Traveling economically through memory space: Characterizing output order in memory for serial order

STEPHAN LEWANDOWSKY
University of Western Australia, Crawley, Australia

GORDON D. A. BROWN
University of Warwick, Coventry, England
and University of Western Australia, Crawley, Australia

AND

JACQUELINE L. THOMAS
University of Western Australia, Crawley, Australia

How do people report the contents of short-term memory when information about order must be retained but items can be retrieved in any order? We report an experiment using an unconstrained reconstruction task in which people can report list items in any order but must place them in their correct serial positions. We found (1) a tendency to report recent items first in immediate but not in delayed reconstruction, (2) a tendency to recall temporally isolated items first, (3) a preference for forward recall order, and (4) a preference for output orders that minimize the length of the path that must be traversed through memory space during retrieval. The results constrain most current models of short-term memory in which retrieval is ballistic and is assumed to run to completion autonomously once initiated.

There has been much research emphasis on the order in which people retrieve items from long-term memory during free recall (see, e.g., Howard & Kahana, 2002; Kahana, 1996). This research has uncovered a number of empirical regularities: First, people tend to commence free recall with the report of one of the last list items (e.g., Howard & Kahana, 2002); second, once an item has been recalled, people's next report tends to be in a forward direction—that is, involving a later list item, preferably from nearby positions (i.e., short lags; see, e.g., Kahana, 1996). This information about output order has been crucial in shaping theories of free recall (e.g., Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Howard & Kahana, 2002), and its examination continues to provide new constraints on theories (Farrell & Lewandowsky, 2008).[1]

In short-term memory, by contrast, virtually nothing is known about the variables that determine people's report order, and no existing quantitative theories consider output order worthy of explanation. At first glance, this may not be entirely surprising, because most tests of short-term memory involve a mandatory report order. For example, in the classic immediate serial recall task, people must reproduce the list in the order of presentation. Nonetheless, several researchers have manipulated report order within

a serial recall task—for example, by asking participants to commence recall with the second half of a list before reporting the first half (see, e.g., Beaman, 2002; Cowan, Saults, Elliott, & Moreno, 2002). This simple manipulation has turned out to be theoretically quite diagnostic. For example, the fact that recency increases at the expense of primacy when the last list items are reported first has been taken to suggest that primacy may, in part, result from output interference (Cowan et al., 2002). Similarly, taking a further step to disentangle presentation order from report order, Oberauer (2003) selectively probed for the recall of specific list positions in random order. The absence of recency when performance was plotted with respect to output position (thus canceling out the effect of input serial position) was interpreted as evidence against response suppression (see Brown, Preece, & Hulme, 2000; Lewandowsky, 1999). These precedents underscore the theoretical diagnosticity of manipulations or examinations of report order; however, existing precedents are limited in one important respect, because participants were unable to choose on their own the order in which to report items. This limitation, which turns out to have considerable theoretical import, can be overcome either by use of unconstrained recall (see, e.g., Tan & Ward, 2007)—a procedure in which people are free to report items in any order

S. Lewandowsky, lewan@psy.uwa.edu.au

but must assign them to their list positions (unlike free recall)—or by use of an unconstrained reconstruction-of-order task (e.g., Lewandowsky, Nimmo, & Brown, 2008; Nairne, 1992). In the present article, we will focus on the latter.

In an unconstrained reconstruction task, participants are re-presented at test with the list items in random order and, rather than recalling them, must rearrange the items back into their order of presentation using an array of response options corresponding to the possible list positions. Crucially, there is no restriction on the order in which items may be selected for report. For example, people may select the last list position first and assign to it the item that they believe was last on the list (i.e., by clicking on a response box associated with the last list position and by clicking on the candidate item). Once the first item has been retrieved, people may then choose to report any of the remaining positions next; for example, they may elect to place an item into the first list position or in the penultimate position, and so on.

We used an unconstrained reconstruction task in the present study to present a detailed analysis of output order in short-term memory. We will proceed as follows: We will first argue that analysis of output order can illuminate metacognitive control processes operating in short-term memory. We will then present two theoretical perspectives that generate testable—albeit qualitative—predictions about output order. We will then present an experiment involving unconstrained reconstruction and analyze the data with particular reference to report order. The principal result is that people seek to minimize the length of the path that must be traversed through memory representations during retrieval.

Unconstrained Reconstruction
and Retrieval Control Processes

In immediate serial recall, report is in strict forward order. Several accounts of serial recall have been proposed, ranging from chaining (Lewandowsky & Murdock, 1989) through context evolution (see, e.g., Brown et al., 2000; Burgess & Hitch, 1999), to a competitive cuing process based on decreasing encoding strength of items across serial positions (e.g., Farrell & Lewandowsky, 2002; Page & Norris, 1998). Although these mechanisms differ considerably from each other, they share the common notion that retrieval is ballistic, so that once engaged, it runs to completion on its own and without much cognitive control.

In tasks such as unconstrained reconstruction, by contrast, report order is free but memory for an item's original position is required. Assuming that people deviate from strict forward retrieval (which can always be modeled by one of the preceding ballistic processes), how might a person choose the first item to report? And, once that first item has been retrieved, how do people transition to the next one? Clearly, these processes cannot be ballistic—that is, running to completion without intervention once launched—but must instead involve some type of control mechanism.

In the present article, we seek to characterize the constraints that govern this control mechanism. We propose two possible sources of constraints that are derived from two alternative, but not necessarily mutually exclusive, approaches to memory—namely, temporal distinctiveness (see, e.g., Brown, Neath, & Chater, 2007) and an event-based positional view (e.g., Lewandowsky & Farrell, 2008).

Output Order: Temporal and Positional Factors

According to the notion of temporal distinctiveness, the temporal separation of events at encoding is a crucial determinant of memory performance. All other things being equal, distinctiveness models predict that the memorability of an event increases with its temporal separation from neighboring events. Hence, given the list structure A....B....C.D, where the letters A, B, and C refer to arbitrary list items and each "." represents a unit of time, item B would be expected to be retrieved more accurately than if it had been presented on the list A.B.C.D. Likewise, distinctiveness predicts that recent items should be retrieved more accurately than temporally distant items. Crowder (1976) proposed an analogy involving telephone poles being viewed from a moving train: In the same way that evenly spaced telephone poles become visually less discriminable as they recede into spatial distance, memories become less discriminable from each other as they recede in time.

A recent computational instantiation of the temporal distinctiveness hypothesis is the scale independent memory, perception, and learning (SIMPLE) model of Brown et al. (2007). In SIMPLE, the success with which items are retrieved is governed by their separation from (and hence lack of confusability with) other items in a temporally organized (and potentially multidimensional) memorial space. Items that are widely separated from their neighbors—such as item B in the preceding example—are retrieved more accurately than their temporally crowded counterparts. Similarly, recent items are retrieved more accurately than earlier ones because time is assumed to be logarithmically transformed, thereby instantiating Crowder's (1976) telephone pole analogy.

A current limitation of SIMPLE is the assumption that retrieval relies on perfectly placed cues in the order mandated by the task. For example, in forward serial recall, each item is retrieved in response to a cue that is placed at exactly that point in the temporal past at which the target item had been presented. The cue is moved forward across the temporal dimension as output progresses, although SIMPLE does not explain the processes underpinning placement of the cue. Likewise, when output order is unconstrained, SIMPLE does not explain the order in which cues are placed, although its basic architecture suggests which orders should be preferred. Specifically, on immediate tests, SIMPLE predicts that early report of recent items would be advantageous. This is because as time passes, recent items lose their superior accuracy by receding into the increasingly crowded past. Items from the beginning of the list, by contrast, are already temporally crowded when retrieval commences, and they thus suffer comparatively less if their retrieval is delayed. On delayed tests, the early-report advantage of recent items is diminished; hence, SIMPLE would predict a reduced

or absent tendency for retrieval to commence with the recency portion of the list.

