CrossMark

The role of control processes in temporal and semantic contiguity

M. Karl Healey[1] · Mitchell G. Uitvlugt[1]

Abstract

Atkinson and Shiffrin (1968) argued that performance on any memory task reflects the combined influence of both the fixed structure of the memory system *and* control processes tailored to the specific task. We investigated the role of control processes in governing the temporal contiguity and semantic contiguity effects in free recall—tendencies to organize recall based on proximity in the study list and pre-existing semantic associations. Subjects studied lists that contained four "clusters", each composed of four semantically associated words but presented in random order such that associates were not in adjacent serial positions. Subjects were given either standard free-recall instructions, instructions to focus on order-based associations (i.e., the original order of presentation), or meaning-based associations (i.e., pre-existing semantic relationships). Replicating previous work, lists with strong semantic relationships resulted in a reduced overall temporal contiguity effect when recalled under standard free-recall instructions. However, under meaning-based recall instructions, the temporal contiguity effect was nearly eliminated. Detailed analyses of within-cluster transitions and an order reconstruction task revealed that temporal information was encoded, but control processes prevented it from dominating memory search. These results point to a need for more empirical work exploring how control processes change recall dynamics and for more theoretical work modeling the computational basis of these processes.

Keywords Episodic memory · Free recall · Temporal contiguity

Introduction

Atkinson and Shiffrin (1968) opened their seminal chapter by drawing a distinction between *memory structure* and *control processes*. They defined memory structure as "the permanent features of memory, ... [which] include both the physical system and the built-in processes that are unvarying and fixed from one situation to another" (p. 90). They contrasted memory structure with control processes, which "are selected, constructed, and used at the option of the subject and may vary dramatically from one task to another even though superficially the tasks may appear very similar" (p. 90).

Both memory structure and control processes are critical to a full understanding of the memory system, but some areas of theory development have focused more on one than the other. In our own area of work, the dynamics of episodic memory search, much of the focus has been on structure (as noted by Lehman & Malmberg, 2013). This includes both the physical system (e.g., whether there are separate short- and long-term stores; Davelaar et al., 2005; Sederberg et al., 2008) and built in processes (e.g., how mental context drifts during encoding and retrieval; Manning et al., 2011; Howard & Kahana, 2002a). These lines of work have been enormously profitable, but we take the anniversary of Atkinson and Shiffrin (1968) as an opportunity to step back and consider their admonition that "control processes are such a pervasive and integral component of human memory that a theory which hopes to achieve a degree of generality must take them into account" (p. 191).

Much of the recent focus on structure has been driven by the *temporal contiguity effect* (TCE). The TCE is the finding that recalling one event, i, tends to cue recall of another event, j, that was presented nearby in the list such that the lag in serial positions, the distance between i and j, is small (Kahana, 1996). The TCE can be readily observed in many laboratory list-learning tasks and even in more naturalistic settings (for a review, see Healey et al., in press). The ubiquity of the TCE has led us to suggest it may arise from processes that are automatic, obligatory, and built into the structure of the memory system (Healey & Kahana, 2014;

✉ M. Karl Healey
 khealey@msu.edu

1   Department of Psychology, Michigan State University,
     316 Physics Road, East Lansing, MI, USA

Healey et al., in press; Uitvlugt & Healey, 2019; Healey, 2018).

This presumed automaticity is reflected in many theoretical accounts of the TCE. For example, under retrieved context models of the TCE (Howard & Kahana, 2002a; Sederberg et al., 2008; Polyn et al., 2009), associations between items and a time-correlated mental context representation form during encoding, and then memory search follows a more or less ballistic trajectory in which one successful recall automatically reinstates its context which in turn automatically cues the next recall (however, see Lohnas et al., 2015; Healey & Kahana, 2016). This is in contrast to models that have descended directly from Atkinson and Shiffrin (1968) which incorporate many control processes to regulate memory encoding and search (e.g., rehearsal, compartmentalization, cue selection, stopping rules, rechecking; Raaijmakers & Shiffrin, 1981; Shiffrin & Steyvers, 1997; Lehman & Malmberg, 2013).

Recent evidence suggests that such control processes are vitally important in determining the magnitude of the TCE. One obvious signature of control processes is the finding that the magnitude of the contiguity effect in free recall increases substantially as subjects accumulate experience with the task (Healey et al., in press). A more dramatic demonstration comes from eliminating the impetus for the subject to engage control processes during encoding. When subjects are asked to process a list of words under some cover task and then receive a surprise free-recall test, the TCE is dramatically reduced (Nairne et al., 2017; Healey, 2018), though not quite eliminated (Healey, 2018). These initial demonstrations show that we should seriously consider the possibility that the TCE may depend as much on task-specific control processes as on the fixed structure of memory (Hintzman, 2016).

Here, we will try to establish some boundary conditions on the degree to which control processes can modulate the TCE. We will do so by manipulating two variables that Atkinson and Shiffrin (1968) identified as likely to influence which control processes are engaged: the nature of the instructions and the meaningfulness of the material.

Task instructions

Task instructions are obviously critical. For example, free recall and serial recall are identical tasks save for a small change in instructions. It has long been established that this small change in instructions produces very different serial position curves (SPC; Raffel, 1936; Jahnke, 1965; Anderson & Crosland, 1933), even when the instructions are not given until after the list has been studied (Deese, 1957). Subtler changes of instruction also influence the size of the primacy and recency effects, such as asking subjects to initiate recall from different parts of the list (e.g., the

beginning, the middle, the end) and then allowing free recall after this constrained-initiation (e.g., Craik, 1969; Murdock, 1968; Dalezman, 1976; Katz, 1968).

Instructions influence primacy and recency, but what about the TCE? Unsurprisingly, serial recall instructions produce a much larger TCE than free-recall instructions (Klein et al., 2005). This change in behavior no doubt reflects the influence of control processes. Nonetheless, many other aspects of behavior on the two tasks are quite similar, suggesting that differing control processes rely on the same underlying structural features in both tasks (Ward et al., 2010; Grenfell-Essam & Ward, 2012). That is, performance reflects an interplay of fixed structure and variable control processes. The extent to which the TCE can be modulated by this interplay remains unclear. For example, if asking subjects to focus on serial order increases the TCE, does asking them to *ignore* serial order decrease the TCE?

Atkinson and Shiffrin (1968) suggested that the focus on temporal-order information in free recall was due to a control process. They argued that because the long-term store is too large to search at random, subjects are likely to try searching along some ordered dimension. The Search of Associative Memory (SAM) theory (Raaijmakers & Shiffrin, 1980, 1981) extended these ideas, arguing that the subject can, based on task instructions, create a retrieval plan that specifies how memory will be searched, including what cues will be used (e.g., temporal-contextual information, item information). SAM also allows different cues to have different weights set via control processes. Indeed, the notion that subjects can selectively weight different dimensions is common in many theories. For example, attentional weights to different dimensions in the SIMPLE model (Brown et al., 2007), "retrieval orientation" in theories of the neural-basis of retrieval (Rugg & Wilding, 2000), and "dimension weighting" in visual search (Memelink & Hommel, 2013). Task instructions could thus lead subjects to weight different information more heavily, perhaps during both encoding and recall. Typical free-recall instructions may encourage the subject to weight temporal-order information particularly heavily because it allows the subject to discriminate items that were on the most recent list from items that were on earlier lists. In other words, searching memory by temporal order could be the result of a task-specific control process. In that case, it may be possible to totally eliminate the TCE simply by asking subjects to focus on something other than temporal-order information during search.

We will test this hypothesis with a subtle manipulation of task instructions. All subjects will be told about two different ways of searching memory: along a temporal-order dimension by allowing one recalled item to remind them of other items studied nearby in the list, or along a semantic

dimension by allowing one recalled item to remind them of other items that were meaningfully related. Then, subjects will diverge into three different conditions. Subjects in a free-recall condition will be asked to recall the items in whatever order they wish. Subjects in an order-focus condition will be asked to focus on the order dimension (essentially serial recall that relaxes the requirement of perfect ordering). Finally, subjects in a meaning-focus condition will be asked to focus on the semantic dimension.

List structure

One challenge is that subjects might prefer the order dimension regardless of instructions because it provides a more reliable cue. For every list item there is always another item that occurred in an adjacent serial position; however, every item *does not* necessarily have a strong semantic associate on the list. To make the semantic dimension more appealing to subjects, we will also manipulate the inherent degree of semantic organization in the lists. Raaijmakers and Shiffrin (1980) showed that many aspects of the influence of semantic information on recall can be explained by the role of control processes in selecting cues. Specifically, they were able to fit a variety of key effects from the literature on recall of categorized lists simply by assuming that when recalling categorized lists, subjects include category information along with context information in the retrieval cue. Use of such semantic cues could reduce the reliance on temporal-order information and attenuate the TCE.

Although there is a large literature on how semantic associations among words influence recall order (Bousfield & Sedgewick, 1944; Tulving, 1962; Jenkins & Russell, 1952; Bousfield, 1953; Pollio et al., 1969), there is surprisingly little work on how such meaning-based associations influence the TCE. The primary exceptions are experiments that have added groups of related words to a list. At one extreme, the entire list can be drawn from a single semantic category such that all items have strong relations to each other. Nairne et al. (1991) showed that categorized lists reduced forward serial recall as measured via input-output correspondence (Asch & Ebenholtz, 1962). Later work using methods that consider both forward and backward transitions and control for which items are available found a robust TCE in single-category lists (Miller et al., 2013; McCluey et al., 2018). At the other extreme, an otherwise random list can include a few pairs of related items separated by some serial position lag. In this case, semantic associations increase the TCE if the related pairs were presented close together in the list but decrease the TCE if the pairs were presented far apart in the list (Healey et al., in press; see also, Howard & Kahana, 2002b).

The most direct evidence comes from Polyn et al. (2011) who found that when lists were composed of eight items

drawn from each of three categories presented in a random order, the TCE *seemed* to be reduced. However, when analyses were restricted to cases where subjects transitioned between items from the same category, it revealed a preference to transition between category members that were presented nearby in the list. This was true both when subjects were free to recall in any order and when they were explicitly given a category name as a recall cue (cf. Tulving and Pearlstone, 1966). Similarly, when they transitioned from one category to another during free recall, they preferred categories that were presented nearby in the list. That is, even when subjects had strong semantic associations to rely on, input order still powerfully influenced recall order.

The Polyn et al. (2011) list design is thus a useful starting point for testing the limits of control processes: if a subject is asked to focus on semantic information and ignore order information, can the TCE be eliminated? One limitation of their study is that it reported only the TCE and not corresponding measures of the *semantic contiguity effect* (SCE), and therefore, it did not fully delineate how adding semantic structure to a list affects the balance between semantic and temporal organization of memory search. Thus, we will conceptually replicate the semantic structure manipulation of Polyn et al. (2011) but use latent semantic analysis (Landauer & Dumais, 1997) to create semantic structure rather than categories, allowing us to simultaneously measure both the TCE and the SCE. By combining this manipulation of semantic structure with a manipulation of task instructions, we can investigate the role of control processes in setting the balance between order and semantic information during memory search.

