rating task offered participants a scale with seven points (−3 to +3 in increments of 1), and for this reason, many items were given the same rating by the participant. Thus there are multiple possible desirability-based paths, with each path permitting a different ordering (but nonetheless restricting the ordering so that more desirable items are always listed before less desirable items). Thus to calculate the mean desirability-based path distance across all feasible desirability-based paths, I again used Monte Carlo methods with 100,000 simulated desirability-based orderings. For each simulation, I generated a strict ordering for listed items based on their desirability ratings, with the positioning of identically rated items being randomized. The overall desirability path distance for the simulation was calculated as the sum of the distances of adjacent items in the ordering.

A comparison between the participants' actual distances and the mean desirability path distances is provided in Figures 4A through 4C. As can be seen in these figures, actual distances were almost always shorter than mean desirability path distances. Specifically, 82% of participants in Experiment 1A, 88% of participants in Experiment 1B, and 88% of participants in Experiment 1C had shorter actual distances than mean desirability distances ($p < .001$ for all three experiments according to a binomial test). Again a paired t test confirmed that these difference were significant, $t(49) = 6.59$, $p < 0.001$ for Experiment 1A; $t(49) = 7.56$, $p < 0.001$ for Experiment 1B; $t(49) = 7.93$, $p < 0.001$ for Experiment 1C.[3] A similar paired t test comparing average desirability paths with average random paths also found that desirability paths were shorter than random paths, supporting my earlier claim that a purely desirability-based retrieval strategy can indirectly lead to the appearance of semantic clustering, $t(49) = 3.23$, $p < 0.01$ for Experiment 1A; $t(49) = 1.64$, $p = .11$ for Experiment 1B; $t(49) = 6.21$, $p < 0.001$ for Experiment 1C. Note that this difference does not reach statistical significance for Experiment 1B.

Now, the finding that the actual paths of decision makers were different from the desirability paths suggests that undesirable items were being retrieved earlier than if decision makers were able to retrieve items only in order of desirability. To rigorously examine this, I compared the relationship between desirability and retrieval order (as in Figures 3A through 3C) with the relationship that one would expect if the items were retrieved strictly in order of their desirability. Specifically, for each participant, I reordered the retrieved set of items based on their desirabilities to generate an ideal ordering, in which the most desirable item is retrieved first, the

[3] The preregistration analysis plan had specified the use of a linear regression, instead of this path analysis. In this regression, each observation would capture the semantic similarity (dependent variable), ordering distance (independent variable), and desirability difference (control variable) between each pair of listed items for each participant. Although such a regression does reveal a significant negative relationship between ordering distance and semantic similarity, controlling for desirability, I have realized that this regression is not statistically sound, as observations cannot be considered to be independent (i.e. the semantic similarity of items A and B and B and C influences the semantic similarity of items A and C). The path analysis provides a better way of disentangling desirability from semantic clustering, which is why I have listed the path analysis results here.