A recent study by Lewandowsky et al. (2008) provided some evidence for the existence of such a mechanism. Lewandowsky et al. presented participants with lists of seven letters that were presented singly and separated by unpredictable temporal gaps that ranged from 50 msec to 1.2 sec. An unconstrained reconstruction task that immediately followed list presentation revealed that, in line with SIMPLE's expectations, people often—but not always—reported the most recent items first. Tan and Ward (2007) reported the same preference using a similar unconstrained task but involving written recall, suggesting that recency-first report is a general characteristic of immediate tests. However, neither study examined the effects of delay: It is thus unknown whether the recency-first report tendency diminishes with retention interval as expected on a temporal distinctiveness view.

Lewandowsky et al. (2008) also found that people tended to report items that were temporally isolated from their neighbors (by larger temporal gaps) earlier than temporally crowded items, all other variables (such as serial position) being equal. Although this tendency is not a necessary consequence of an isolation view, it supports the notion that temporal factors are an important determinant of people's report orders.

A final aspect of the results of Lewandowsky et al. (2008) and of Tan and Ward (2007) was that notwithstanding the tendency to report recent items first, people nonetheless showed a strong preference overall for forward report of all items (although in Tan & Ward, 2007, this tendency was limited to six-item lists, whereas it was largely absent with eight-item lists). This preference for forward retrieval is not readily reconciled with the distinctiveness notion (see Brown, Chater, & Neath, 2008) and suggests that criteria other than temporal crowding also determine output order.

The preference for forward retrieval is reminiscent of other empirical precedents that point to the importance of positional proximity in retrieval transitions. In free recall, there is considerable evidence that people prefer to report positionally proximal items in succession (see, e.g., Kahana, 1996), although this preference is sometimes not monotonically decreasing (Farrell & Lewandowsky, 2008). In serial recall, Kahana and Caplan (2002) furthermore showed a strong associative asymmetry favoring forward over backward transitions. Those two precedents are consistent with the preference for strict forward retrieval found by Lewandowsky et al. (2008) and Tan and Ward (2007). In the present article, we build on those precedents and develop the notion of "path length"—defined as the sum of all pairwise transition lags during retrieval of a list—and postulate that people seek to minimize path length during retrieval.

What does it mean to minimize path length in memory retrieval? In many models of memory, it is assumed (either implicitly or explicitly, as in the case of SIMPLE) that memory traces are located within a multidimensional psychological space, with their proximity in that space determined by their similarity. Thus, to-be-remembered items will occupy nearby locations in memory if they are semantically similar, phonologically similar, or presented close together in time or serial position. Minimizing path length thus implies that people favor transitions within clusters of items that are proximal in psychological space. In free recall, there is ample evidence that similar items tend to be recalled together even when they were not presented together. This clustering at output appears to hold, whatever the dimension of similarity (see Brown et al., 2008). For example, Bousfield (1953) found that semantically similar items tend to cluster in free recall, and the same is true of items studied in similar contexts (Hintzman, Block, & Inskeep, 1972)—and of course it is particularly true of items in positional proximity (Kahana, 1996). A natural explanation of all those forms of clustering is that participants seek to minimize the distance of the path they must travel through psychological space during retrieval. Direct evidence for this idea was provided by Romney, Brewer, and Batchelder (1993), who showed that people preferentially recalled items from long-term memory in an order that minimized the path traversed through the multidimensional semantic similarity space associated with the list (which was obtained by similarity scaling). It is, however, unknown whether the notion of path length also applies to retrieval from short-term memory, when order information must be retained and when path length is defined with respect to positional information rather than semantic similarity.

Why might it be advantageous to minimize path length? Nairne, Ceo, and Reysen (2007) recently examined the effects of one retrieval on another in probed short-term recall. In their studies, specific items were cued for successive retrieval, thus controlling lag as well as serial positions of the two targets and permitting analysis of the success of transitions. Nairne et al. found a large accuracy advantage for forward transitions over backward transitions (i.e., if the first-retrieved item preceded the second one on the list, recall performance on the latter was improved), and, irrespective of direction, transitions between adjacent items (lag ± 1) were more accurate than transitions across an intervening item. Further evidence—albeit from a different task (recognition memory)—comes from a study by Schwartz, Howard, Jing, and Kahana (2005), who found that recognition was more accurate when a studied item was tested immediately following the test of an adjacent item. Similar distance effects in recognition, using latency rather than accuracy as the primary measure, have been obtained by McKoon and Ratcliff (1980; see also Lewandowsky, 1986). These results point to a memorial cost that is associated with increasing transition distances, and they support our contention that people may seek to minimize path length because it may maximize accuracy.

In summary, although very little is known about report order in immediate memory retrieval, the few available precedents (Lewandowsky et al., 2008; Tan & Ward, 2007) suggest that both temporal and positional considerations govern people's retrieval strategies. We therefore postulate that people's retrieval strategies are built around two strong—and not necessarily compatible—constraints. First, in immediate retrieval, it is advantageous for one to report recent items first, before their relative advantage has dissipated over time. This advantage diminishes with

delay. Second, it is advantageous for one to minimize "path length" by keeping pairwise transitions small and in a forward direction. At present, the way in which those two constraints interact is completely unknown. We will now present an experiment that explored those constraints with respect to the variables that determine report order in short-term memory.

Experimental Goals and Hypotheses

The experiment pursued three specific goals. First, we tested the possibility that early report of recency items might be reduced after a delay. Accordingly, the experiment used two retention intervals, which were 0.5 sec for immediate reconstruction and 8 sec for delayed reconstruction.

Second, on the basis of the observation by Lewandowsky et al. (2008) that people's report order was influenced by temporal isolation, the experiment again used unpredictable temporal gaps in between list items. Our analysis focused on the role of isolation as well as serial position in determining output order. Unlike recency-first reports, the preferential early report of isolated items might be invariant across retention intervals because the relative advantage of isolated items over other list items—all other factors, such as serial position, being equal—is unaffected by retention interval.

Third, we tested the prediction that participants will seek to minimize path length during retrieval. Any tendency to minimize path length should be unaffected by delay, because positional transition distances are time invariant. To facilitate examination of the role of path length, our analysis went beyond the conventional examination of pairwise transitional contingencies in retrieval. To date, experiments in free recall have typically focused on which item i will be recalled on recall attempt n, given that item j was recalled on attempt $n-1$, a measure known as conditional recall probability (CRP). Although analysis of CRPs can be quite diagnostic, a focus on pairwise transition probabilities runs the risk of missing characteristics of the output sequence as a whole. For example, a preference for output orders that minimize path length cannot be captured by consideration of pairwise contingencies. Our analysis, therefore, focused on the overall report sequence in addition to pairwise transitions.

METHOD

The experiment varied the temporal isolation of list items and permitted retrieval in any order using an unconstrained reconstruction task. Lists were followed by brief (0.5 sec) or long (8 sec) retention intervals at random. Because we were interested in the possible effects of time per se, without a contribution from other processes such as rehearsal, participants engaged in articulatory suppression throughout list presentation, retention interval, and retrieval.

Participants

A total of 24 members of the authors' campus community participated voluntarily in exchange for course credit or without remuneration.

Stimuli and Apparatus

All lists consisted of seven letters that were sampled randomly without replacement from a pool of 13 consonants (B, F, H, J, K, L, M, N, R, S, W, X, and Z). Six interstimulus intervals (ISIs) of 50-, 100-, 200-, 400-, 800-, and 1,200-msec duration were used for every list. All possible permutations of the ISIs were used, resulting in 720 unique sequences of intervals. The 720 sequences were split into 12 sets of 60, with each ISI presented in each possible position approximately 10 times (range, 8–12) per set. Participants were randomly assigned to one of the sets, which they completed once with immediate reconstruction and once with delayed reconstruction.

On immediate-reconstruction trials, the last item on the list was followed by the response screen after 500 msec. In delayed reconstruction, 8 sec elapsed between study and test. The order of the 60 trials at each retention interval and the interleaving of the two sets were rerandomized for each participant.