Methods

Data sharing

All data analyzed in this report are freely available on the author's website (https://cbcc.psy.msu.edu/data/Heal18ctrl.csv).

Subjects

Power calculations reveal that a sample size of 9 per condition provides a $1 - \beta$ power of 0.95 to detect (via a two-tailed one-sample t test) a TCE of the size reported in a meta-analysis (Sederberg et al., 2010). To ensure we had sufficient power to detect even smaller effects, we set a target sample size of 70 subjects per condition.

A total of 439 Michigan State University undergraduate students signed up to participate in the experiment in exchange for course credit. Of these, there were 70

males, 363 females, one transgender individual, and five individuals who preferred not to answer. There were 422 native English speakers, 15 non-native English speakers, and two individuals who preferred not to answer. The mean age was 19.33 (SD = 2.22); two individuals declined to provide their age. Of the original 439 subjects who signed up, three were not run due to technical difficulties.

The remaining 436 subjects were randomly assigned to one of six distinct conditions in a 2 (list structure: random lists, related lists) × 3 (instruction: free-recall, order-focus, meaning-focus) between-subjects design. Table 1 shows sample sizes for each condition. Two of the subjects (one from the related lists/order-focus condition; one from the random lists/meaning-focus condition) completed the majority of the free-recall trials but abandoned the study before starting the order reconstruction task—all of their available data is included in the analyses.

Procedure

Free-recall task

To begin, subjects were instructed to memorize lists of words for a free-recall test. Each list was composed of 16 words, with the constraint that no word was used more than once for a given subject. During the study phase, words were presented one at a time on a computer screen for 2 s each with a jittered (i.e., variable) interstimulus interval of 400–1200 ms (uniform distribution).

Following the final word presentation and a jittered 300–700 ms (uniform distribution) pause, the recall phase began. Subjects were given 60 s to recall as many words as possible from the preceding list by typing in each recalled word and pressing enter to submit their responses. A spell-checking algorithm corrected any typos in subjects' responses and scored their recall accuracy (for a detailed description of the algorithm, see Healey, 2018). Subjects completed 18 trials of the free-recall task.

List structure manipulation

Two between-subject conditions varied in the semantic relatedness of the words within a list:

Related lists In the related lists condition, each word list contained four clusters which were each composed of four semantically-related words. Unlike most previous work, our lists were not created by drawing from distinct semantic categories (hence our use of the term "cluster" rather than "category"). Instead, semantic relatedness was manipulated using latent semantic analysis (LSA; Landauer & Dumais, 1997) which measures similarity of words as the cosine of the angle between their high-dimensional vector representations. The lists were constructed by randomly sampling words from the pool of 1638 words that we have used in past free-recall studies (Healey & Kahana, 2014, 2016; Healey, 2018) according to two constraints. First, each possible pairing of words within the same cluster had to have a $\cos(\theta) > 0.48$. Second, each possible pairing of words from different clusters had to have a $\cos(\theta) < 0.20$. Together, these constraints ensured that words within a cluster were highly related to one another while words from different clusters were relatively unrelated. Note that the generated lists used only 711 unique words from the original 1638-word pool. As an example, one semantically-related word list contained the following clusters: *video, computer, keyboard, calculator*; *mouth, tongue, head, nose*; *palace, king, throne, queen*; and *potato, broccoli, spinach, fruit*.

At the time of presentation, the order of words in a list was randomly shuffled with the constraint that no members of the same cluster were presented in adjacent serial positions. This ensured that strict serial order versus strict semantic retrieval strategies would produce different recall orders. In other words, related lists were composed of groups of words that are close together (i.e., clustered) in the LSA semantic space but randomly distributed across serial positions.

Table 1 Sample size, overall recall, accuracy on the order reconstruction, and PFR-scores by condition

List structure	Instructions	n	Overall recall (SD)	Order acc. (SD)	PFR-score (SD)
Related lists					
	Free-recall	74	.543 (.113)	.158 (.075)	2.571 (3.139)
	Meaning-focus	72	.555 (.124)	.144 (.072)	2.819 (2.765)
	Order-focus	72	.470 (.114)	.178 (.109)	-2.911 (4.435)
Random lists					
	Free-recall	73	.477 (.118)	.181 (.091)	2.706 (2.989)
	Meaning-focus	72	.492 (.123)	.184 (.102)	2.850 (2.729)
	Order-focus	73	.450 (.120)	.205 (.125)	-2.169 (4.298)

We created enough 18-list stimuli sets so that within each of the three instruction conditions (described below) each subject saw a unique stimuli set, but the same sets were used in all three instruction conditions. We did this to ensure that differences across instruction conditions could not be explained by differences in stimuli.

Random lists In the random lists condition, lists were created by randomly selecting words with no constraints except for restricting the word pool to the 711 words that were used for the related lists condition. Thus, these lists are comparable to the randomly generated lists used in many previous free-recall studies, in that they may contain semantic associates by chance but have no consistent semantic structure. As with the related lists condition, within each instruction condition each subject saw a unique 18-list stimuli set, but the same sets were repeated across instruction conditions.

Instruction manipulation

All subjects were told that both order-based and meaning-based information can help guide memory search. Specifically, they were told: "Memory search is helped by the order in which you memorized the words. Remembering one word might remind you of the very next word in the list", and "Memory search is also helped by meaning—that is, the definition of the word. Remembering one word might remind you of a word that is semantically (meaningfully) related". They were given examples to illustrate each of these approaches. After this common introduction, the remaining instructions were manipulated to encourage subjects to guide their memory searches in different manners. There were three instruction conditions: free-recall, order-focus, and meaning-focus.

Free-recall Subjects were instructed to use either method to help search memory. Specifically, subjects were told: "When remembering the words, you are free to use either the order of the words or the meaning to help you remember. After the list, you will be prompted to type in any words that you can remember from the list. When prompted, type in whatever word comes to mind first, then type in the next word that comes to mind".

Order-focus Subjects were instructed to use the original order of the list to guide recall. Specifically, subjects were told: "When remembering the words, try to use only the original order of the list to guide your memory search. When one word comes to mind, try to think of other words that were presented nearby it in the list. Completely ignore any semantic relationships. After the list, you will be prompted to type in any words that you can remember from the list.

When prompted, type in whatever word comes to mind first, then allow the order in which you saw the words to guide your memory search".

Meaning-focus Subjects were instructed to use the semantic associations of items on the lists to guide recall. Specifically, subjects were told: "When remembering the words, try to use only the semantic associations of items on the list to guide your memory search. When one word comes to mind, try to think of other words that were related to it in the list. Completely ignore the order in which you saw the words. After the list, you will be prompted to type in any words that you can remember from the list. When prompted, type in whatever word comes to mind first, then allow the semantic associations to guide your memory search".

Order reconstruction task

After recalling the final list, subjects completed a surprise order reconstruction task. Here, subjects were given all of the words from one of the previously studied lists and were asked to reconstruct the list's original serial order from the study phase (for similar procedures, see Nairne, 1991, 1990).

All the words of a list were presented on screen simultaneously in a numbered list; however, the words were in a constrained randomized order such that at most one pair of words that were adjacent in the study list could be adjacent in the randomized list and that the first item of the original list was not in the first position of the randomized list. The words were numbered 1–16. To reconstruct the study list, subjects began by typing the number of the word they thought was presented first, then the number of the word they thought was presented second, and so on. As subjects submitted their responses, the word list they were creating was presented on screen. Words could only be submitted once. If an error was made, subjects could reset their created list and start over.

There was a total of four order reconstruction trials. The first trial tested the most recently studied list (i.e., the 18th list) to take advantage of the recency effect and maximize the chance of detecting any retention of order information. Then, subjects reconstructed the 1st, 6th, and 12th lists in that order.

Quantifying the temporal contiguity and semantic contiguity effects

The TCE is generally illustrated by examining *conditional-response probability* as a function of lag (lag-CRP). The lag-CRP gives the probability that recall of an item studied in position i of a study list will be followed by recall of an item studied in position $i + lag$. For example, if recall of the

item from position 5 was followed by recall of the item from position 6, then $lag = +1$. If, however, it was followed by recall of the item from position 3, then $lag = -2$. For each lag, the CRP is computed by dividing the number of times a transition of that lag was *actually* made by the number of times it *could* have been made (e.g., it could not have been made if the item $i + lag$ was already recalled; Kahana, 1996).

The SCE can be illustrated by measuring *conditional-response probability* as a function of semantic similarity (semantic-CRP) instead of lag. Semantic-CRPs are usually calculated by taking the distribution of $\cos(\theta)$ similarity values between all possible word pairs and dividing it into some number of bins, and then for each bin, the number of transitions that actually corresponded to that bin and the number of transitions that could have potentially corresponded to that bin is computed. This method is problematic with the current design because average $\cos(\theta)$ values will naturally be higher in the related lists condition than in the random lists condition, which makes comparisons across conditions difficult.

Therefore, we will compute the semantic-CRP in a slightly different way. Each time a participant recalls a word, we will find the semantic similarity between that recalled word and all of the words that could have been transitioned to. We will then order these possible transitions from most to least similar, essentially creating a *relative semantic lag*, such that transitioning to the most similar item would be a $lag = 1$ transition, transitioning to the second most similar would be a $lag = 2$ transition, and so on. We will consider semantic lags 1–5 (there are, of course, no negative semantic lags). Then, as with the lag-CRP, we will divide the number of times a subject *actually* made a transition of each semantic lag by the number of times they *could* have made transitions of that semantic lag. In calculating semantic similarity for the analyses reported below, we used LSA. We note that almost identical results are obtained when the same analyses are conducted using Word Association Space (WAS; Steyvers et al., 2004) $\cos(\theta)$ values which some evidence suggests better predicts subjects' transitions than LSA (Morton & Polyn, 2016).

The lag- and semantic-CRPs provide a visual representation of the TCE and SCE, but it is useful to have a single number that quantifies the size of the effects. For this purpose, *factor scores* are typically used (Sederberg et al., 2010; Polyn et al., 2009). The temporal factor score is computed by ranking the absolute value of the lag of each actual transition with respect to the absolute values of the lags of all transitions that were possible at that time, which provides a percentile score for each transition. Averaging these percentile scores across all of a subject's transitions provides the temporal factor score. A semantic factor score is computed in exactly the same way except the distribution of

possible semantic lags is used instead of temporal lags. For both the temporal and semantic factors scores, a value of 0.5 indicates that transitions are random with respect to temporal/semantic lag, and values above 0.5 indicate a TCE/SCE. To control for the fact that the size of the TCE can be influenced by the shape and level of the SPC, we also conducted a shuffling procedure by computing factor scores for both the actual recall sequence and for many random permutations of the sequence and then testing whether the actual scores falls in the most extreme 2.5% of the positive tail of the shuffle distribution (Healey, 2018; Polyn et al., 2011).