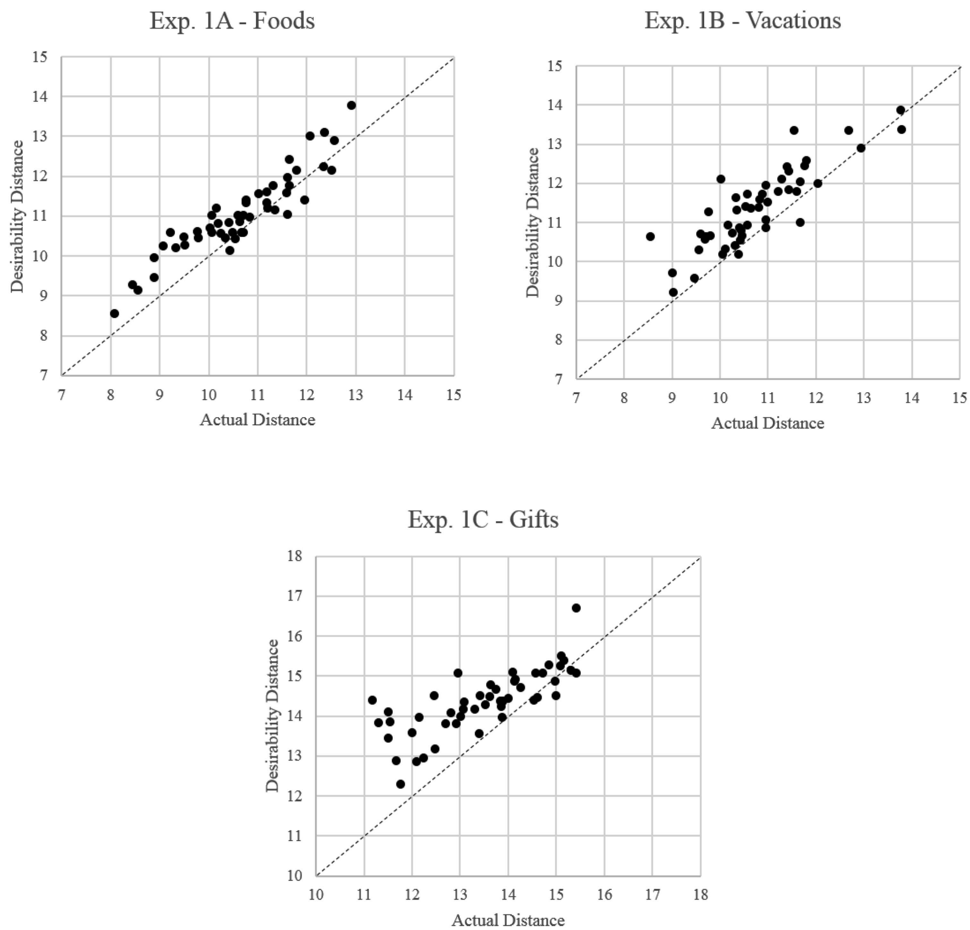

Figure 4. A (top left), B (top right), and C (bottom): Scatter plots of distances of participants' sequence of retrieved items versus average distances if items were listed in order of desirability, for Experiments 1A, 1B, and 1C. Each point corresponds to a single participant in each of the three experiments.

second most desirable item is retrieved second, and so on. I then subtracted the desirability of each item in each of the positions in the participant's retrieved list from the desirability of the item in the corresponding position in the ideal list. This gave us a list with differences in desirabilities in each of the 20 retrieval positions. If decision makers did retrieve undesirable items earlier and desirable items later (than ideal), then the differences in this list would be decreasing in order (e.g., a positive difference would be observed for the first position and a negative difference would be observed for the last position). I tested for this decreasing relationship in regressions in which the dependent variable was the desirability difference and the independent variable was the corresponding order position. These regressions considered each of the 20 retrieval positions for each of the participants to be a separate observation, and controlled for participant-heterogeneity with random effects. They revealed a significant negative effect of position for all three experiments ($\beta = -2.76$, $z = -26.92$, $p < 0.001$, 95% CI [−2.96, −2.56] for Experiment 1A; $\beta = -3.44$, $z = -30.19$, $p < 0.001$, 95% CI [−3.67, −3.22] for Experiment 1B; $\beta = -2.94$, $z = -38.04$, $p < 0.001$, 95% CI [−3.09, −2.79] for Experiment 1C). This confirms that in the present experiments, undesirable items were being retrieved significantly earlier than ideal (i.e., significantly earlier than if decision makers were able to retrieve items strictly in order of desirability).[4]

Discussion

The goal of Experiments 1A through 1C was to test for semantic clustering effects in the retrieval of choice sets from memory. It used two existing techniques to quantify semantic clustering: the first, based on the conditional response probability approach outlined by Howard and Kahana (2002), involved testing whether the retrieval of one item increased the likelihood of similar items being retrieved in the successive time period. The second, based on the path analysis approach outlined by Romney et al. (1993) involved calculating the total distance traveled in semantic space by a participant's retrieval sequence, and comparing this to the distance expected if retrieval order was randomized.

Both of these techniques indicated strong semantic clustering effects in Experiments 1A through 1C, however neither of these

[4] This analysis was not mentioned in the preregistered analysis plan, but is necessary for testing whether undesirable items were retrieved earlier than would be predicted by desirability-based retrieval.

techniques controlled for the effect of desirability on retrieval. Desirable items were more likely to be retrieved first, suggesting that the semantic clustering observed using the CRP and the path analysis techniques could be due to decision makers merely having similar preferences for similar items, and not due to any direct semantic effects in retrieval. To control for desirability, I thus ran a variant of the path analysis that calculated the average semantic distances of retrieval sequences that listed items purely in order of desirability. I found that such desirability-based distances were still much larger than the distances of the participants' actual retrieval sequences, suggesting the existence of direct semantic clustering effects in addition to desirability-based retrieval biases.

The presence of semantic clustering suggests that undesirable items that are semantically related to previously retrieved items can occasionally intrude and be retrieved before more desirable (but semantically unrelated) items. This hints at a potential inefficiency in the decision process—With perfect memory retrieval, decision makers would be quicker and less prone to error if they did not cluster retrieval semantically, and instead first considered the most desirable items before examining other less desirable items.

Experiments 2A, 2B, and 2C: Choice Context

In Experiments 1A through 1C, there was a robust semantic clustering effect, showing that items that are semantically related to previously retrieved items are themselves more likely to be retrieved in successive time periods. This illustrates the key role of semantic relatedness in determining the items that are considered by decision makers in decisions from memory. In Experiments 2A through 2C, I examine a second semantic effect, this time involving the influence of choice context on the retrieval of items. These experiments consider choice domains that are identical to those used in Experiments 1A through 1C but vary the contextual cues given to participants in these domains. If semantic processes do guide retrieval, one should expect items that are semantically related to the choice context to be more likely to be retrieved.

Experiments 2A through 2C involve identical procedures to each other, and vary only in terms of the decision domain, and thus I present their methods and results simultaneously. Again, all three experiments have been preregistered and approved by the University of Pennsylvania IRB.[5]

Method

Participants. Participants (N = 100; mean age = 31; 44% female) in Experiment 2A, participants in Experiment 2B (N = 100; mean age = 30; 47% female), and participants in Experiment 2C (N = 100; mean age = 30; 41% female), recruited from Prolific Academic, performed the experiment online. All participants were residents of the United States. They were compensated at a rate of approximately $US7.50/hr.