The stimuli were displayed and responses were recorded via a Windows-based computer running a MATLAB program designed using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

Procedure

The experiment consisted of 120 study–test trials involving unconstrained reconstruction. Each trial commenced with the presentation of a fixation cross for 400 msec. The list items were then presented on the screen one at a time for 400 msec each. The time between items was determined by the permutation of ISIs for that particular trial. Following the condition-dependent retention interval for that trial, retrieval commenced by displaying the response screen for the reconstruction task. Participants could not anticipate the retention interval during list presentation.

The top of the response screen showed seven boxes that contained the list items in random order. In addition, a row of seven empty "response" boxes was shown at the bottom of the screen. Participants used the mouse to select one of the boxes containing a list item by clicking on it. Participants then clicked on one of the empty response boxes at the bottom of the screen to indicate where in the list the selected item had occurred (with box position denoting serial position). Once a response box was chosen, the item appeared in the response box in black font, and its box at the top of the screen turned gray to indicate that that item could not be selected again.

Participants could not place an item into an already-filled response box, nor could they move an item after it had been placed into a response box. Participants were instructed that they could fill the response boxes in any order. The next trial commenced 3.5 sec after the last response was made.

Participants repeated the word "salt" aloud throughout list presentation, retention interval, and reconstruction. To ensure that this articulatory suppression continued throughout the experiment, the experimenter remained present for all trials. Participants completed four practice trials before commencing the experiment. A self-paced break was scheduled after every 30 experimental trials.

RESULTS

We will present the results in four stages. We will begin by reporting conventional serial position curves and examining the probability with which items were reported first. We will next report pairwise contingencies between successive reports and examine the accuracy associated with various transitions. We will then move beyond pairwise contingencies and analyze overall output orders for the entire list, seeking predictors for the frequency of use of those report orders. Finally, we will relate report order to accuracy of reconstruction.

Serial Position and First-Report Analysis

Individual participants' performance across all trials ranged from .38 to .69, safely away from both ceiling and
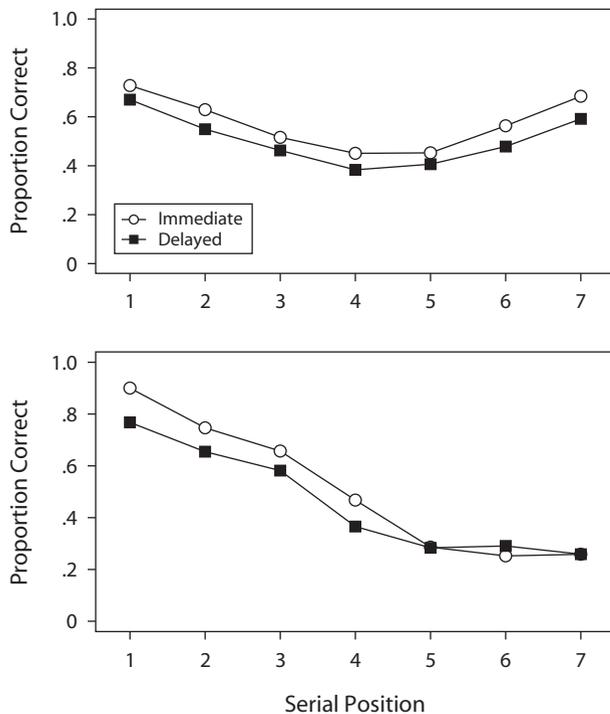
Figure 1. Serial position curves showing correct-in-position performance for immediate and delayed reconstruction. The top panel shows data aggregated across all observed output orders, and the bottom panel shows performance for only those lists that were reconstructed in strict forward order.

floor; hence, all participants were retained for analysis. The top panel of Figure 1 shows the serial position curves for immediate and delayed reconstruction, aggregated across all observed output orders. The apparent reduction in accuracy after a delay was confirmed by a within-subjects ANOVA, which showed a significant effect of both delay [$F(1,23) = 35.12$, $MS_e = .011$, $p < .0001$] and serial position [$F(6,138) = 25.92$, $MS_e = .021$, $p < .0001$], but no interaction between the two [$F(6,138) = .87$].[2]

The bottom panel of Figure 1 shows the serial position curves when consideration was restricted to only those trials on which reconstruction was in strict forward order. As would be expected on a temporal-distinctiveness account, under those circumstances, recency was entirely absent, and performance on the primacy portion of the list was slightly higher than that for the overall data in the top panel.

We next examined people's preferred starting point of reconstruction by considering the first-retrieval probabilities (FRPs; see, e.g., Howard & Kahana, 2002). Figure 2 displays the FRPs in the conventional manner, by plotting the average probability across participants with which items from the various serial positions were reported first. In order for this analysis to be commensurate with the free-recall literature, only correct first reconstructions were considered; the pattern is virtually unchanged if erroneous responses are also included (not shown in detail in the present article). It is clear that there was a strong tendency

to initiate retrieval from the beginning of the list irrespective of retention interval, and that first-reports from the end of the list (Positions 6 and 7) were more frequent in immediate than in delayed reconstruction. The pattern for immediate reconstruction largely mirrored the results of Tan and Ward (2007), although Tan and Ward observed more first responses from the later list positions than we did in the present study.

The contrast between the pattern in Figure 2 and the results typically observed in free recall is striking: Whereas in free recall FRP functions tend to be characterized by extreme recency (see, e.g., Howard & Kahana, 2002), the present FRP functions strongly favored the first item—albeit to a lesser extent than the last item is favored in free recall. One variable that is likely to contribute to this difference in outcome is list length: Most free recall experiments involve lists with a dozen items or more. In confirmation, Tan and Ward (2007) observed considerably more first retrievals from later list positions with eight-item lists (their Experiment 1) than with six-item lists (their Experiment 2).

The use of varying temporal intervals between list items permitted an examination of the role of temporal isolation in determining the starting point for reconstruction. To obtain a suitable temporal isolation score, we first computed the total temporal isolation of each item by adding the intervals surrounding it (doubling the interval following or preceding it, respectively, for the first and last item, to permit inclusion of those items in the analysis), and then used its position within the ascending sequence of all such isolation values—that is, its temporal isolation rank—as a predictor of first report. By this relative isolation measure, the temporally most crowded item has Rank 1, and the temporally most isolated item has Rank 7.[3] Figure 3 shows
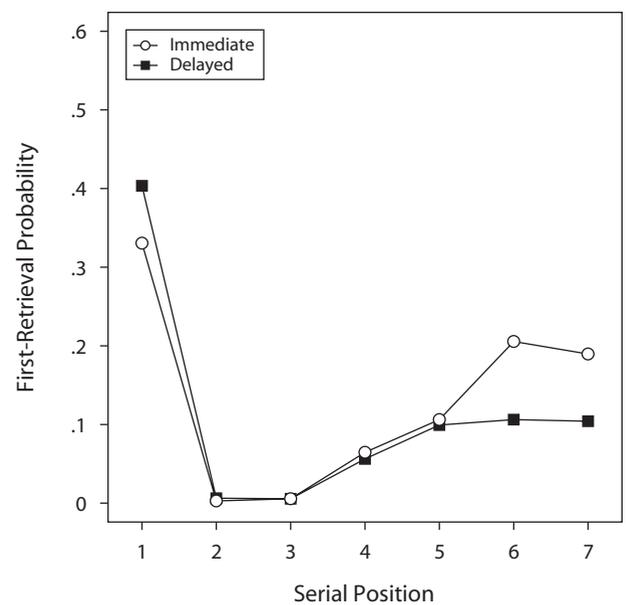
Figure 2. First-retrieval probabilities based on correct responses only as a function of serial position for immediate and delayed reconstruction.
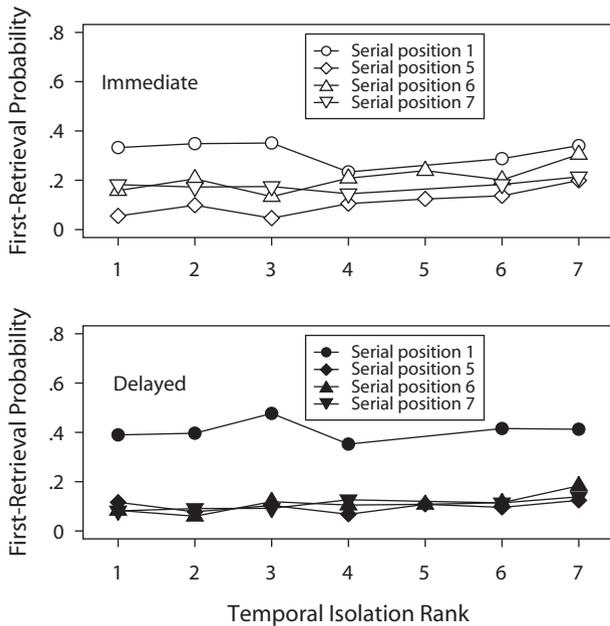
Figure 3. First-retrieval probabilities based on correct responses only as a function of an item's temporal isolation rank for immediate and delayed reconstruction, broken down by serial position. Serial positions 2, 3, and 4 are not shown because they did not contribute a sufficiently large number of observations; see Figure 2.

the FRPs as a function of isolation rank and broken down by serial position of the first-reported item.