Data analysis

We will rely primarily on confidence intervals to assess whether the TCE is significant within a condition and whether it differs among conditions. Bootstrapped 95% confidence intervals (CI) for a given condition with n subjects were computed by randomly sampling n subjects from that condition with replacement (i.e., the same subject could be sampled more than once) and computing the mean of that random sample. This resampling procedure was repeated 1000 times, providing a distribution of sample means. The 95% CI has endpoints at the 2.5 and 97.5 percentiles of the sampling distribution (Efron & Tibshirani, 1993). For measures where each subject provides multiple data points (e.g., the lag-CRP), resampling was done at the level of subjects, such that if a subject was sampled, all their data for that measure was selected (i.e., pairings of the rows and columns of the subject $\times$ data point matrix were maintained). To supplement the CIs, the Appendix reports traditional significance tests as well as measures of effect size for key effects.

Results

How do control processes change recall initiation and recall accuracy?

We begin by examining how the instruction manipulation influenced some basic measures of recall dynamics: the probability of first recall curve (PFR) and the SPC (See Table 1 for overall recall probabilities). PFRs reveal how subjects begin searching memory by showing the probability of starting recall at each serial position. The PFR in immediate free recall tends to show a strong recency effect and a more modest primacy effect (Hogan, 1975; Howard & Kahana, 1999). SPCs reveal the overall probability of successfully recalling an item as a function of serial position and also typically show both a primacy and a recency effect (Murdock, 1962). We will first consider how the instruction manipulation influenced these measures in
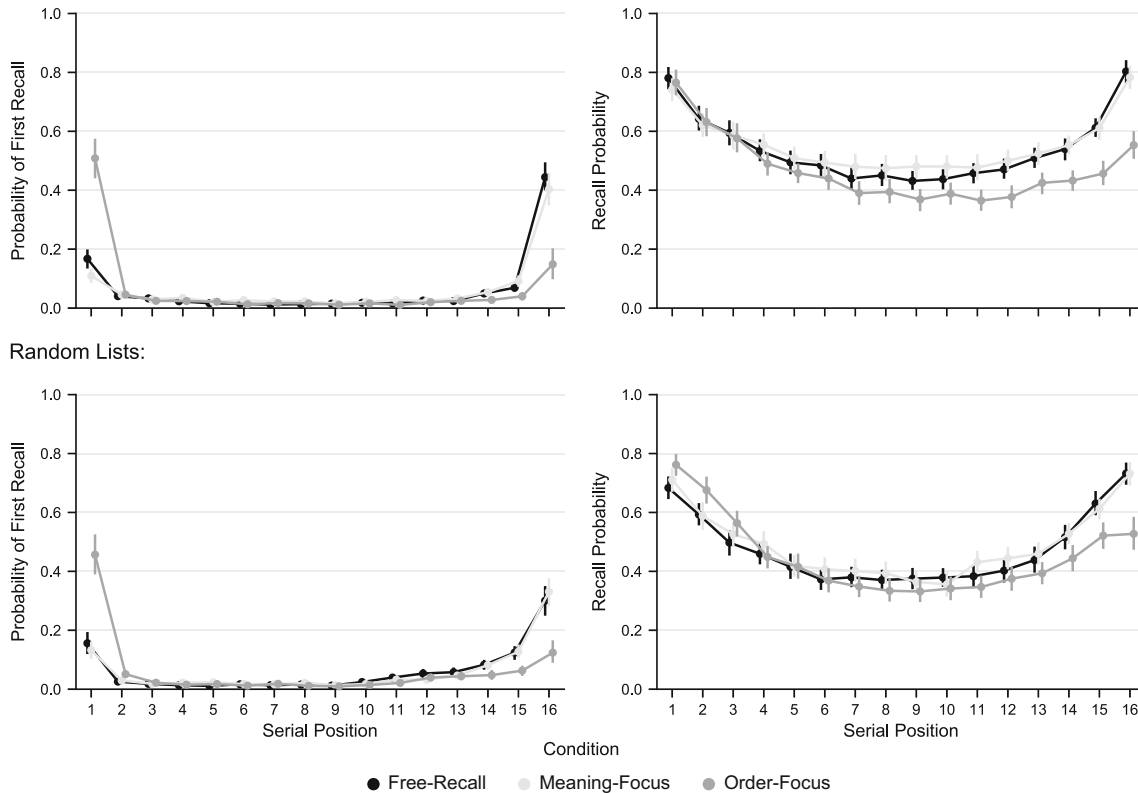
Related Lists:

Random Lists:

Fig. 1 Probability of first recall curves (*left panels*) and serial position curves (*right panels*) by instruction condition for related lists (*top row*) and random lists (*bottom row*). All *error bars* are bootstrapped 95% confidence intervals

the related lists condition (Fig. 1, top row). In the PFR, the order-focus instructions lead to a stronger primacy effect (comparing initiation probabilities across conditions for serial position 1) and a reduced recency effect (comparing initiation probabilities across conditions for serial position 16) relative to the free-recall control condition.[1] The SPC reveals a similar effect on recency, though not on primacy. By contrast, the meaning-focus instructions had little effect on either the PFR or the SPC. Turning to the random lists condition (Fig. 1, bottom row), we see a very similar pattern. The fact that order-focus instructions increased recency and reduced primacy indicates that the manipulation successfully biased subjects toward searching memory along an order-based dimension.

The shape of the SPCs warrants some discussion. The SPCs in Fig. 1 show roughly equal primacy and recency effects. Although other studies of immediate free recall have occasionally found roughly balanced

SPCs, (e.g., Ward et al., 2003; Ward & Maylor, 2005; Unsworth et al., 2011; Bousfield et al., 1958; Jahnke, 1965), the recency effect usually tends to be larger than the primacy effect (Murdock, 1962; Healey & Kahana, 2014). Our smaller recency effect may arise from several factors. First, it may reflect level of practice. Classical studies such as Murdock (1962) used highly practiced subjects who have completed multiple sessions of the task, and it is known that with practice subjects tend to shift from a primacy to a recency strategy (Unsworth et al., 2011; Dallett, 1963; Hasher, 1973; Goodwin, 1976). Our own subjects completed only a single session. We return to the issue of primacy versus recency strategies below when we consider individual differences and show that a subset of subjects who favor primacy are indeed attenuating the average recency effect. Second, when lists contain sequential dependencies (Deese & Kaufman, 1957) or semantic structure (Polyn et al., 2011), recency tends to be less dominant. Thus, by emphasizing order and semantics, the instructions in the current study may have discouraged subjects from adopting a strategy of maintaining the last few items in short-term memory and immediately outputting them at the beginning of the recall period.

---

[1]In the main text we rely on confidence intervals to judge the reliability of effects. In the Appendix, we provide ANOVAs and *t* tests for key comparisons.

How do control processes change the role of order and semantic information during memory search?

Contiguity in related lists

Looking first at the lag-CRPs in the related lists condition (Fig. 2, top row), we see that whereas the order-focus instructions produced a familiar lag-CRP peaked at $|lag| = 1$ with a forward asymmetry, the meaning-focus instructions produced a flattened lag-CRP. Remarkably, even the free-recall instructions produced a flattened lag-CRP—simply including semantic structure in the list (and mentioning order-based and meaning-based search in the instructions) is enough to prompt subjects to engage control processes to focus on the semantic dimension (Polyn et al., 2011). The temporal factor scores quantify the magnitude of the TCE: the temporal factor score was well above 0.5 under order-focus instructions, was marginally above 0.5 under free-recall instructions, and not significantly different than 0.5 under meaning-focus

instructions. However, all of these factor scores, including the meaning-focus score, were significantly higher than the expected value from the shuffling procedure described above. In other words, meaning-focus instruction eliminated, or nearly eliminated, the TCE depending on how significance is assessed. The semantic-CRPs and semantic factor scores mirror this pattern: the SCE was highest under meaning-focus instructions, at an intermediate level under free-recall, and lowest under order-focus instructions. However, unlike the TCE, the SCE was significantly above both 0.5 and the mean of the shuffled distribution in all conditions.

It is noteworthy that the flattening of the lag-CRP under meaning-focus instruction in the related lists condition was largely due to a dramatic reduction in the number of $lag = +1$ transitions and to a lesser extent $lag = -1$ transitions. As a result, the lag-CRP in this condition has a distinct non-monotonic shape. In interpreting this finding, it is important to bear in mind that the related lists were constructed such that no items from the same cluster were presented
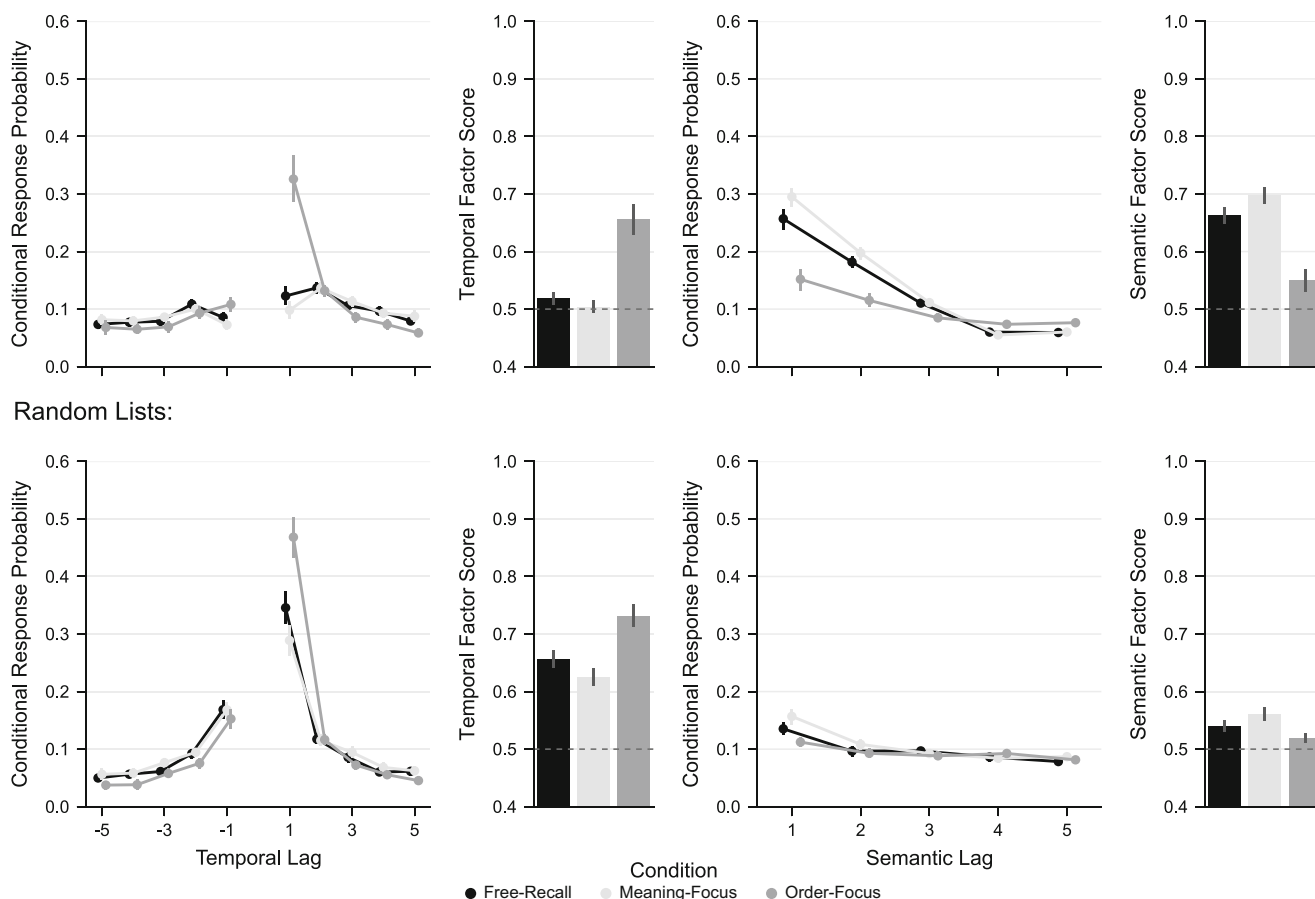
Related Lists:

Random Lists:

Fig. 2 Lag-conditional response probability functions along with temporal factor scores (*two leftmost panels*) and semantic-conditional response probability functions along with semantic factor scores (*two rightmost panels*) by instruction condition for related lists (*top row*) and random lists (*bottom row*). These analyses are based on all transitions. All *error bars* are bootstrapped 95% confidence intervals

in adjacent serial positions. Thus, it is expected that any subject relying on semantic associations would show a disproportionate decrease in $|lag| = 1$ transitions. It is also possible that demand characteristics discouraged subjects given meaning-focus instructions from making obviously near-lag transitions. While this likely plays some role, the fact that an almost identical pattern is seen in the free-recall condition suggests that the non-monotonicity is mainly due to relying on semantic associations.

Contiguity in random lists

Turning to the random lists in the bottom row of Fig. 2, we see task instructions had a similar but less dramatic effect even when the list did not contain a clear semantic structure. In all cases, the lag-CRPs were peaked with forward asymmetry and the temporal factor scores were above both 0.5 and the mean of the shuffled distribution. However, temporal contiguity was highest under order-focus instructions and significantly lower under meaning-focus instructions. Again, the semantic-CRPs and semantic factor scores mirrored this pattern. These results demonstrate that control processes modulate the size of both the TCE and the SCE even when lists lack rich semantic associations.

The finding that the related lists under meaning-focus instructions produced a small but significant TCE and a large, robust SCE indicates that most transitions were guided by semantic associations. That is, subjects were able to use control processes to search memory along a semantic dimension and largely ignore the order dimension. However, the fact that memory *search* was not dominated by input order does not mean that input order was ignored during *encoding*. It is possible that order information was encoded and available for use when semantic associations are less helpful. Therefore, we conducted two additional sets of analyses to test for the use of order information. In the first, we subset transitions into those between words in the same semantic cluster and those between words from different clusters (Polyn et al., 2011). In the second, we consider the surprise order reconstruction task.

Does the TCE reemerge when transitioning within and between clusters?

The foregoing analyses show that when recalling related lists under meaning-focus instructions, subjects were not using order information to guide them directly from one recalled item to another item studied nearby in the list. Instead, they were using semantic information to guide them to items nearby in semantic space—items from the same semantic cluster. But it could be that they *are* using order information to decide which same-cluster items they will

recall next. Following Polyn et al. (2011), we can isolate transitions made between items from the same cluster and define a within-cluster lag, such that $lag_{within} = 1$ denotes transitioning between two same-cluster items that were presented in succession ignoring any intervening different-cluster items. For example, if the study list presented the items $king$, $keyboard$, $fruit$, $queen$, transitioning from $king$ to $queen$ would be a $lag_{within} = 1$ transition even though $keyboard$ and $fruit$ separated them in the study list. Polyn et al. (2011) found that these within-cluster transitions showed a robust TCE. Does this within-cluster TCE still occur when subjects are explicitly asked to focus on semantics?

The top row of Fig. 3 shows the results. Note that because each cluster was composed of four items, the only within-cluster lags that are possible are of absolute values 1–3. The lag-CRPs show a peak at $lag_{within} = 1$ with a clear forward asymmetry. And the temporal factor scores were above both 0.5 and the mean of the shuffled distribution under all instruction conditions, though it was still significantly higher under order-focus than free-recall or meaning-focus instructions. The semantic-CRPs and semantic factor scores show a significant within-cluster SCE—subjects still used semantic information when making transitions within a cluster. In sum, within-cluster transitions reveal that although control processes allow subjects to focus primarily on semantic associations, they still have access to order information and use it during memory search.

A similar analysis can be conducted on between-cluster transitions (i.e., transitions made between items from different semantic clusters). Here, we might expect order information to be particularly important because the lists were constructed to minimize semantic associations among items in different clusters—that is, order information may provide the most reliable path from one semantic cluster to another. To conduct this analysis, we define between-cluster lag by considering the number of serial positions intervening between one item and all other items that are from different clusters, ignoring any intervening items from the same cluster. For example, if the study list included the following sequence of items $computer$, $nose$, $keyboard$, $fruit$, then transitioning from $computer$ to $fruit$ would be a $lag_{between} = 2$ transition even though the same-cluster item $keyboard$ intervened between them.

The bottom row of Fig. 3 shows the results. A robust TCE was observed under all instruction conditions, though again temporal factor scores were significantly higher under order-focus than free-recall or meaning-focus instructions. The semantic-CRPs were quite flat regardless of instruction condition, though the semantic factor scores reveal a small but significant SCE under all instructions except order-focus. In sum, between-cluster transitions suggest that when moving between two clusters that are widely separated in
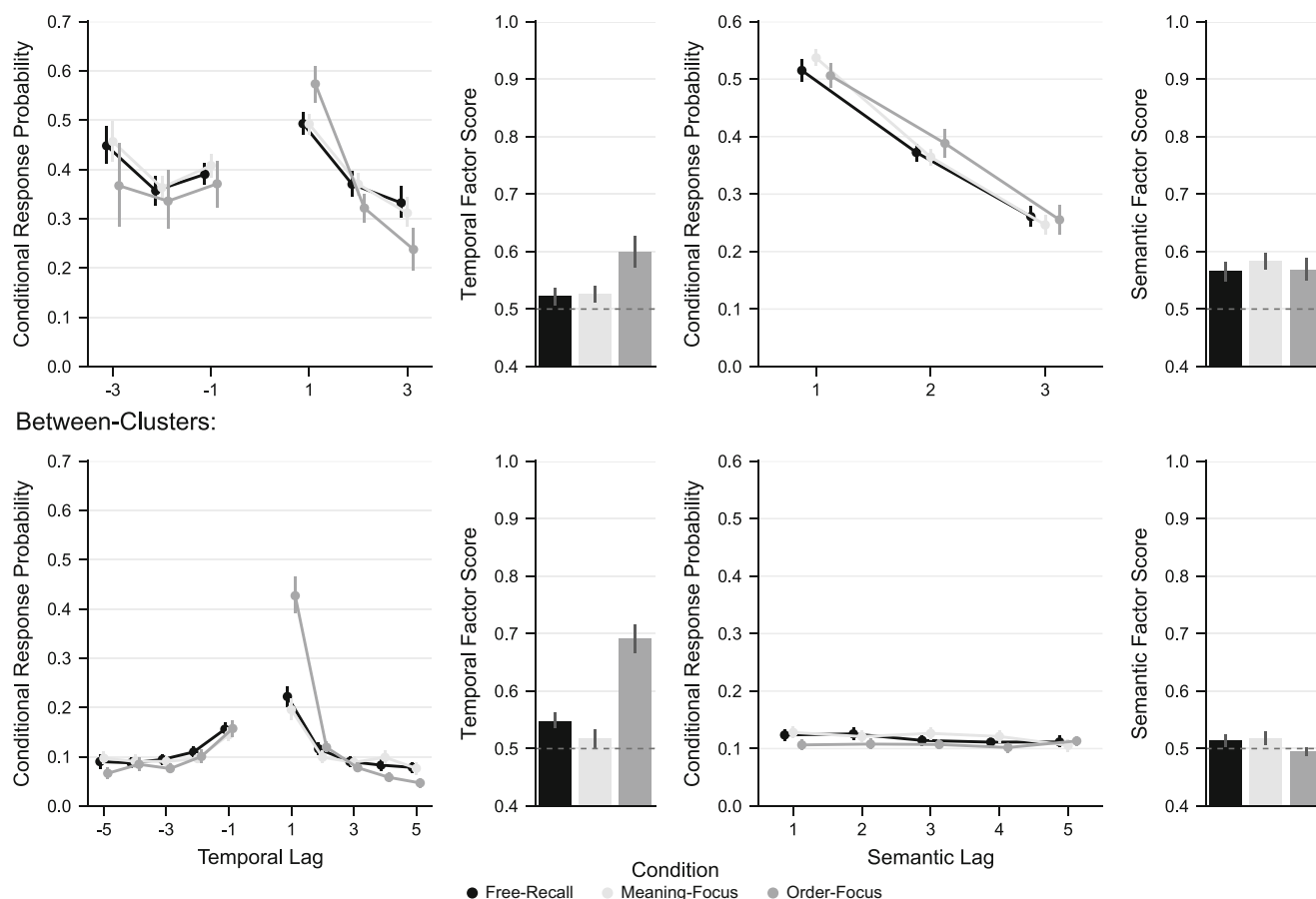
Within-Cluster:

Between-Clusters:

Fig. 3 Lag-conditional response probability functions along with temporal factor scores (*two leftmost panels*) and semantic-conditional response probability functions along with semantic factor scores (*two rightmost panels*) by instruction condition for transitions within a semantically related cluster (*top row*) and transitions between different clusters (*bottom row*). These data are based only on the related lists. All *error bars* are bootstrapped 95% confidence intervals

semantic space, subjects tend to rely strongly on order information but are still able to use control processes to change the balance between temporal and semantic dimensions depending on task instructions.