Procedures. In all three experiments participants were shown a description of a decision setting and were asked to list 20 items that came to their mind as they considered making their decisions. The specific choice context used in the decision setting was, however, randomly assigned to participants, with each participant receiving one of two contextual cues. The instructions (with the randomized choice context underlined) were as follows:

> Experiment 2A (foods): Imagine that you could eat anything that you wanted for [dinner/breakfast] tomorrow. In the boxes below please list 20 food items that come to your mind as you consider what to eat. Please make sure to list all foods that you think of, regardless of whether you would eventually want to eat them. Please list these foods in the order that they come to your mind (i.e., the first food that comes to your mind listed first, the second listed second).

> Experiment 2B (vacations): Imagine that you could go on a [wine tasting/camping] trip to any destination that you wanted. In the boxes below please list 20 vacation destinations that come to your mind as you consider where to go. Please make sure to list all destinations that you think of, regardless of whether you would eventually want to go the destination. Please list these vacation destinations in the order that they come to your mind (i.e., the first destination that comes to your mind listed first, the second listed second, etc.).

> Experiment 2C (gifts): Imagine that you have to purchase a gift for your [partner for Valentine's day/friend for her baby shower]. In the boxes below please list 20 items that come to your mind as you consider what gift to purchase. Please make sure to list all items that you think of, regardless of whether you would eventually want to purchase them. Please list these items in the order that they come to your mind (the first item that comes to your mind listed first, the second listed second).

Participants listed all the items on a single screen. After these items were listed participants were taken to a second screen on which they rated each of their 20 items for desirability in their corresponding choice context, on a scale from −3 to +3 (−3 corresponding to extremely undesirable, 0 to neither desirable nor undesirable, and +3 to extremely desirable). Thus participants in Experiment 2A who were given the dinner context rated their 20 items based on how much they would like to eat each of the items for dinner, and participants in this experiment who were given the breakfast context rated their items on how much they would like to eat each of the items for breakfast. Likewise, participants in Experiment 2B rated their items based on how much they would like to go there for a wine tasting or a camping trip, and participants in Experiment 2C rated their items based on how much they would like to give the item as a Valentine's day or baby shower gift.

Results

Context dependence. Our primary tests for Experiments 2A through 2C involved the effect of the choice context on the set of retrieved items. For this purpose, I used the Word2Vec representations to obtain vectors for each pair of contextual cues.[6] I then calculated the relative cosine similarities between these vectors and each of the items listed by the participants to measure the relative semantic relatedness between listed items and the two choice contexts in each of the three experiments. More formally, for each item x in Experiment 2A, I calculated $S(x, \mathit{dinner}) - S(x, \mathit{breakfast})$, for each item x in Experiment 2B, I calculated $S(x, \mathit{wine\ tasting}) - S(x, \mathit{camping})$, and for each item x in Experiment 2C, I

[5] Preregistration materials are available at: https://osf.io/zudp7/.

[6] Note that the Word2Vec representations had individual vectors for *Valentines_Day*, *baby_Shower*, and *wine_tasting*, and I used these individual vectors instead of averaging the vectors for component words in these phrases.

calculated $S(x$, *Valentine's day*$) - S(x$, *baby shower*). I then compared these relative cosine similarity values across the two groups of participants in each experiment. If choice context does influence the items retrieved, one would expect $S(x$, *dinner*$) - S(x$, *breakfast*) to be larger for participants in Experiment 2A in the dinner context relative to the breakfast context. Likewise $S(x$, *wine tasting*$) - S(x$, *camping*) should be larger for participants in Experiment 2B in the wine tasting context relative to the camping context, and $S(x$, *Valentine's day*$) - S(x$, *baby shower*) should be larger for participants in Experiment 2C in the Valentine's day context relative to the baby shower context.

The relationship between item retrieval and choice context can be seen in Figures 5A through 5C, which plot the average cosine similarity differences across participants, for the three experiments, as a function of condition. The solid lines always lie above the dashed lines, indicating that retrieved items were more similar to their choice context compared with the choice context in the alternate condition. More rigorously, I tested for these differences using three regressions (one for each experiment) in which the main dependent variable was the cosine similarity difference ($S[x$, *dinner*$] - S[x$, *breakfast*] for Experiment 2A, $S[x$, *wine tasting*$] - S[x$, *camping*] for Experiment 2B, and $S[x$, *Valentine's day*$] - S[x$, *baby shower*] for Experiment 2C) and the main independent variable was the experimental conditions to which the participants were assigned (1 if condition was *dinner* in Experiment 2A, *wine tasting* in Experiment 2B, and *Valentine's Day* in Experiment 2C; 0 otherwise). I also controlled for participant-heterogeneity with random effects. This regression illustrated strong positive effects of condition on cosine similarity, showing that choice context does in fact influence item retrieval ($\beta = 0.04$, $z = 9.16$, $p < 0.001$, 95% CI [0.03, 0.05] for Experiment 2A; $\beta = 0.08$, $z = 8.33$, $p < 0.001$, 95% CI [0.06, 0.10] for Experiment 2B; $\beta = 0.08$, $z = 15.72$, $p < 0.001$, 95% CI [0.07, 0.09] for Experiment 2C).