Considering delayed reconstruction first, it is apparent that isolation rank had no effect at the later serial positions, although there was a hint of an effect for the first item. Intriguingly, that effect appeared to be that the first item was more likely to be reconstructed first if it was temporally less isolated. Turning to immediate reconstruction, the pattern for the first item resembled that for the delayed trials but was further accentuated: There was a notable tendency for the first item to be reported first when it was temporally crowded (Ranks 1–3), and to a lesser extent, when it was temporally isolated (Ranks 6 and 7). For the remaining serial positions, by contrast, temporal isolation had a monotonically increasing effect; the more isolated a late list item was, the more likely people were to commence immediate reconstruction with that item.

These data suggest that people particularly prefer to initiate retrieval with the first item if it is temporally crowded, thus perhaps exploiting the unique positional distinctiveness of the first item (which, unlike other items, has only one neighbor) in those circumstances. Conversely, recent items are preferentially reported first when their relative temporal isolation is particularly large (i.e., in immediate but not in delayed reconstruction).

Pairwise Report Contingencies

Characterizing transition preferences: Lag-CRP analysis. To date, examinations of output contingencies have largely focused on the examination of pairwise transitions in the form of a lag-CRP analysis—that is, the

conditional probability with which a subsequent item is reported, as a function of its lag (i.e., difference in serial positions) from the previously retrieved item. Figure 4 shows the average lag-CRP functions obtained in the present experiment. The figure shows the transition probability between the first two retrieved items as a function of the lag between them; readers not familiar with the intricacies of CRP analyses may consult the Appendix for details.

Not unexpectedly, the figure reveals a bias for immediate (+1) forward transitions at both retention intervals. However, in contrast with many lag-CRPs in free recall (see, e.g., Howard & Kahana, 2002; Kahana, 1996), the present CRPs show a striking upturn for extreme negative lags at both retention intervals, with the largest backward transitions (−6) being nearly as frequent as immediate forward transitions. This upturn is contrary to an apparent earlier consensus in the literature that lag-CRP functions (at least in free recall) are always monotonically decreasing, but it confirms a recent re-analysis of numerous free-recall studies that showed that lag-CRP functions in free recall can also be nonmonotonic and, like the one obtained here, often show an upturn for extreme transitions (Farrell & Lewandowsky, 2008). The present lag-CRPs differ from those obtained in free recall in several other respects. In particular, most transitions are to the immediately adjacent items, with very few transitions at lags of ±2 or ±3. The dominance of adjacent transitions points to people's
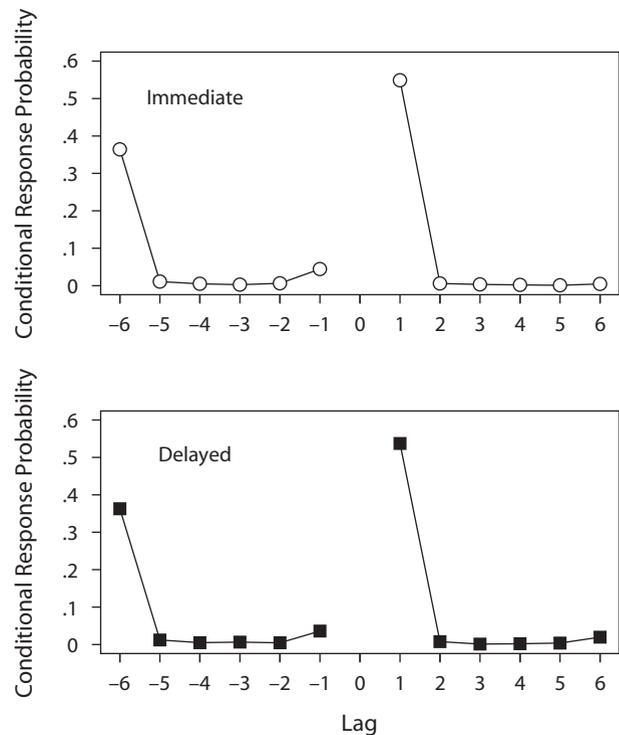
Figure 4. Lag-CRP functions describing the transitions between the first two reported items for immediate and delayed reconstruction. Only correct responses are considered. Positive lags occur when the second retrieved item followed the first one on the list, and negative lags occur when the second item preceded the first one on the list.

tendency to commence their report with ordered retrieval of several list items, a tendency that we will examine in more detail below.

The transitions can be further broken down by the serial position of the first-reconstructed item. Figure 5 shows the resultant pattern, limited to the serial positions that were reconstructed first by at least 20 participants (to ensure stability). The figure shows that for all but the last list item, people strongly preferred forward transitions: This was true even for the sixth list item, for which virtually all transitions were +1 and hence to the last list item. If the last list item (Position 7) was reported first, by contrast, people either transitioned to the first item (−6) or to the second-to-last item (−1).

Another intriguing feature of Figure 5 is that the transition probabilities did not differ with delay. Once people completed their first report, they transitioned to the next item in the same manner, irrespective of retention interval. (Note that the earlier differences in FRPs between retention intervals do not show up in the CRPs because the latter probabilities are conditionalized on report of the first item and normalized; see the Appendix.)

Transitional success: Lag-accuracy analysis. Unlike free recall, the reconstruction methodology permits examination of the success with which a transition is made from one item to the next. Specifically, one can exam-

Figure 5. Lag-CRP functions describing the transitions between the first two output positions for immediate and delayed reconstruction. Only correct responses are considered. Transitions are broken down by serial position of the first-reconstructed item. Only those first-reconstructed serial positions that are based on data from more than 20 participants are considered. Transition probabilities are normalized to sum to 1 for each serial position separately.

ine accuracy of reconstruction at the kth output position (i.e., whether or not the correct item is being placed in the response box chosen for the kth report) as a function of the positional lag between the $k−1$th and the kth item. Unlike for the lag-CRP analysis, to ensure a sufficiently large number of observations, we considered Output Positions 2–7 to compute the recall accuracies associated with transitions of lags +1, +2, −1, and −2.[4] Figure 6 shows the resultant lag-accuracy functions in two ways: The "unconditional" curves plot accuracy at output position k irrespective of the success of the previous response, whereas the "previous response correct" curves consider responses at output position k only if the preceding response $k−1$ was correct. Because both types of responses were considered with respect to the same denominator (see the Appendix for details), the latter were necessarily lower than the former.