Is global order information encoded?

So far, we have seen that when recalling related lists with instructions to focus on meaning, subjects rely primarily on semantic associations. That is, instructions can cause subjects to largely ignore input order when searching memory. Nonetheless, it is still possible that subjects encode information about input order, even though they do not use it to guide memory search. Indeed, analyses of the subset of transitions made between items in the same cluster show that subjects retain and use information about the order of items *within* a semantic cluster. Similarly, they retain and use information about the order of items at the borders

of semantic clusters as shown by the analysis of between-cluster transitions. However, do subjects retain information about the overall, global order of the list? If so, they should be able to deploy control processes to reconstruct the serial order of a list regardless of instruction condition or list composition. To test this prediction, we examine the data from the end-of-experiment surprise order reconstruction task. In this task, subjects were shown all the items for a previously studied list but in a scrambled order and were then asked to reconstruct the original order by assigning each item to a serial position.

We must approach the order reconstruction data with some caution because it was administered after subjects had already recalled the lists. That is, even though subjects are told to reconstruct the order in which they studied the items, they are likely to also be influenced by the order in which they recalled them. This makes it hard to compare reconstruction accuracy between conditions that differed in

output order. Therefore, we will use the reconstruction data to answer one limited question for each condition: is order reconstruction accuracy above chance. If it is, subjects in that condition must have formed some memory of input order.

We begin by looking at reconstruction accuracy as a function of serial position (Fig. 4; see Table 1 for overall accuracy). Although there are some small differences among the conditions, we do not consider them given the confound with original recall order. What is most striking is that accuracy was highest for early serial positions and dropped off for later serial positions. This contrasts with previous work on order reconstruction tasks which has generally found a bow-shaped curve with better accuracy for both early and late positions (e.g., Nairne, 1990; Healy, 1974). This difference may reflect the fact that whereas most previous work has asked subjects to reconstruct the order of each list before another

Related Lists:

Random Lists:

Fig. 4 Order reconstruction accuracy as a function of serial position by instruction condition for related lists (*top row*) and random lists (*bottom row*). All *error bars* are bootstrapped 95% confidence intervals

is presented, our reconstruction task came at the end of the experiment after many lists had been previously recalled. It is possible that this introduced a negative recency effect similar to that observed in final free recall (Craik, 1970). Second, whereas past work has allowed subjects to assign items to serial position in any order (e.g., they could have assigned an item to the final serial position first), in our task subjects were required to decide which item should be in the 1st position before moving on to the 2nd and so on, which could have produced output interference. Nonetheless, the data allow us to ask whether reconstruction accuracy was above chance.

To test whether accuracy is above chance, we will analyze the reconstruction data in exactly the same way we have analyzed recall data: by computing CRPs and factor scores. For example, we can take the items that a subject assigned to serial positions 1 and 2, find their true serial positions, and calculate the lag between them. If the subject accurately reconstructed the order, then the lag will be +1, but if instead the item the subject assigned to serial position 2 was actually from serial position 5, then the lag will be +4. These are similar to the transposition gradients often used in the serial recall literature (Nairne, 1990; Healy, 1974; Solway et al., 2012).

The results for related lists are shown in the top row of Fig. 5. We see that the lag-CRP was peaked with a forward asymmetry and all the temporal factor scores were above both 0.5 and the mean of the shuffled distribution. The results for the random lists, shown in the bottom row of Fig. 5, reveal the same pattern, except the contiguity effect was somewhat larger. As we have noted above, the fact that order was reconstructed after subjects had already recalled the lists makes it difficult to interpret the magnitude of the TCE and differences among conditions. Nonetheless, the fact that order reconstruction was above chance in all conditions supports one unambiguous conclusion: it is clear that subjects have access to some information about input order regardless of list composition or task instructions. That is, control processes influenced how much order information was used during memory search but did not prevent if from being encoded.

Individual differences

Previous work has shown that individual differences in how subjects organize memory search, especially the tendency toward temporal clustering, predicts recall success (Healey et al., in press; Healey et al., 2014; Sederberg et al., 2010; Unsworth, 2009). However, we do not know whether these correlations arise because memory for order information is of fundamental importance when searching memory, or if it is because serial recall happens to be a good strategy when free-recalling random lists. Successful
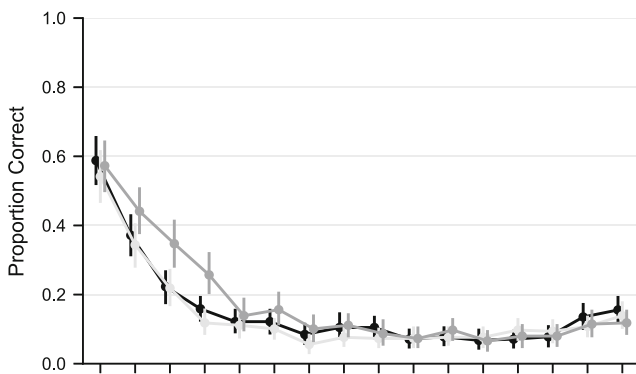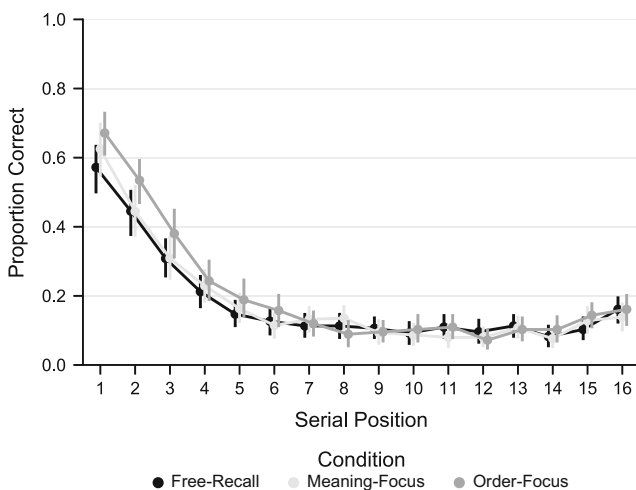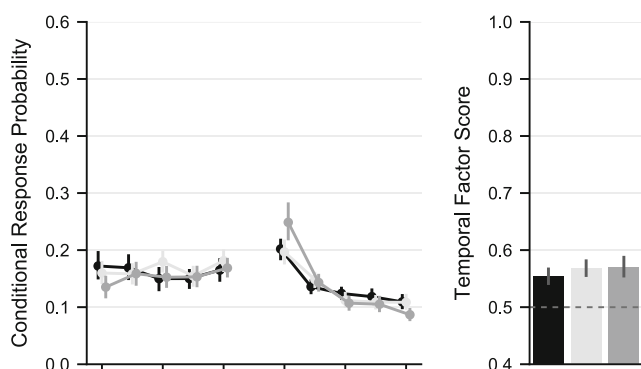
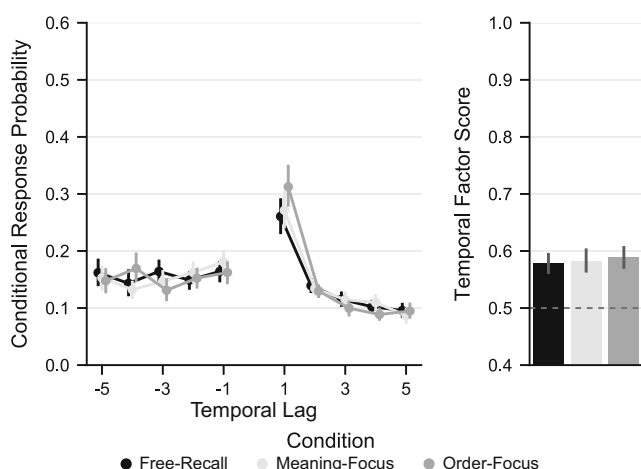Related Lists:

Random Lists:

Fig. 5 Lag-conditional response probability functions along with temporal factor scores by instruction condition for related lists (*top row*) and random lists (*bottom row*). These analyses are based on all transitions of the order reconstruction task. All *error bars* are bootstrapped 95% confidence intervals

memory search may depend on using control processes to focus on the most appropriate associative dimension. Here, we will examine how the pattern of correlations among measures of recall dynamics changes across conditions. We will consider four measures: Overall recall (i.e., the percent of words recalled), temporal factor scores, semantic factor scores, and a measure of how subjects initiate recall.

We have already discussed the first three measures but have only considered recall initiation with respect to the PFR curve. To provide a single-number summary of recall initiation for a given participant, we began by relabeling serial positions so they centered around zero such that negative values reflect a tendency toward primacy and positive values reflect a tendency toward recency. Specifically, we relabeled positions 1–8 as -8 – -1 and positions 9 – 16 as +1 – +8. We then used these new labels

to weight a participant's PFR curve and then summed across serial positions:

$$\text{PFR-score} = \sum_{i=1}^{ll}(PFR_i \times weight_i). \tag{1}$$

Where, ll is the list length (i.e., 16 items) and i indexes serial positions. For example, if a participant initiated recall at serial position 1 90% of the time and position 2 10% of the time, then their *PFR-score* would be $(.9 \times -8) + (.1 \times -7) = -7.9$, whereas a participant that initiated recall at position 16 90% of the time and position 15 10% of the time would get a PFR-score of $(.9 \times 8) + (.1 \times 7) = +7.9$. Table 1 shows these PFR-scores as a function of condition.

Consistent with previous work (Healey & Kahana, 2014; Unsworth et al., 2011), we found that the distribution of PFR-scores was somewhat bimodal: although most subjects tend to initiate recall from recency positions, a substantial proportion favor a primacy initiation pattern. Figure 6 shows SPCs, PFRs, CRPs, factor scores separately for subjects with PFR-score $\geq$ 3 and $\leq$ -3. The general patterns are the same as when all subjects are considered; however, subjects who show a recency initiation pattern have SPCs much closer to the classic recency heavy pattern (cf. Unsworth et al., 2011). We turn now to correlations among the measures.

Figure 7 shows the correlations among each pair of measures broken down by condition. The figure caption provides guidelines for assessing the significance of individual correlations as well as differences between correlations. However, because the number of correlations presents a multiple comparison problem, we will focus on a few theoretically motivated comparisons.