We also ran variants of the preceding regressions with three additional independent variables: the first variable was the retrieval position of the item (1 to 20, based on whether the item was listed first or last). The second variable was an interaction between retrieval position and the condition variable. This variable tested whether the effect of choice context diminished or increased as additional items were retrieved. The final variable was the choice item's rated desirability. I found that the effect of the condition variable on cosine similarity persisted with these additional covariates ($p < 0.001$ for all). Moreover, there was a negative interaction effect between the condition and the item ordering in Experiments 2B (vacations) and 2C (gifts), indicating that the effect of the choice context diminished as additional items were retrieved ($\beta = -0.002$, $z = -2.94$, $p < 0.01$, 95% CI $[-0.003, -0.001]$ for Experiment 2B; $\beta = -0.006$, $z = -9.90$, $p < 0.001$, 95% CI $[-0.007, -0.005]$ for Experiment 2C). This effect can also be observed in Figures 5B and 5C which show that the differences in average cosine similarities between the solid and dashed lines decrease for items retrieved later on during deliberation. There was no such interaction effect for Experiment 2A (foods; $p = .42$).

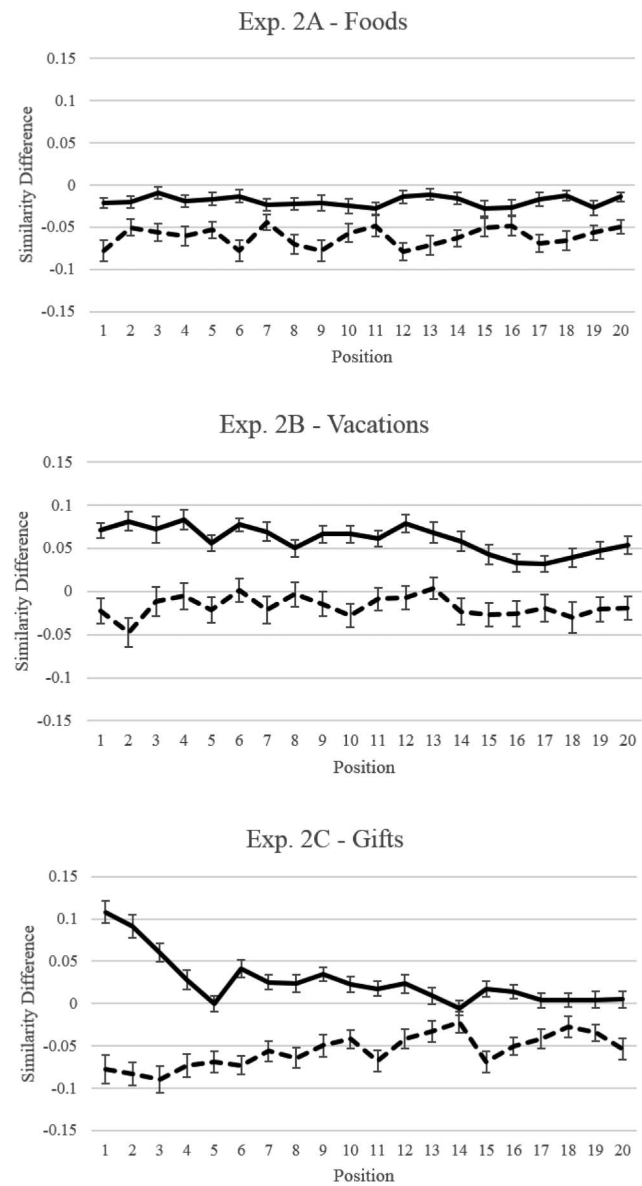

Figure 5. A (top), B (middle), and C (bottom): The similarity differences for items listed in different positions (1 for listed first to 20 for listed last), as a function of condition in Experiments 2A, 2B, and 2C. Here the solid lines correspond to choice contexts of dinner, wine tasting, and Valentine's day, and dashed lines correspond to contexts of breakfast, camping, and baby shower, in Experiment 2A, 2B, and 2C respectively. Errors bars display +/–1 standard error.

Semantic clustering and desirability. We also analyzed the data from Experiments 2A through 2C to test if the semantic clustering effects observed previously were replicated. For this purpose, I analyzed each of the conditions in each of the experiments separately. I first calculated the relationship between conditional response probabilities and sematic relatedness with retrieved items (in a manner that was identical to that reported for Experiments 1A through 1C). For this, I ran six regressions (one for each condition in each experiment), in which the dependent variables were the calculated CRP values and the independent variables were the bin numbers (we also included random effects on the participant-level). As in Experiments 1A through 1C, these regressions found a very strong positive effect of semantic relatedness on CRP ($p < 0.001$ for all six tests).

We replicated this result using the path analysis approach outlined in the preceding text. For this, I again compared the semantic distance over the path of items listed by participants with the distance expected if item retrieval order was random and if the item retrieval order was determined entirely in order of desirability (with the most desirable item listed first). Again, I found that the actual distances were lower than random distances and desirability distances, indicating the presence of semantic clustering. These differences were significant, as evaluated by paired *t* tests across participants in each condition in each experiment ($p < 0.001$ for all six tests).