The figure reveals two interesting regularities: First, transitions from immediately adjacent positions (i.e., lags ±1) resulted in more accurate responses than did transitions across an intervening item (i.e., lags ±2). Second, this advantage for adjacent positions was particularly pronounced for forward transitions, extending the recent finding by Nairne et al. (2007) to situations in which transition distance was under the participants' control. These results reinforce the notion that people may seek to minimize path length in order to maximize accuracy. In order to examine the role of path length, we will now consider the pattern of overall output orders. In contrast with previous examinations of pairwise transitions, our new approach allows characterization of the output sequence as a whole.
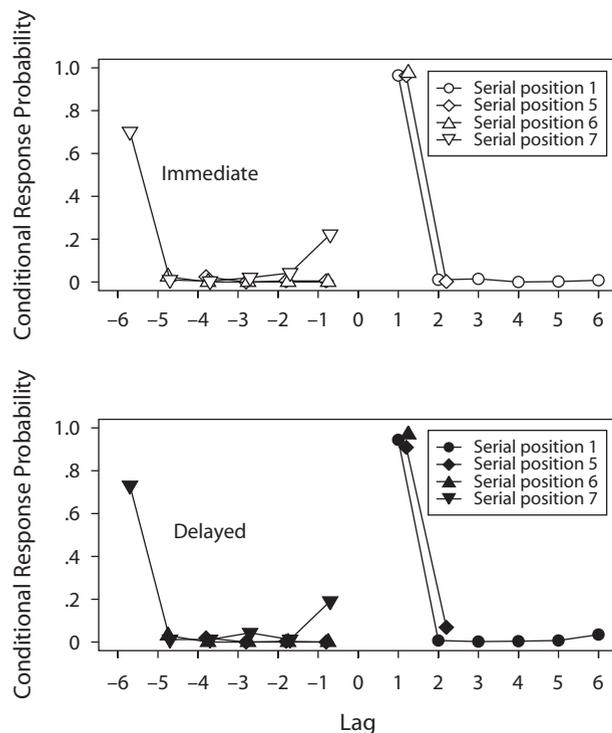
Overall Output Order

Across all trials and participants, 356 unique output orders were observed out of a possible 2,880.[5] Of those, 114 occurred at both delays, 96 occurred uniquely with immediate reconstruction, and 146 occurred uniquely with a delay. Thus, there were 210 different output orders altogether in immediate reconstruction and 260 different orders in delayed reconstruction.
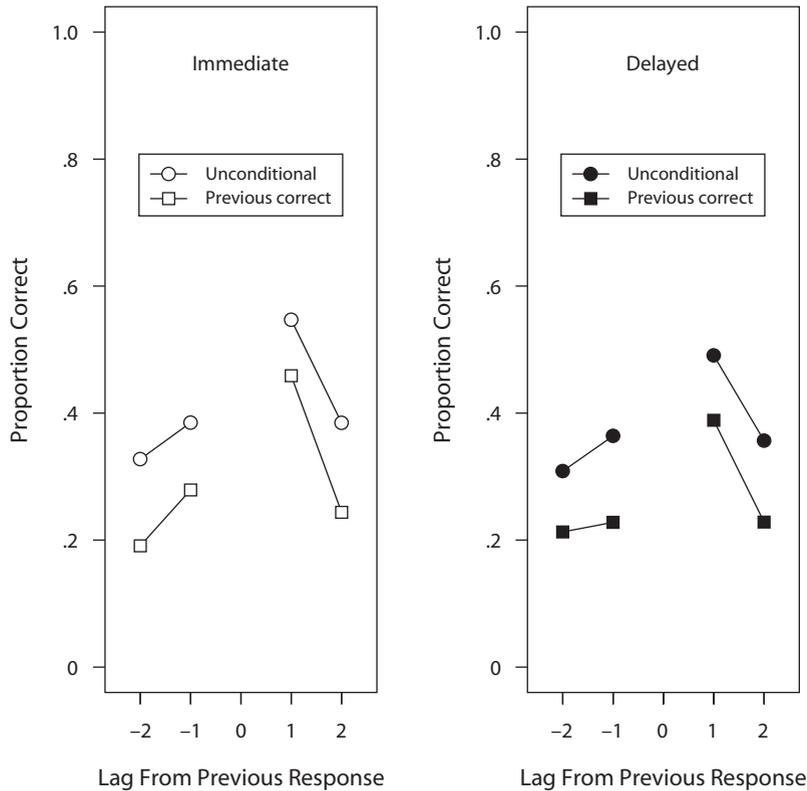
The modal output order in both conditions involved strict forward serial reconstruction (i.e., reporting the list items in the order 1234567), which occurred on 14.7% and 12.8% of all trials in the immediate and delayed reconstruction conditions, respectively. This difference did not approach significance [$t(23) = 1.15, p = .26$].

In confirmation of the FRPs in Figure 2, participants initiated recall with the first list item more frequently after a delay ($M = .46$) than in immediate reconstruction ($M = .36$) [$t(23) = 3.98, p = .001$]. Conversely, participants started retrieval with the last item significantly more often in the immediate ($M = .20$) than in the delayed ($M = .11$) conditions [$t(23) = 3.01, p = .006$]. (Note that unlike the data shown in Figure 2, this analysis did not conditionalize on the first response being correct; in consequence, these means are higher than those shown in the figure.)

Tables 1 and 2 show the 10 most frequent report orders for immediate and delayed reconstruction, respectively.

Figure 6. Lag-accuracy functions describing the success of transitions between consecutive output positions for immediate and delayed reconstruction. Data are aggregated across Output Positions 2–7 and show the proportion of correct responses at output position k as a function of positional transition lag from the item in output position $k-1$, either irrespective of the success at $k-1$ (unconditional) or when retrieval $k-1$ was successful (previous correct).

Output orders are represented by listing the serial positions of the items (i.e., response boxes) in the order in which they were reconstructed; hence, 4512... means that the fourth response box was filled first, then the fifth, then the first, then the second, and so on.

The tables also show the accuracy associated with those report orders that are discussed in a later section. The tables reveal that when people did not use strict forward report, they tended to exploit the recency advantage by commencing report with the last few items. To summarize

this effect, we considered the following output orders to exploit the recency advantage: 7123456 (report the last item first, then continue from the beginning of the list), 6712345 (report the last two items before reconstructing in forward order from the beginning), 5671234, and 4567123. When those orders are considered together, their incidence was clearly greater in immediate (.36) than in delayed (.28) reconstruction.

A further perspective on those recency-based output orders is provided in Figure 7, which shows the associated

Table 1
**Most Frequent Output Orders and Associated Accuracy in Immediate Reconstruction**

Output Order	Frequency	Proportion	Accuracy
1 2 3 4 5 6 7	212	.147	.52
6 7 1 2 3 4 5	196	.136	.65
7 1 2 3 4 5 6	122	.085	.59
5 6 7 1 2 3 4	119	.083	.64
4 5 6 7 1 2 3	81	.056	.72
7 6 1 2 3 4 5	41	.029	.53
1 2 3 4 7 6 5	33	.029	.65
6 7 1 2 3 5 4	28	.019	.67
6 7 4 5 1 2 3	28	.019	.67
1 2 3 4 7 5 6	27	.019	.59

Note—Accuracy values were obtained by averaging across all serial positions, trials, and participants for a given output order.

Table 2
Most Frequent Output Orders and Associated Accuracy in Delayed Reconstruction

Output Order	Frequency	Proportion	Accuracy
1 2 3 4 5 6 7	184	.128	.46
6 7 1 2 3 4 5	120	.083	.53
5 6 7 1 2 3 4	119	.083	.56
7 1 2 3 4 5 6	83	.058	.35
4 5 6 7 1 2 3	74	.051	.72
1 2 3 6 7 4 5	59	.041	.62
1 2 3 7 4 5 6	43	.030	.53
1 2 6 7 3 4 5	38	.026	.58
1 2 3 4 7 5 6	36	.025	.65
1 2 3 4 7 5 6	30	.021	.66