We focus first on correlations with overall recall. Overall recall was not correlated with PFR-scores in any condition (Fig. 7a). This replicates our previous finding that although recall initiation shows substantial inter-individual variation, this variation is unrelated to overall recall success (Healey et al., 2014; but see, Unsworth et al., 2011). Overall recall *was* correlated with temporal factor scores (Fig. 7b), but the direction of that correlation varied across conditions. For random lists under free-recall instructions, the two were positively correlated, replicating previous work with standard free recall (Healey et al., in press; Healey et al., 2014; Sederberg et al., 2010; Unsworth, 2009). For related lists, however, the two were *negatively* correlated in the free-recall and meaning-focus conditions. This suggests that it is not so much the ability to reconstruct temporal order that is critical in determining recall success but rather the ability to focus on whichever associative dimension provides the strongest retrieval cues.

The correlations among overall recall and semantic factor scores tell a similar story (Fig. 7c). For related
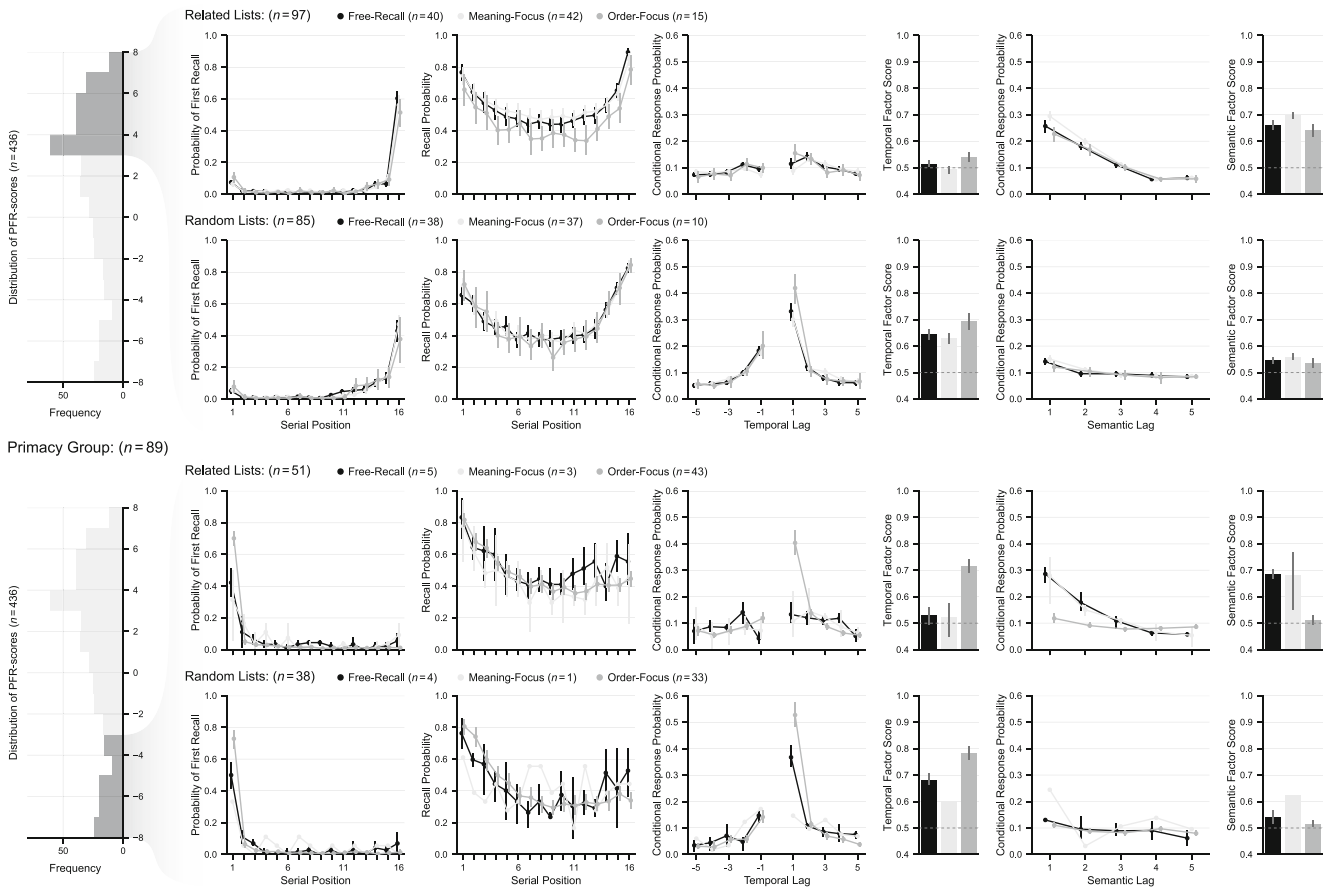
Fig. 6 The influence of recall initiation. The *leftmost panels* both show the distribution of PFR-scores across all subjects, which is broken down in to a Recency Group (*top panels*: PFR-score $\geq$ 3) and a Primacy Group (*bottom panels*: PFR-score $\leq$ −3). Within each group, we further divided subjects by list condition and instruction condition and examined their (*continuing left to right in the figure*) probability of first recall curves, serial position curves, lag-conditional response probability functions, temporal factor scores, semantic-conditional response probability function, and semantic factor scores. All *error bars* are bootstrapped 95% confidence intervals

lists, semantic factor scores tend to predict recall success, whereas for random lists, they tend to be unrelated. Again, this replicates and extends earlier findings that semantic contiguity was unrelated to recall success when lists are randomly constructed (Healey et al., 2014).

Temporal factor and semantic factor scores were generally uncorrelated or negatively correlated with each other (Fig. 7d; cf., Healey et al., 2014;). For related lists these correlations were significant regardless of instruction; for random lists, the correlation was significant only under meaning-focus instructions. It is not surprising that these types of clustering are negatively correlated as they are to a certain extent mutually exclusive when lists are presented in random order with respect to the semantic associations among items—you can not recall in perfect temporal and perfect semantic order at the same time.

Finally, PFR-scores tended to be uncorrelated with either temporal (Fig. 7e) or semantic (Fig. 7f) factor scores, except under order-focus instructions where low PFR-scores (i.e., a tendency to initiate recall from the beginning of the list) predicted more temporal clustering but less semantic clustering (at least for related lists). In other words, encouraging subjects to focus on order was associated with an increased tendency to initiate from the beginning of the list and an increased tendency to recall in order.

In sum, the correlations suggest that neither temporal clustering nor semantic clustering is universally related with successful memory search. Instead, the determinants of successful memory search depend on list composition and instructions. Memory success depends on using control processes to implement an appropriate strategy.

Discussion

Our aim was to illuminate the role of control processes in determining the relative influences of semantic versus
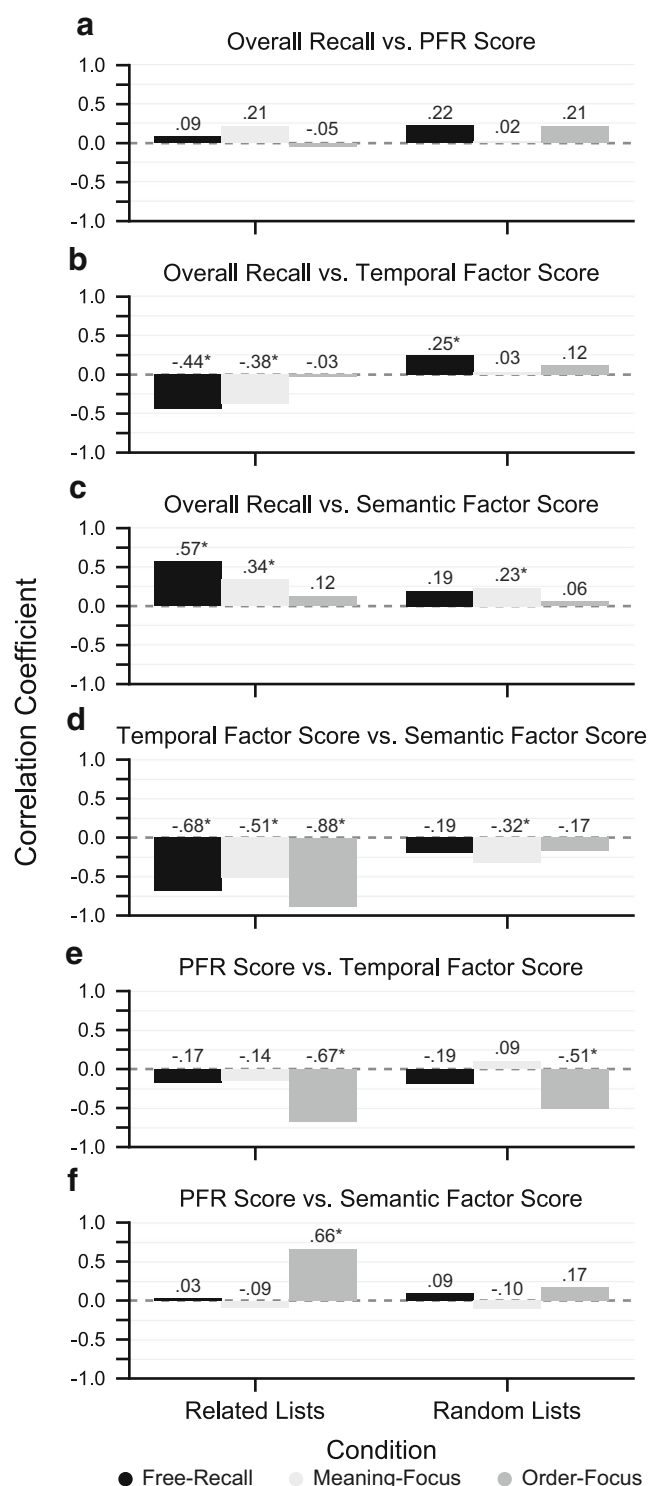
Fig. 7 Correlations among overall recall, pfr-scores, temporal factor scores, and semantic factor scores as a function of list type and instruction condition. Correlations that are significantly different than zero are marked with a *. All those with $|r| \geq .44$ are significant at $p < .001$ and those with $|r| \geq .23$ are significant at $p < .05$. If the difference between two correlations is $|r_2 - r_1| > 0.32$, then $p < .05$. If $|r_2 - r_1| > 0.54$, then $p < .001$. If, however, $|r_2 - r_1| < 0.20$, then $p \geq .05$. Because of differences in sample size across conditions, significance varies for differences of intermediate values

order information during memory search. We had subjects study lists that varied in their degree of semantic structure. Some were created by randomly selecting words from a word pool and thus had little semantic structure. Others were constructed such that there were four clusters of highly related words (presented in a scrambled order) and thus had considerable semantic structure. Replicating previous work (Polyn et al., 2011), we found that adding semantic structure reduced the overall TCE but that a robust TCE remained when subjects transitioned between items from the same semantic cluster *and* when they transitioned between items from different clusters. We made the novel discovery that the reduction in the overall TCE is accompanied by an increase in the SCE. We went beyond prior work by manipulating the task instructions to encourage subjects to engage different control processes. Some subjects were encouraged to use meaning-based associations to guide memory search and others were encouraged to use order-based associations to guide memory search. Subjects were quite successful at using control processes to follow these instructions. Asking them to focus on meaning *essentially eliminated* the overall TCE and increased the SCE. Asking them to focus on order-based associations accentuated the TCE and reduced (but did not eliminate) the SCE. A final set of analyses showed that subjects in all conditions could reconstruct the original serial order of studied lists, suggesting that although control processes powerfully modulated the use of order information during memory search, they do not prevent order information from being stored during study. Taken as a whole, these findings point to two important theoretical insights.