A final analysis involved examining the effect of desirability on retrieval order. For this purpose, I ran a regression analysis in which the dependent variable was retrieval order (1 to 20, based on whether the item was listed first or last), and the independent variable was the participant's rated item desirability. I also included a control for the cosine similarity of the item with the contextual cue. This analysis counted each retrieved item by each participant as a separate observation and controlled for participant-heterogeneity with random effects. Replicating the findings of Experiments 1A through 1C, I found positive significant effects for desirability ($p < 0.001$ for all six tests), indicating that desirable items are more likely to be retrieved earlier, even when controlling for their similarity with the choice context.[7]

Discussion

Experiments 2A through 2C tested for the effect of contextual cues on item retrieval. They specified semantic relatedness using cosine similarity between Word2Vec word vectors. As expected they found that decision makers were more likely to retrieve items that are semantically related to the choice context. Additionally, the effect of choice context was stronger for items retrieved early on during deliberation for Experiments 2B (vacations) and 2C (gifts). As deliberation continued, retrieved items in these experiments drifted further away from the contextual cue. Additionally, as in Experiments 1A through 1C, Experiments 2A through 2C also noted a strong degree of semantic clustering, with retrieved items increasing the likelihood of semantically related items being retrieved in the subsequent time period. I also observed the effect of desirability on item retrieval: Highly desirable items were more likely to be retrieved early on in the decision. Overall Experiments 2A through 2C find evidence for another important semantic determinant of item retrieval: choice context. These experiments also illustrate the value of semantic space models for formalizing semantic relatedness between contextual cues and item listings.

General Discussion

Semantic processes play a key role in guiding behavior in a wide range of cognitive tasks. I have shown that these processes also influence the choice items that come to mind when decision makers have to generate choice sets from memory. Particularly, in Experiments 1A through 1C, I found that retrieved items increase the retrieval probability of other items that they are semantically related to. This generates a semantic clustering effect according to which semantically related items are retrieved together. This clustering effect exists in addition to desirability-based retrieval biases. In Experiments 2A through 2C, I found that contextual cues, such as choice context, increase the retrieval probabilities of items that are semantically related to these cues. This effect is usually stronger early on in deliberation.

Memory-Based Choice Sets

Most experimental research in psychology presents explicit choice sets to participants, and is consequently focused on understanding the decision rules that are used to evaluate items in these choice sets (see, e.g., Busemeyer & Rieskamp, 2014; Oppenheimer & Kelso, 2015). However, many everyday decisions do not involve externally provided choice sets; rather decision makers must retrieve feasible choice items from memory (Alba & Hutchinson, 1987; Lynch & Srull, 1982). It is these decisions from memory that are sensitive to the influence of semantic processes.

Our results suggest that to build more comprehensive theories preferential decision making, psychologists should attempt to model the semantic memory mechanisms responsible for choice item retrieval in the absence of exogenous choice sets. There are already theories of preferential choice that involve a significant memory component, and permit the influence of probes and cues to guide memory retrieval (e.g., Dougherty et al., 1999; Glöckner et al., 2014; Johnson et al., 2007). Out of these, the associative accumulation model (Bhatia, 2013) has perhaps the most explicit semantic component for simple preferential choice, in that it allows for salient choice items and other contextual variables in the choice task to bias the activation of associated attributes, and thus increase the activation of other (semantically related) choice items. It is possible that such a model could be combined with existing theories of semantic memory to more rigorously describe the retrieval patterns documented in this article.

One benefit of this endeavor would be a more detailed analysis of the effect of noise in memory retrieval and memory-based choice. In the current article, I did not consider the possibility that some of the findings could be influenced by noise in memory. Although there is no reason to expect this type of noise to artificially generate semantic clustering or context dependence, noise could interact in complex ways with underlying semantic structure. A formal model of memory retrieval that allows for the effect of noise would be able to better control for this interaction.

Semantic processes are not the only memory mechanisms at play in memory-based decisions. The decision maker's familiarity with individual items (which can be measured by the frequency of the item's occurrence in the decision maker's environment) is likely to guide retrieval as well. This type of familiarity-based memory effect has already been shown to play a role in a number of decision making domains, and existing models of memory-based judgment and decision making explicitly allow for the influence of such factors on item activation and retrieval (Gigerenzer & Goldstein, 1996; Marewski & Schooler, 2011).

Another important variable influencing the retrieval of items in memory-based choice tasks is the desirability of the items. In these experiments, I found that items rated as being highly desirable were also the ones that were most likely to be listed earlier during deliberation. The causes of this desirability bias are currently unclear. For example, desirability could influence the baseline activation of items making desirable items easier to retrieve. This would imply that more desirable items are also more likely to be

[7] The semantic clustering and desirability-based analysis outlined in this subsection was not mentioned in the preregistration analysis plan, however it does provide a convenient replication of the results of Experiments 1A, 1B, and 1C.

retrieved in memory tasks without an explicit preferential component (congruent with the finding that word emotionality increases word recall—see Rubin & Friendly, 1986). Alternatively, desirability could act as a contextual cue; a cue that selectively increases the activation of desirable choice items only when decision makers have to make preferential decisions (see Talmi & Moscovitch, 2004 for a related point). This would imply that the desirability biases observed in these experiments are unique to preferential decision making contexts. Understanding the influence of desirability on retrieval in memory-based choice, as well as the interaction between desirability and various semantic and nonsemantic memory processes, is an important topic for future work.