Note—Accuracy values were obtained by averaging across all serial positions, trials, and participants for a given output order.
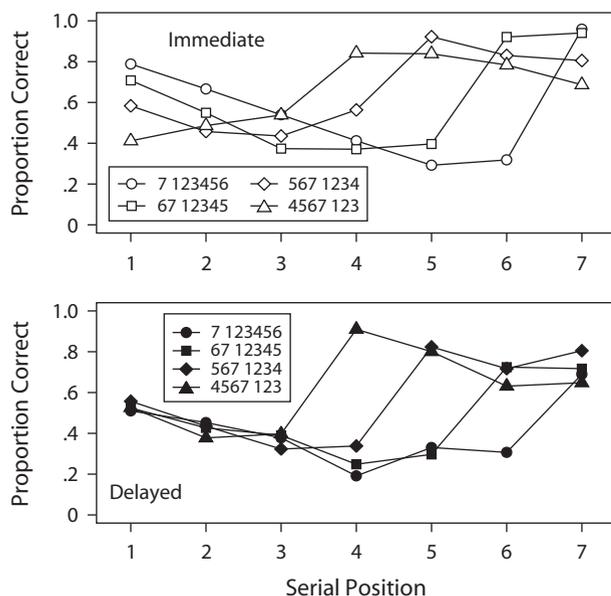
Figure 7. Serial position curves for output orders that exploited recency for immediate and delayed reconstruction. Legend entries describe output orders; see text for explanation.

serial position curves, obtained by first averaging performance across each participant's relevant trials and then aggregating across participants. The figure illustrates that irrespective of retention interval, early report of terminal list items was associated with high accuracy. The top panel of the figure furthermore shows that for immediate reconstruction, primacy was reduced as the number of first-reported terminal items increased from 1 through 4; for delayed recall, by contrast, primacy (i.e., performance on the first three items) remained constant across output orders. The trade-off between recency and primacy that we observed for immediate reconstruction is consonant with results by Cowan et al. (2002) and Beaman (2002), which were obtained with forward serial recall that commenced with later list items and then "wrapped around" to report the earlier list positions.

Tan and Ward (2007) reported that similar large differences between their unconstrained serial position curves were eliminated when performance was normalized and considered as a function of output position. We applied their analysis in the present article by examining performance as a function of the ordinal sequence in which the response boxes were filled and by using the proportion of total errors as the response measure (thus removing absolute differences in performance). In confirmation of the observation by Tan and Ward (2007), the analysis (not shown in detail here) revealed that the performance patterns were notably similar for all orders and showed a monotonic decrease across output positions. The absence of recency resembled the finding by Oberauer (2003) obtained with random successive cuing of list positions.

Output order and temporal isolation. To examine whether the temporal isolation profile of a list predicted its likely reconstruction order, we sorted the list items on each trial from least to most isolated for the earlier analy-

ses of temporal FRPs and assigned each item a temporal isolation rank. These isolation ranks were then correlated (by Spearman rank correlation) with the observed output order for each participant trial. Note that if temporal isolation contributed to report order by favoring early report of isolated items, then these correlations would be negative because isolation ranks increased with increasing isolation. Across subjects, the mean correlations differed from 0 for both immediate ($M = -.10$) [$t(23) = -8.20$, $p < .0001$] and delayed ($M = -.10$) [$t(23) = -10.42$, $p < .0001$] reconstruction, but the two means did not differ from each other [$t(23) = .19, p > .1$].

These results suggest that people's overall output orders were at least partially governed by the temporal isolation profile of the lists. Although the absolute magnitudes of the correlations were small, they were significant at both retention intervals. The analysis thus extended the earlier result—namely, that temporal isolation contributed to the choice of first report—to the remaining output positions. We conclude that people show a tendency to report lists in decreasing order of isolation, beginning with the more isolated items and terminating with the most crowded items. Unlike the tendency to report recent items first, the early report of isolated items was not diminished by a longer retention interval.

All remaining analyses considered only those orders that occurred at least seven times in either immediate or delayed reconstruction, because it would not have been feasible to compute accuracies or other summary statistics for less frequent report orders.

Output order and path length. For the reasons noted at the outset, we assumed that there was a cost to moving from one part of psychological space to another between retrievals. This cost was assumed to increase with the distance traveled; that is, it was assumed to increase with transition lag (see Figure 6 for evidence that this was the case in the present study; compare lags ± 1 and ± 2). We therefore expected that participants would show a tendency to prefer output orderings that minimized the overall cost of traveling through the space. So, for example, the output recall sequence 4512 involves three transitions that are—in order—+1, −4, and +1, whereas 1234 involves three consecutive transitions of +1, and so on. We expect the latter to be preferred because its total absolute path length (3) is considerably less than that of the former (6).

What is the psychological cost of making a transition of a given size? We assumed that Fitt's Law (Fitts, 1954) applies, and, hence, that the cost of traveling a distance d through memory space depended on $\log(d)$. On the basis of the results reported by Nairne et al. (2007), we made the further assumption that backward transitions were more costly than forward transitions. To accommodate this asymmetry, backward transitions were weighted by a parameter w. Thus, the cost of a transition of −4 would be $\log(w \times 4)$, whereas the cost of a transition of +2 would be $\log(2)$. The total cost of an output order, then, was the sum of the costs of each of the six transitions (for the seven-item lists used presently).

We estimated the value of w by routine Simplex methods to maximize the correlation between total path cost

and frequency of usage of the various output orders. There were 27 and 28 output orders that occurred seven or more times in immediate and delayed reconstruction, respectively, and that were considered in two separate analyses. The best-fitting estimates of w were 3.55 and 4.16 for immediate and delayed reconstruction, respectively, which gave rise to correlations of $r = -.60$ and $r = -.71$, respectively, between path cost and usage.

The large magnitude of those correlations suggests that people sought to avoid output orders that involved transitions between distant serial positions—especially in a backward direction—and instead preferred output orders consisting of short, preferably forward, steps. These results are consonant with the earlier analyses: The lag-CRP functions (see Figure 4) suggested that people primarily favored small (forward) transitions, and the lag-accuracy functions confirmed that immediate transitions were more accurate than more distant transitions. We conclude that in line with our expectations, people seek to minimize path length when retrieving lists from short-term memory.

However, path cost on its own cannot account for the fact that people also often exploited the recency advantage by retrieving the last list item(s) first; any report order that commenced with end-of-list items and then resumed from the beginning of the list necessarily involved at least one large backward transition, and the lag-CRP functions (Figure 5) confirm that those large transitions occurred quite frequently. We therefore sought to augment the role of path cost by considering additional potential predictor variables that captured the exploitation of recency. This next analysis, therefore, provides a window into the trade-off between the two constraints we proposed at the outset—namely, minimizing path length versus maximizing the advantage afforded by temporal distinctiveness.

Path cost and recency advantage. We designed an index of people's tendency to exploit recency, which we call *lifo* from here on ("last-in-first-out"), and which consisted of the number of items that were reported from the recency portion of the list, in any order, before retrieval continued with early list items. Thus, if the first two items reported were the last two list items, and if their report was followed by contiguous report of the early list items, then the lifo score for that output order was 2 (e.g., 6712345 or 7612345). Likewise, if the first four items reported were from the end of the list, then the lifo score for that output order was four (e.g., 4567123).[6] To illustrate, the earlier Figure 7 showed serial position curves for several output orders with lifo scores of 1 through 4; note, however, that in addition to the orders shown in that figure, there are numerous others with identical lifo scores, because lifo is not sensitive to the order in which items from the recency portion of the list are reported.

We conducted a regression analysis with the (logarithmically transformed) usage proportions as the criterion variable (to illustrate, Tables 1 and 2 show those usage proportions for a subset of output orders), and lifo and path cost, as described in the previous section, as potential predictors (using a common estimate of $w = 3.78$ to compute path costs for both retention intervals). Data from both retention intervals were considered simultaneously.

Table 3
Results of Regression Analyses to Predict (Logarithm of) Usage Proportion of Output Orders Across Both Retention Intervals Simultaneously

Model	Predictors	Estimate	t	p	Adjusted r^2
1	Path length	−.32	−6.62	<.0001	.44
2	Path length	−.30	−6.82	<.0001	.53
	lifo	.18	3.16	.003	
3	Path length	−.34	−7.55	<.0001	.52
	I-lifo	.26	3.19	.002	
4	Path length	−.32	−7.19	<.0001	.54
	I-lifo	.29	3.51	.001	
	D-lifo	.12	1.75	<.10	

Note—*df*s are 53, 52, 52, and 51, respectively, for Models 1 through 4.