First, control processes exert a tremendous influence over the dimensions along which memory is searched. In some ways, this finding is obvious (e.g., the behavioral differences between free recall and serial recall are surely due to control processes). However, this finding is critical to our understanding of the TCE. The effects on the TCE reported here are extremely large, much larger than the effects of other variables known to powerfully influence recall, such as the length of the retention interval, the subject's level of expertise with the task, the subject's intellectual ability, and the subject's age (for a review, see Healey et al., in press). The only other manipulation known to reduce the TCE to the levels reported here is the complete elimination of control processes during encoding by giving subjects a surprise free-recall test (Healey, 2018; Nairne et al., 2017). Despite the size of the effect, many models specifically designed to account for the TCE do not specify control processes that could account for it. In the next section, we discuss what these control processes might be.

The second insight is that the mechanisms that encode information about input order seem to be automatic. The evidence for this claim is that despite the massive

effect of task instructions on the overall TCE, both within-cluster transitions and order reconstruction still showed a TCE, regardless of instruction conditions. Related evidence comes from recent work in which we showed that incidental encoding, which should eliminate any non-automatic encoding of order information, does not eliminate the TCE (Healey, 2018). Together, these findings suggest that subjects are always encoding order information but can flexibly decide whether or not to use that information based on the nature of the task. That is, control processes modulate the size of the TCE, but encoding of order information per se may be automatic. This complex picture of automatic encoding processes interacting with voluntary search processes is reminiscent of Atkinson and Shiffrin's (1968) original conception of a tight interplay between fixed structure and flexible control processes. A fuller understanding of recall dynamics will require more models that simulate this interplay (e.g., Lehman & Malmberg, 2013).

What are the control processes?

The results reported here suggest that control processes play a vital role in determining the balance between temporal and semantic contiguity. What are the control processes? The Atkinson and Shiffrin (1968) model and its descendants suggest multiple possibilities. The SAM theory (Raaijmakers & Shiffrin, 1980, 1981) suggests that at the beginning of each search attempt, the subject creates a retrieval plan that, among other things, specifies how retrieval cues will be generated: which types of information will be incorporated into the cue and the relative weighting of different types of information.

These ideas allow the theory to account for the influence of list composition. For example, Raaijmakers and Shiffrin (1980) conducted simulations showing that SAM can explain many aspects of free recall from categorized lists such as the classic Tulving and Pearl-stone (1966) findings by assuming that the categorized nature of the lists is taken into account when creating a retrieval plan. Their simulations were based on a retrieval plan in which for the first recall, subjects use only context as a cue. However, once a first item is retrieved, its category name is also retrieved and then both context *and* category information are used to cue further same-category recalls until some number of retrieval failures occur (presumably because most category-members have been recalled). At that point, the subject switches back to a context-only cue until a member of a new category is retrieved, at which point the new category is incorporated into the cue.

A similar retrieval plan could account for the present findings on how the semantic structure of the lists influenced memory search under free-recall instructions.

Specifically, as subjects discover either that the lists contain strong semantic structure (related lists) or little semantic structure (random lists), they could update their retrieval plans to include more or less semantic information in the cue. This would naturally modulate the size of the TCE. A similar mechanism could account for the effect of the instruction condition. Raaijmakers and Shiffrin (1980, 1981) argued that the retrieval plan is determined in part based on task instructions, thus the effect of instructions could be modeled by assuming that instructions to focus on either temporal order or semantic information changes the weighting of item-order versus semantic information as subjects generate retrieval cues.

These ideas also help explain the finding from the order reconstruction task that subjects in all conditions encoded information about item-order even if they did not use it during retrieval. Because retrieval plans govern search and not encoding, they allow subjects to encode information about item-order, leaving it available for an end-of-experiment test, but ignore it during the end-of-list test.

Control processes at encoding may also contribute to the effects. Ideas with their roots in Atkinson and Shiffrin (1968) are helpful here too. For example, instructions or stimuli that emphasize semantic information may encourage deeper levels of processing, which could be implemented by increasing the encoding of item features without changing the encoding of item-order information (Malmberg & Shiffrin, 2005; Lehman & Malmberg, 2013). This would cause memory search to be dominated by semantic associations while leaving item-order information available for use in the order reconstruction task. Another possibility is that upon being presented with a given word, the subject is reminded of its previously-presented associates, which are then brought into a buffer and corehearsed. This automatic reminding process could be supplemented with a deliberate compartmentalization process (Lehman & Malmberg, 2011, 2013) that actively drops unrelated items from the buffer. This would tend to produce strong within-cluster associations and perhaps even unitization that could account for within-cluster recall order. Yet another possibility is that instructions to focus on order versus meaning cause control processes to modulate the rate of contextual fluctuation (e.g., Mensink & Raaijmakers, 1988) during a list: slower contextual fluctuation would decrease the TCE.

Thus, models descended from Atkinson and Shiffrin (1968) include multiple control processes that could account for the present findings. These control processes could also help competing models of the TCE account for the findings. For example, retrieved context models (RCM; Howard & Kahana, 2002a; Sederberg et al., 2008; Polyn et al., 2009; Lohnas et al., 2015) have focused on formalizing the mechanisms that govern contextual dynamics and would

likely have to incorporate control processes to account for the current findings. One possibility is that presenting strong associates in a random order effectively decorrelates changes in temporal context from the passage of time, which would naturally attenuate the TCE and accentuate the SCE in a way that may be amplified by control processes (Polyn et al., 2011). Another possibility is that subjects create a retrieval plan which specifies how parameters of the memory system, such as rate of context drift or the balance of episodic versus semantic associations, should be tuned to meet the demands of a task (Healey et al., 2014; Healey & Kahana, 2014).

In sum, many control processes could produce the results reported here. And that is the point: we have well-developed models of the memory structures that generate the TCE but know little about the control processes that modulate it.

Conclusions

We have provided some preliminary data showing how control processes determine what types of associative information are used during memory search. We argue that these data reflect a mix of automatic processes that are invariant in the face of changing task demands and control processes that dynamically adapt to the details of a particular task. The field has many sophisticated models of how automatic processes and other aspects of the fixed structure of the memory system give rise to the TCE, the SCE, and other classic effects such as primacy and recency. However, our understanding of the control processes that allow us to apply this fixed structure to infinite variations in task demands has lagged behind. We hope that the current data serve to remind us of one of the central themes of Atkinson and Shiffrin (1968): that control processes are pervasive and will be an integral part of a truly general theory of memory.

Appendix

Detailed analyses

Below we report hypotheses tests for key comparisons. These are intended to supplement the confidence intervals we relied on in the main text.

Overall recall

Overall recall (see Table 1) was higher for related lists than random lists ($F(1, 430) = 18.921, p < .001, \eta_p^2 = .042$), and also differed as a function of instruction condition ($F(2, 430) = 11.476, p < .001, \eta_p^2 = .051$) with performance tending to be lower in the order-focus

condition than the other two conditions (in related lists, order-focus vs. free-recall: $t(144) = -3.890, p < .001$; order-focus vs. meaning-focus: $t(142) = -4.281, p < .001$. In random lists, order-focus vs. free-recall: $t(144) = -1.366, p = .174$; order-focus vs. meaning-focus: $t(143) = -2.075, p = .040$). There was no list structure × instruction condition interaction ($F(2, 430) = 1.727, p = .179, \eta_p^2 = .008$).

Recall initiation

PFR-scores (see Table 1) did not differ between related lists and random lists ($F(1, 430) = 0.835, p = .361, \eta_p^2 = .002$). They did, however, differ as a function of instruction condition ($F(2, 430) = 112.334, p < .001, \eta_p^2 = .343$) with the order-focus instructions encouraging more primacy (i.e., more negative values of PFR-score) than either the meaning-focus or free-recall conditions (in related lists, order-focus vs. free-recall: $t(144) = -8.639, p < .001$; order-focus vs. meaning-focus: $t(142) = -9.304, p < .001$. In random lists, order-focus vs. free-recall: $t(144) = -7.957, p < .001$; order-focus vs. meaning-focus: $t(143) = -8.383, p < .001$). There was no list structure × instruction condition interaction ($F(2, 430) = 0.445, p = .641, \eta_p^2 = .002$).

Overall contiguity

Temporal contiguity

Temporal factor scores (see Fig. 2) were lower for related lists than random lists ($F(1, 430) = 236.481, p < .001, \eta_p^2 = .355$), and also tended to differ as a function of instruction condition ($F(2, 430) = 120.181, p < .001, \eta_p^2 = .359$) with scores being higher in the order-focus condition than the other two conditions (in related lists, order-focus vs. free-recall: $t(144) = 9.553, p < .001$; order-focus vs. meaning-focus: $t(142) = 10.384, p < .001$. In random lists, order-focus vs. free-recall: $t(144) = 5.780, p < .001$; order-focus vs. meaning-focus: $t(143) = 8.382, p < .001$). There was a list structure × instruction condition interaction ($F(2, 430) = 6.524, p = .002, \eta_p^2 = .029$) reflecting the larger impact of instructions for related lists.

Semantic contiguity

Semantic factor scores (see Fig. 2) were higher for related lists than random lists ($F(1, 430) = 308.959, p < .001, \eta_p^2 = .418$), and also tended to differ as a function of instruction condition ($F(2, 430) = 100.344, p < .001, \eta_p^2 = .318$) with scores being lower in the order-focus condition than the other two conditions (in related lists,

order-focus vs. free-recall: $t(144) = -9.302, p < .001$; order-focus vs. meaning-focus: $t(142) = -12.175, p < .001$. in random lists, order-focus vs. free-recall: $t(144) = -2.988, p = .003$; order-focus vs. meaning-focus: $t(143) = -5.554, p < .001$). There was a list structure × instruction condition interaction ($F(2, 430) = 36.249, p < .001, \eta_p^2 = .144$) which was driven by the larger impact of instructions for related lists.

Within-cluster contiguity

Note that within-cluster scores (see Fig. 3) can only be computed for related lists, therefore the following analyses consider only related lists.

Temporal contiguity

Within-cluster temporal factor scores were higher in the order-focus condition than the other two conditions (order-focus vs. free-recall: $t(143) = 4.787, p < .001$; order-focus vs. meaning-focus: $t(142) = 4.482, p < .001$), which did not differ from each other (free-recall vs. meaning-focus: $t(143) = -0.422, p = .673$).