Efficient Memory

The tests performed in this article controlled for the influence of desirability on retrieval, and demonstrated the existence of a distinguishable, direct, semantic clustering effect. One consequence of this effect is that decision makers often retrieved undesirable items that were semantically related to previously retrieved items before retrieving more desirable but semantically unrelated items. This type of semantic interference can be seen as being detrimental to the decision: If decision makers needed to maximize accuracy and minimize decision time, retrieval would be entirely in order of desirability.

Although these experiments suggest the presence of a type of inefficiency in memory-based decisions, semantic clustering can alternatively be seen as reflecting the effortless use of an existing memory system. For example, a purely desirability-based retrieval strategy would likely put additional cognitive demands on the decision maker, as decision makers would have to suppress the associative activation processes responsible for semantic clustering, while simultaneously searching for and activating only the most desirable items stored in their memories. In fact, to the extent that similar items are similarly desirable, semantic clustering can even approximate desirability-based retrieval, implying that the effects observed in this article may reflect adaptive memory search in decision making. Interestingly, executive control has been shown to lead to increased semantic clustering in memory tasks (Hills, Mata, Wilke, & Samanez-Larkin, 2013; Hills, Todd, & Goldstone, 2010), suggesting that individuals with higher working memory span may be especially likely to display the types of patterns documented in this article.

Another way in which semantic processes facilitate good decisions involves the use of contextual cues. As I have shown in these experiments, decision makers selectively retrieve items that are semantically related to the choice context. Thus when deliberating over what to eat for breakfast, they are able to easily and quickly retrieve breakfast-related food items. In contrast, when deliberating over what to eat for dinner, they are able to easily and quickly retrieve dinner-related food items. Preferences for choice items depend on context and automatically responding to choice context likely leads quicker and more accurate decisions.

Decision Rules

A complete understanding of the relationship between semantic memory and efficient or inefficient decision making, however, depends on the choice rules that are used by decision makers to evaluate retrieved items. If item retrieval is sequential then decision makers could utilize alternative-wise decision rules such as the satisficing heuristic. Such rules consider the overall desirability of an item and accept or reject the item before considering other items (e.g., Simon, 1956). Alternatively, if retrieval is in parallel, or if decision makers evaluate items only after multiple items have been retrieved, choice may involve attribute-wise decision strategies such as the lexicographic heuristic. Such rules evaluate the attributes of choice alternatives one after the other and accept or reject items based on their values on these attributes (e.g., Gigerenzer & Goldstein, 1996; Tversky, 1972). Other decision rules, involving a mix of attribute and alternative-wise processing are also possible (see, e.g., Bhatia, 2017c; Glöckner et al., 2014; Roe, Busemeyer, & Townsend, 2001), as are decision rules that select "typical" items such as those that are overall most similar to the entire set of retrieved items. It is possible that the observed retrieval strategies are optimal with one set of decision rules but not with another (also see Payne, Bettman, & Johnson, 1993).

Of course the study of the decision rules used by decision makers in memory-based choice tells us about more than just whether memory is efficient. Understanding these rules allows researchers to fully characterize the memory-based decision process and thus predict not only the items that decision makers consider while deliberating, but also the items that they eventually choose and how long they take to make their choices. Although one would expect the findings documented here to generalize to settings in which choices are required (as items retrieved by decision makers fundamentally constrain their choice sets) a direct study of decision rules used in memory-based choice can provide a more detailed understanding of the effect of semantic clustering and context dependence on actual decision outcomes. It is possible that these effects are diminished with some decision rules, and amplified with others.

Although a full characterization of the memory-based decision process is still lacking, there has been some work that has contrasted memory-based decision making against stimulus-based decision making in areas such as marketing (Lynch, Marmorstein, & Weigold, 1988; Lynch & Srull, 1982; Nedungadi, 1990; Rottenstreich, Sood, & Brenner, 2006). Much of this work involves providing information regarding the available choice items and their attributes to decision makers prior to the decision task, and then asking decision makers to recall this information as they decide. Behavior in such memory-based tasks is compared with behavior in equivalent stimulus-based tasks in which information is provided to decision makers as they decide (and need not be retrieved from memory). This work has documented systematic differences in choice strategies and choice outcomes between memory-based and stimulus-based decisions, implying that existing psychological theories of decision making may not be directly applicable to the types of memory-based choices studied in the current article.