The results from four relevant regression models are summarized in Table 3. The first model simply restates the results of the preceding path cost analysis in regression terms (and aggregating across retention intervals) and serves as a comparison point. The second model confirms that lifo contributed a significant additional amount of variance, suggesting that people sought to exploit the recency advantage in addition to minimizing path length. The third and fourth models are of particular interest because they examined the role of lifo for immediate and delayed reconstruction separately. The *I-lifo* and *D-lifo* variables selectively coded the effect of lifo for immediate and delayed reconstruction, respectively (and were dummy coded as 0 for the other retention interval). A comparison of Models 2–4 clarifies that lifo had an effect for immediate but not for delayed reconstruction. There are several sources of support for this conclusion. First, the proportion of variance explained by lifo across both retention intervals was virtually equal to that explained by I-lifo alone (compare r^2 for Models 2 and 3). Second, adding D-lifo as a third explanatory variable to path length and I-lifo did not increase the proportion of explained variance (compare Models 3 and 4).

The regression analysis converges on the conclusion that in addition to minimizing path length, people prefer to report items from the recency portion of the list first, but they do so only when reconstruction is immediate. When there is a delay that reduces the temporal-distinctiveness advantage of the recency items, they are no longer preferentially reported first. Together, those two variables capture more than half of the total variance in people's choice of different output orders.

Accuracy and Output Order
This final analysis considered the relationship between output order and accuracy by computing the mean correct-in-position performance (averaged across participants and serial positions) separately for each output order (with $n \geq 7$, as was noted earlier). The correlations failed to reach significance for both immediate ($r = .24$) and delayed ($r = .04$) reconstruction.

We next repeated the preceding regression analysis, but using proportion correct rather than frequency of use as the criterion. Across all four models, only a single effect straddled significance: In Model 3 (see Table 3 for an explana-

Table 4
Usage and Accuracies for Different Classes of Output Orders for Both Delay Conditions

Output order(s)	Immediate			Delayed		
	Frequency	Proportion	Accuracy	Frequency	Proportion	Accuracy
Begin with first item	446	.31	.56	573	.40	.55
Begin with last item	218	.15	.53	135	.09	.34
Exploit recency	518	.36	.65	396	.28	.54

Note—Only orders that occurred seven or more times at either retention interval are considered.

tion of models), the effect of I-lifo was .03 [$t(51) = 1.99$, $p = .052$], suggesting that in immediate reconstruction, accuracy could be boosted by reporting recency items first. This conclusion is confirmed by the summary in Table 4 (orders that exploited recency were defined as before). It is clear that those report orders led to greater accuracy in immediate but not in delayed reconstruction (see Tables 1 and 2 for a breakdown by specific output orders).

These results are entirely in accord with the expectations of temporal distinctiveness cited at the outset; early report of recent items maximizes accuracy because it prevents those items from receding into the temporally crowded past. Of course, if retrieval does not commence until after a considerable delay, then those same items have already lost some of their distinctiveness, thus rendering their early report less advantageous. Indeed, Table 4 shows that for delayed reconstruction, initiating retrieval with the last item even engendered a clear accuracy deficit as compared with initiating retrieval with the first item. This result is consistent with the preceding section, which showed that minimizing path cost is an overriding consideration in delayed reconstruction: Initiating retrieval with the last item necessarily involves a large cost because it inevitably entails at least one (large) backward transition, and this cost may give rise to the observed accuracy disadvantage.

DISCUSSION

Summary of Results

Our principal results are readily summarized. When people were free to choose the order in which to report items from short-term serial order memory, their output sequences satisfied two, sometimes competing constraints.

First, there was a memorial cost associated with transitions between serial positions. That cost was greater for backward than for forward transitions, and it increased with transition lag. Accordingly, people consistently sought to minimize that path cost. This tendency was invariant across retention intervals.

Second, a clear memorial benefit resulted from reporting temporally distinct items first (where increased temporal distinctiveness can result either from recency or from isolation during presentation). People therefore preferentially reported distinct items earlier than they did temporally crowded items. If the items' distinctiveness resulted from their relative isolation, this preference for early report persisted even after a notable retention interval. By contrast, if the items were distinct because they were presented late on the list, the tendency to report them

early declined with delay, and so did the accuracy advantage associated with early report.

Theoretical Implications

All current models of short-term serial order memory assume that retrieval proceeds autonomously once initiated, with no role for online decisions about output order. That is, irrespective of whether models rely on temporal context (Brown et al., 2000, 2007; Burgess & Hitch, 1999), positional markers (Henson, 1998; Lewandowsky & Farrell, 2008), or a primacy gradient alone (Farrell & Lewandowsky, 2002; Page & Norris, 1998), current theories are at least tacitly ballistic.

Our results, by contrast, clearly suggest that people's choice of output order is not ballistic but quite heterogeneous. The data thus underscore the need to develop theories that can explain people's heterogeneous—but nonetheless lawful and principled—retrieval strategies. One model that points in this direction is the SAM theory of free recall (Raaijmakers & Shiffrin, 1981), in which retrieval is clearly not ballistic.

To what extent, then, can mechanisms in existing models nonetheless shed light on the present results? Consider first the role of temporal distinctiveness. Items can be temporally distinct for two quite different reasons: They can be distinct because they were presented recently and therefore have not yet receded into the temporally crowded past, or they can be distinct because they were temporally separated from their list neighbors. A critical attribute of recency-based distinctiveness is that it diminishes over time: Accordingly, we found that people preferentially commenced retrieval with recent items in the immediate condition but not in the delayed condition. Moreover, whereas there was an accuracy advantage associated with reporting recent items first in immediate reconstruction, commencing retrieval with the last item was disadvantageous in delayed reconstruction. It follows that people are generally finely tuned to the memorability of recency items and structure their report accordingly. Isolation-based distinctiveness, by contrast, withstands the passage of time because even as all events become more temporally crowded, the relative advantage of isolated items persists. Accordingly, people preferentially reported isolated items early (and more accurately), irrespective of delay.

Turning to the role of path length, we found clear evidence that people seek to minimize the total distance of transitions between successively reported items. Accordingly, output orders that involve numerous small transitions were more likely to be chosen than were output orders with large transitions. Moreover, people clearly preferred

forward to backward transitions, and this was again reflected in the relative paucity of output orders involving backward transitions—with the exception of reports that exploited the recency advantage, as was just discussed. Our results thus extend the finding of Romney et al. (1993) from semantic clustering in long-term memory to serial retrieval in short-term memory, suggesting that people seek to economize their travel through memory space.

Why do people prefer small transitions over larger ones, and hence shorter path lengths through memory space over longer ones? The finding appears to sit well with multidimensional models of memory, according to which items occupy locations within a memory space in which items that are similar to one another along various dimensions (e.g., semantic, phonological, temporal, positional) tend to be stored in close proximity to one another. Given the intuitively reasonable assumption that there is a cost (in either time or error probability) to making larger jumps through such a memory space, it is unsurprising that participants attempt to minimize total transition distance during retrieval. Consistent with this idea, our data show that there is an accuracy advantage associated with small transitions (see also Nairne et al., 2007), as well as a preference for output orders that involve shorter overall path lengths. Analogous results have been found in other tasks: In absolute identification, for example, a stimulus presented on trial n—Sn— is identified both faster (Petzold & Haubensak, 2001) and more accurately (Luce, Nosofsky, Green, & Smith, 1982) if its magnitude is similar to that of stimulus $Sn-1$ (i.e., the item presented on the immediately preceding trial).

The idea of a multidimensional memory space provides a natural perspective on the observed tendency to prefer output orders that minimize path length, but by itself does not offer an account of the observed forward bias. We can draw on two existing theoretical proposals that may explain the forward bias. In free recall, the TCM model of Howard and Kahana (2002) offers an account of the forward bias based on the notion of a "retrieved context." Essentially, the idea is that when a given item is retrieved, that item's encoding context is also retrieved, and that context in turn provides a relatively strong cue for the item that followed the just-retrieved item on the list. However, Farrell and Lewandowsky (2008) noted some potential difficulties with the TCM approach. An alternative perspective on the forward bias was provided by the temporal sampling model sketched by Brown et al. (2008). According to their model, items can be viewed as occupying *segments*—rather than points, as in temporal-distinctiveness models such as SIMPLE— of a time line. Because the time axis is logarithmically compressed, recent items occupy larger segments than do more temporally distant items. On the additional assumption that retrieval proceeds by sampling evidence along the temporal axis, a temporal-sampling model of this type predicts a forward bias in recall. That is, when item n is being cued, evidence will also be accumulating for items $n+1$ and $n-1$. However, because item $n+1$ (being more recent) occupies a larger proportion of the time line than does item $n-1$, it will accumulate more evidence and is thus most likely to be recalled next.