Semantic contiguity

Within-cluster semantic factor scores did not differ among instruction conditions (order-focus vs. free-recall: $t(143) = 0.217, p = .829$; order-focus vs. meaning-focus: $t(142) = -1.186, p = .238$; free-recall vs. meaning-focus: $t(143) = -1.527, p = .129$).

Between-cluster contiguity

Note that between-cluster scores (see Fig. 3) can only be computed for related lists, therefore the following analyses consider only related lists.

Temporal contiguity

Between-cluster temporal factor scores were highest in the order-focus condition and lowest in the meaning-focus condition (order-focus vs. free-recall: $t(144) = 9.930, p < .001$; order-focus vs. meaning-focus: $t(142) = 11.375, p < .001$; free-recall vs. meaning-focus: $t(144) = 2.689, p = .008$).

Semantic contiguity

Within-cluster semantic factor scores were lower in the order-focus condition than in the other conditions (order-focus vs. free-recall: $t(144) = -2.963, p = .004$; order-focus vs. meaning-focus: $t(142) = -3.272, p < $

.001), which did not differ from each other (free-recall vs. meaning-focus: $t(144) = -0.496, p = .621$).

Order reconstruction accuracy

Overall order reconstruction accuracy (see Table 1) was higher for random lists than related lists ($F(1, 430) = 10.403, p = .001, \eta_p^2 = .024$), and also tended to differ as a function of instruction condition ($F(2, 430) = 3.113, p = .045, \eta_p^2 = .014$). There was no list structure × instruction condition interaction ($F(2, 430) = 0.309, p = .735, \eta_p^2 = .001$). Whereas the order-focus condition performed better under random lists (order-focus vs. free-recall: $t(144) = 1.241, p = .216$; order-focus vs. meaning-focus: $t(142) = 2.193, p = .030$), under random lists, there was no difference (order-focus vs. free-recall: $t(144) = 1.292, p = .199$; order-focus vs. meaning-focus: $t(143) = 1.089, p = .278$).

References

Anderson, I., & Crosland, H. R. (1933). A method of measuring the effect of primacy of report in the range of attention experiment. *The American Journal of Psychology*, *45*(4), 701–713.

Asch, S. E., & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society*, *106*, 135–163.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: a proposed system and its control processes. In Spence, K.W., & Spence, T. (Eds.) *The psychology of learning and motivation* (Vol. 2, pp. 89–105). New York: Academic Press.

Bousfield, W., Whitmarsh, G., & Esterson, J. (1958). Serial position effects and the Marbe effect in the free recall of meaningful words. *Journal of General Psychology*, *59*(2), 255–262.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, *49*(2), 229–240.

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of sequences of restricted associative responses. *Journal of General Psychology*, *30*, 149–165.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539–576.

Craik, F. I. M. (1969). Modality effects in short-term storage. *Journal of Verbal Learning and Verbal Behavior*, *8*(5), 658–664. https://doi.org/10.1016/S0022-5371(69)80119–2

Craik, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, *9*, 658–664.

Dalezman, J. (1976). Effects of output order on immediate, delayed, and final recall performance. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(5), 597–608.

Dallett, K. M. (1963). Practice effects in free and ordered recall. *Journal of Experimental Psychology*, *66*(1), 65–71.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations

of recency effects. *Psychological Review*, *112*, 3–42. https://doi.org/10.1037/0033-295X.112.1.3

Deese, J. (1957). Serial organization in the recall of disconnected items. *Psychological Reports*, *3*(3), 577–582.

Deese, J., & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology*, *54*, 180–187.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

Goodwin, J. (1976). Changes in primacy and recency with practice in single-trial free recall. *Journal of Verbal Learning & Verbal Behavior*, *15*, 119–132.

Grenfell-Essam, R., & Ward, G. (2012). Examining the relationship between free recall and immediate serial recall: the role of list length, strategy use, and test expectancy. *Journal of Memory and Language*, *67*(1), 106–148.

Hasher, L. (1973). Position effects in free recall. *American Journal of Psychology*, *86*(2), 389–397.

Healey, M. K. (2018). Temporal contiguity in incidentally encoded memories. *Journal of Memory and Language*, *102*, 28–40.

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. https://doi.org/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*, 575–596. https://doi.org/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. https://doi.org/10.1037/rev0000015

Healey, M. K., Long, N. M., & Kahana, M. J. (in press). Contiguity in episodic memory. *Psychonomic Bulletin & Review*.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(6), 644–655. https://doi.org/10.1016/S0022-5371(74)80052-6

Hintzman, D. L. (2016). Is memory organized by temporal contiguity? *Memory & Cognition*, *44*, 365–375.

Hogan, R. M. (1975). Interitem encoding and directed search in free recall. *Memory & Cognition*, *3*, 197–209.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 923–941. https://doi.org/10.1037/0278-7393.25.4.923

Howard, M. W., & Kahana, M. J. (2002a). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299.

Howard, M. W., & Kahana, M. J. (2002b). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, *46*, 85–98.

Jahnke, J. C. (1965). Primacy and recency effects in serial-position curves of immediate recall. *Journal of Experimental Psychology*, *70*(1), 130.

Jenkins, J. J., & Russell, W. A. (1952). Associative clustering during recall. *Journal of Abnormal and Social Psychology*, *47*, 818–821.

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

Katz, L. (1968). The limited-capacity hypothesis: effects of sequence length and instructions in free recall. *Journal of Verbal Learning and Verbal Behavior*, *7*(5), 942–944.

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

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.

Lehman, M., & Malmberg, K. J. (2011). Overcoming the effects of intentional forgetting. *Memory & Cognition*, *39*(2), 335–347.

Lehman, M., & Malmberg, K. J. (2013). A buffer model of memory encoding and temporal correlations in retrieval. *Psychological Review*, *120*(1), 155–189. https://doi.org/10.1037/a0030851

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

Malmberg, K. J., & Shiffrin, R. M. (2005). The one-shot hypothesis for context storage. *Journal Experimental Psychology: Learning, Memory and Cognition*, *31*(2), 322–336. https://doi.org/10.1037/0278-7393.31.2.322

Manning, J. R., Polyn, S. M., Baltuch, G., Litt, B., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences*, *108*(31), 12893–12897. https://doi.org/10.1073/pnas.1015174108

McCluey, J. D., Burke, J. F., & Polyn, S. M. (2018). Temporal and semantic structure of a study list alters temporal organization in free recall. *Manuscript in Preparation*.

Memelink, J., & Hommel, B. (2013). Intentional weighting: a basic principle in cognitive control. *Psychological Research*, *77*(3), 249–259. https://doi.org/10.1007/s00426-012-0435-y

Mensink, G.-J. M., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review*, *95*, 434–455. https://doi.org/10.1037/0033-295X.95.4.434

Miller, J. F., Lazarus, E., Polyn, S. M., & Kahana, M. J. (2013). Spatial clustering during memory search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 773–781.

Morton, N. W., & Polyn, S. M. (2016). A predictive framework for evaluating models of semantic organization in free recall. *Journal of Memory and Language*, *86*, 119–140.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482–488. https://doi.org/10.1037/h0045106

Murdock, B. B. (1968). Serial order effects in short-term memory. *Journal of Experimental Psychology Monograph Supplement*, *76*, 1–15.

Nairne, J. S. (1990). Similarity and long-term memory for order. *Journal of Memory and Language*, *29*, 733–746.

Nairne, J. S. (1991). Positional uncertainty in long-term memory. *Memory & Cognition*, *19*(4), 332–340.

Nairne, J. S., Cogdill, M., & Lehman, M. (2017). Adaptive memory: temporal, semantic, and rating-based clustering following survival processing. *Journal of Memory and Language*, *93*, 304–314.

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning Memory, and Cognition*, *17*(4), 702.

Pollio, H. R., Richards, S., & Lucas, R. (1969). Temporal properties of category recall. *Journal of Verbal Learning and Verbal Behavior*, *8*, 529–536.

Polyn, S. M., Erlikhman, G., & Kahana, M. J. (2011). Semantic cuing and the scale-insensitivity of recency and contiguity. *Journal Experimental Psychology: Learning, Memory and Cognition*, *37*(3), 766–775.

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. https://doi.org/10.1037/a0014420

Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: a theory of probabilistic search of associative memory. In Bower, G. H. (Ed.)

The psychology of learning and motivation: advances in research and theory (Vol. 14, pp. 207–262). New York: Academic Press.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 93–134. https://doi.org/10.1037/0033-295X.88.2.93

Raffel, G. (1936). Two determinants of the effect of primacy. *The American Journal of Psychology*, *48*(4), 654–657.

Rugg, & Wilding (2000). Retrieval processing and episodic memory. *Trends in Cognitive Sciences*, *4*(3), 108–115.

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, *115*(4), 893–912. https://doi.org/10.1037/a0013396

Sederberg, P. B., Miller, J. F., Howard, W. H., & Kahana, M. J. (2010). The temporal contiguity effect predicts episodic memory performance. *Memory & Cognition*, *38*(6), 689–699. https://doi.org/10.3758/MC.38.6.689

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin and Review*, *4*, 145. https://doi.org/10.3758/BF03209391

Solway, A., Murdock, B. B., & Kahana, M. J. (2012). Positional and temporal clustering in serial order memory. *Memory & Cognition*, *40*(2), 177–190.

Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In Healy, A. F. (Ed.) *Cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer Washington, DC. American Psychological Association*.

Tulving, E. (1962). Subjective organization in free recall of unrelated words. *Psychological Review*, *69*(4), 344–354.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, *5*, 381–391. https://doi.org/10.1016/S0022-5371(66)80048-8

Uitvlugt, M. G., & Healey, M. K. (2019). Temporal proximity links unrelated news events in memory. *Psychological Science*, *30*(1), 92–104.

Unsworth, N. (2009). Variation in working memory capacity, fluid intelligence, and episodic recall: a latent variable examination of differences in the dynamics of free recall. *Memory & Cognition*, *37*(6), 837–849.

Unsworth, N., Brewer, G., & Spillers, G. (2011). Inter- and intra-individual variation in immediate free recall: an examination of serial position functions and recall initiation strategies. *Memory*, *19*(1), 67–82.

Ward, G., & Maylor, E. (2005). Age-related deficits in free recall: the role of rehearsal. *Quarterly Journal of Experimental Psychology*, *58A*(1), 98–119. https://doi.org/10.1080/02724980443000223

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial recall: the effects of list length and output order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1207–1241.

Ward, G., Woodward, G., Stevens, A., & Stinson, C. (2003). Using overt rehearsals to explain word frequency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 186–210.