Semantic Space Models

One important contribution of this article is that it shows how decisions involving everyday choice items can be rigorously studied with the use of semantic space models (Griffiths et al., 2007; Jones & Mewhort, 2007; Kwantes, 2005; Landauer & Dumais, 1997; Lund & Burgess, 1996; Mikolov et al., 2013). These decisions typically involve a large, unconstrained, choice domain: In these experiments, participants were free to list any choice items that came to their mind. Psychological research often has difficulties quantifying relevant variables (such as the semantic relatedness between pairs of items) in such unconstrained settings, making naturalistic memory-based decisions very hard to study. This is perhaps one reason why most existing decision making research utilizes artificial experimenter provided choice sets. However, semantic space models make it is possible to obtain semantic representations for a very large set of words and concepts. These representations provide measures of semantic relatedness for nearly any pair of choice items, and between choice items and a wide range of choice contexts.

Our application of semantic space models to measure semantic relatedness was influenced by similar existing applications to free recall and free association (e.g., Hills et al., 2012; Howard & Kahana, 2002). Besides applications to memory, such models are also useful for predicting behavior in a variety of other psychological tasks, including similarity judgment, categorization, text comprehension, and semantic priming (see Bullinaria & Levy, 2007 or Jones et al., 2015 for a review), as well as various high-level cognitive tasks studied by scholars of judgment and decision making (Bhatia, 2017a, 2017b).

Of course, semantic space models do have limitations. As they rely on large natural language corpora for model training, most such models are unable to accommodate individual differences in semantic representations. For example, some of the participants in Experiment 1A may have considered *rice* to be more semantically related to *stir fried vegetables*, whereas others may have considered it to be more semantically related to *refried beans*. As these tests involved a single set of pretrained vector representations, they predicted the same pairwise similarity measures for items across participants, and thus were unable to accommodate any differences in recall across participants that may have stemmed from differences in item representation. Another issue involves items composed of multiple words. In these tests, I merely averaged the word vectors for the component words in multiword items, however this type of averaging may not always be the best way to obtain composite representations.

Despite these limitations, semantic space models provide a unique and promising approach to the study of preferential choice, an area of research that is concerned primarily with how individuals aggregate and evaluate information about real-world items. The psychological processes involved in such tasks cannot be easily understood by presenting decision makers with explicit choice sets composed of abstracted artificial experimenter-generated stimuli. Rather the study of real-world choice requires the use of both naturalistic stimuli and naturalistic methods for eliciting preferences. By providing a way to approximate the representations that individuals have for complex everyday choice items, semantic space models allow for the extension of existing psychological theories (such as theories of free recall) to a wide range of real-world phenomena. I look forward to further applications of semantic space models to the study of decision making in psychology and related fields.

References

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *The Journal of Consumer Research, 13,* 411–454. http://dx.doi.org/10.1086/209080

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38,* 341–380. http://dx.doi.org/10.1006/jmla.1997.2553

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation, 2,* 89–195. http://dx.doi.org/10.1016/S0079-7421(08)60422-3

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review, 120,* 522–543. http://dx.doi.org/10.1037/a0032457

Bhatia, S. (2017a). The semantic representation of prejudice and stereotypes. *Cognition, 164,* 46–60. http://dx.doi.org/10.1016/j.cognition.2017.03.016

Bhatia, S. (2017b). Associative judgment and vector space semantics. *Psychological Review, 124,* 1–20. http://dx.doi.org/10.1037/rev0000047

Bhatia, S. (2017c). Choice rules and accumulator networks. *Decision, 4,* 146–170. http://dx.doi.org/10.1037/dec0000038

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of sequences of restricted associative responses. *The Journal of General Psychology, 30,* 149–165. http://dx.doi.org/10.1080/00221309.1944.10544467

Brenner, L., Rottenstreich, Y., Sood, S., & Bilgin, B. (2007). On the psychology of loss aversion: Possession, valence, and reversals of the endowment effect. *The Journal of Consumer Research, 34,* 369–376. http://dx.doi.org/10.1086/518545

Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods, 39,* 510–526. http://dx.doi.org/10.3758/BF03193020

Busemeyer, J. R., & Rieskamp, J. (2014). Psychological research and theories on preferential choice. In S. Hess & A. Daly (Eds.), *Handbook of choice modelling* (pp. 49–72). Cheltenham, UK: Edward Elgar Publishing.

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion, 1,* 5–11. http://dx.doi.org/10.1037/1528-3542.1.1.5

Dougherty, M. R., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review, 106,* 180–209. http://dx.doi.org/10.1037/0033-295X.106.1.180

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650–669. http://dx.doi.org/10.1037/0033-295X.103.4.650

Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about adaptive decision making? A parallel constraint satisfaction account. *Cognition, 133,* 641–666. http://dx.doi.org/10.1016/j.cognition.2014.08.017

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review, 114,* 211–244. http://dx.doi.org/10.1037/0033-295X.114.2.211

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 225–240. http://dx.doi.org/10.1037/0278-7393.6.3.225

Hare, M., Jones, M., Thomson, C., Kelly, S., & McRae, K. (2009). Activating event knowledge. *Cognition, 111,* 151–167. http://dx.doi.org/10.1016/j.cognition.2009.01.009

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review, 119,* 431–440. http://dx.doi.org/10.1037/a0027373

Hills, T. T., Mata, R., Wilke, A., & Samanez-Larkin, G. R. (2013). Mechanisms of age-related decline in memory search across the adult life span. *Developmental Psychology, 49,* 2396–2404. http://dx.doi.org/10.1037/a0032272