We have sketched out how existing theoretical ideas may account for most aspects of our data in isolation. However, a full account of the present data awaits the development of a new class of models that depart from conventional ballistic modeling and replace it with a more flexible approach to retrieval.

REFERENCES

Beaman, C. P. (2002). Inverting the modality effect in serial recall. *Quarterly Journal of Experimental Psychology*, **55A**, 371-389.

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

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.

Brown, G. D. A., Chater, N., & Neath, I. (2008). Serial and free recall: Common effects and common mechanisms? A reply to Murdock (2008). *Psychological Review*, **115**, 781-785.

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

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, **107**, 127-181.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, **106**, 551-581.

Cowan, N., Saults, J. S., Elliott, E. M., & Moreno, M. V. (2002). Deconfounding serial recall. *Journal of Memory & Language*, **46**, 153-177.

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

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.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, **9**, 59-79.

Farrell, S., & Lewandowsky, S. (2008). Empirical and theoretical limits on lag-recency in free recall. *Psychonomic Bulletin & Review*, **15**, 1236-1250.

Fitts, P. M. (1954). The information capacity of the human movement system in controlling the amplitude of movement. *Journal of Experimental Psychology*, **47**, 381-391.

Henson, R. N. A. (1998). Short-term memory for serial order: The Start–End model. *Cognitive Psychology*, **36**, 73-137.

Hintzman, D. L., Block, R. A., & Inskeep, N. R. (1972). Memory for mode of input. *Journal of Verbal Learning & Verbal Behavior*, **11**, 741-749.

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

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, **24**, 103-109.

Kahana, M. J., & Caplan, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition*, **30**, 841-849.

Lewandowsky, S. (1986). Priming in recognition memory for categorized lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 562-574.

Lewandowsky, S. (1999). Redintegration and response suppression

in serial recall: A dynamic network model. *International Journal of Psychology*, **34**, 434-446.

LEWANDOWSKY, S., & FARRELL, S. (2008). Short-term memory: New data and a model. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 49, pp. 1-48). London: Elsevier.

LEWANDOWSKY, S., & MURDOCK, B. B. (1989). Memory for serial order. *Psychological Review*, **96**, 25-57.

LEWANDOWSKY, S., NIMMO, L. M., & BROWN, G. D. A. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory & Language*, **58**, 415-428.

LUCE, R. D., NOSOFSKY, R. M., GREEN, D. M., & SMITH, A. F. (1982). The bow and sequential effects in absolute identification. *Perception & Psychophysics*, **32**, 397-408.

MCKOON, G., & RATCLIFF, R. (1980). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning & Verbal Behavior*, **19**, 369-386.

NAIRNE, J. S. (1992). The loss of positional certainty in long-term memory. *Psychological Science*, **3**, 199-202.

NAIRNE, J. S., CEO, D. A., & REYSEN, M. B. (2007). The mnemonic effects of recall on immediate retention. *Memory & Cognition*, **35**, 191-199.

OBERAUER, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory & Language*, **49**, 469-483.

PAGE, M. P. A., & NORRIS, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, **105**, 761-781.

PELLI, D. G. (1997). The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.

PETZOLD, P., & HAUBENSAK, G. (2001). Higher order sequential effects in psychophysical judgments. *Perception & Psychophysics*, **63**, 969-978.

RAAIJMAKERS, J. G. W., & SHIFFRIN, R. M. (1981). Search of associative memory. *Psychological Review*, **88**, 93-134.

ROMNEY, A. K., BREWER, D. D., & BATCHELDER, W. H. (1993). Predicting clustering from semantic structure. *Psychological Science*, **4**, 28-34.

SCHWARTZ, G., HOWARD, M. W., JING, B., & KAHANA, M. J. (2005). Shadows of the past: Temporal retrieval effects in recognition memory. *Psychological Science*, **16**, 898-904.

TAN, L., & WARD, G. (2007). Output order in immediate serial recall. *Memory & Cognition*, **35**, 1093-1106.

NOTES

1. Contrary to an apparent consensus in the literature, the preference for short over long lags is not universal. In a reanalysis of numerous free recall studies, Farrell and Lewandowsky (2008) showed that transition probabilities often increase again for extremely large lags. A more appropriate summary of people's transition behavior would therefore cite a strong preference for immediate transitions (lags ± 1) followed by extreme lags (i.e., approximately five or greater), and finally transitions involving intermediate lags (i.e., approximately three to five).

2. An additional analysis confirmed previous findings that temporally isolated items were more accurately recalled. This analysis involved a multilevel regression that regressed performance on items in Positions 2, 4, and 6 onto total temporal isolation (i.e., the sum of the intervals surrounding each of those items). The analysis revealed a significant isolation effect overall, with each second of temporal isolation boosting performance by .04 [$t(23) = 3.09$, $p < .001$]. In two further separate analyses, the isolation advantage was found to be of equal magnitude (.04) for both retention intervals. We do not consider the effects of temporal isolation on performance further.

3. We also considered the effects of absolute temporal isolation by examining the intervals preceding and following an item as predictors in two additional analyses. These analyses gave rise to the same conclusions and are not reported in detail.

4. Transitions involving more than two intervening list positions (not shown here) were associated with increasing accuracy due to confounds with primacy effects, leading to a nonmonotonic function when longer transitions were considered. In the present data, this advantage for greater lags was particularly strong for backward transitions. For example, the only way in which a lag of -6 can occur is by a transition from the last to the first list item, and the first list item is of course retrieved with great accuracy. (Likewise, for forward transitions, larger lags necessarily involve recency items that are also retrieved more accurately.)

5. There were 5,040 theoretically possible orders (i.e., permutations of seven items), but the total number of trials across all participants was 2,880.

6. Most report orders can be unambiguously assigned a lifo score. One exception involves strict forward report, which could plausibly be considered to have a lifo score of either 0 or 7. For the analyses reported here, we assigned a score of 7 to forward report; however, none of the conclusions are altered if that order is assigned a 0 instead.

APPENDIX
Details of Lag-CRP Analysis

In free recall, it is common that one considers only the first two output positions in CRP analyses; see Farrell and Lewandowsky (2008) for details. The focus on the first two output positions is particularly warranted in the present experiment, in which the use of short lists rapidly constrained the range of available lags across output positions. We considered only trials on which the first two reconstructions were correct (the results do not change appreciably if erroneous responses are included), and the CRPs were computed by taking into account the number of possible transitions. Specifically, for each participant p, two vectors were updated across trials for each retention interval separately. Both vectors contained $2L - 1$ elements, where L is the list length, with each element representing a lag of a particular displacement—namely, from $-(L - 1)$ to $+(L - 1)$. The first vector, $\mathbf{r}_p$, kept track of the frequency of transitions of various lags between the first two responses by incrementing the corresponding cell by 1. The second vector, $\mathbf{d}_p$, was a denominator vector and kept track of all lags that could have been produced given the serial position of the first response (e.g., if the second list item was reconstructed first, then the possible lags for the second output position ranged from -1 to $+5$). All possible lags in $\mathbf{d}_p$ were incremented by 1 whenever a single cell in $\mathbf{r}_p$ was incremented. The lag-CRP for a participant was then obtained by dividing $\mathbf{r}_p$ by $\mathbf{d}_p$ in an elementwise fashion, and the average of individual lag-CRPs across participants yielded the mean lag-CRP functions shown in Figure 4. (For plotting, the probabilities were normalized to sum to unity across all possible lags for each delay separately.)