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2010). The central executive as a search process: Priming exploration and exploitation across domains. *Journal of Experimental Psychology: General, 139,* 590–609. http://dx.doi.org/10.1037/a0020666

Hills, T. T., Todd, P. M., & Jones, M. N. (2015). Foraging in semantic fields: How we search through memory. *Topics in Cognitive Science, 7,* 513–534. http://dx.doi.org/10.1111/tops.12151

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers, 16,* 96–101. http://dx.doi.org/10.3758/BF03202365

Howard, M. W., Jing, B., Addis, K. M., & Kahana, M. J. (2007). Semantic structure and episodic memory. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 121–142). http://dx.doi.org/10.4324/9780203936399.ch7

Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language, 46,* 85–98. http://dx.doi.org/10.1006/jmla.2001.2798

Hutchinson, J. (1983). *Expertise and the structure of free recall*. Ann Abor, MI: ACR North American Advances.

Hutchinson, J. W., Raman, K., & Mantrala, M. K. (1994). Finding choice alternatives in memory: Probability models of brand name recall. *Journal of Marketing Research, 31,* 441–461. http://dx.doi.org/10.2307/3151875

Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 461–474. http://dx.doi.org/10.1037/0278-7393.33.3.461

Jones, M. N., & Mewhort, D. J. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review, 114,* 1–37. http://dx.doi.org/10.1037/0033-295X.114.1.1

Jones, M. N., Willits, J., Dennis, S., & Jones, M. (2015). Models of semantic memory. In J. Townsend, Z. Whang, & A. Eidels (Eds.), *Oxford handbook of mathematical and computational psychology* (pp. 232–254). Oxford, UK: Oxford university press.

Kahana, M. J. (2012). *Foundations of human memory*. Oxford, UK: Oxford university press.

Kwantes, P. J. (2005). Using context to build semantics. *Psychonomic Bulletin & Review, 12,* 703–710. http://dx.doi.org/10.3758/BF03196761

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240. http://dx.doi.org/10.1037/0033-295X.104.2.211

Lattin, J. M., & Roberts, J. H. (1992). *Testing for probabilistic independence in consideration of ready-to-eat cereals*. Palo Alto, CA: Graduate School of Business, Stanford University.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28,* 203–208. http://dx.doi.org/10.3758/BF03204766

Lynch, J. G., Jr., Marmorstein, H., & Weigold, M. F. (1988). Choices from sets including remembered brands: Use of recalled attributes and prior overall evaluations. *The Journal of Consumer Research, 15,* 169–184. http://dx.doi.org/10.1086/209155

Lynch, J. G., Jr., & Srull, T. K. (1982). Memory and attentional factors in consumer choice: Concepts and research methods. *The Journal of Consumer Research, 9,* 18–37. http://dx.doi.org/10.1086/208893

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review, 118,* 393–437. http://dx.doi.org/10.1037/a0024143

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds), *Advances in Neural Information Processing Systems* (pp. 3111–3119). Red Hook, NY: Curran associates.

Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 863–883. http://dx.doi.org/10.1037/0278-7393.21.4.863

Nedungadi, P. (1990). Recall and consumer consideration sets: Influencing choice without altering brand evaluations. *The Journal of Consumer Research, 17,* 263–276. http://dx.doi.org/10.1086/208556

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition, 28,* 887–899. http://dx.doi.org/10.3758/BF03209337

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

Oppenheimer, D. M., & Kelso, E. (2015). Information processing as a paradigm for decision making. *Annual Review of Psychology, 66,* 277–294. http://dx.doi.org/10.1146/annurev-psych-010814-015148

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York, NY: Cambridge University Press.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review, 116,* 129–156. http://dx.doi.org/10.1037/a0014420

Roberts, J. H., & Lattin, J. M. (1997). Consideration: Review of research and prospects for future insights. *Journal of Marketing Research, 34,* 406–410. http://dx.doi.org/10.2307/3151902

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review, 108,* 370–392. http://dx.doi.org/10.1037/0033-295X.108.2.370

Roediger, H. L., III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8,* 385–407. http://dx.doi.org/10.3758/BF03196177

Romney, A. K., Brewer, D. D., & Batchelder, W. H. (1993). Predicting clustering from semantic structure. *Psychological Science, 4,* 28–34. http://dx.doi.org/10.1111/j.1467-9280.1993.tb00552.x

Rottenstreich, Y., Sood, S., & Brenner, L. (2006). Feeling and thinking in memory-based versus stimulus-based choices. *Journal of Consumer Research, 33,* 461–469.

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,* 79–94. http://dx.doi.org/10.3758/BF03209231

Shapiro, S., MacInnis, D. J., & Heckler, S. E. (1997). The effects of incidental ad exposure on the formation of consideration sets. *The Journal of Consumer Research, 24,* 94–104. http://dx.doi.org/10.1086/209496

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63,* 129–138. http://dx.doi.org/10.1037/h0042769

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

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 281–299. http://dx.doi.org/10.1037/h0032955

Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin & Review, 1,* 89–106. http://dx.doi.org/10.3758/BF03200763

Received December 27, 2017
Revision received March 20, 2018
Accepted March 24, 2018 ■