Journal of Experimental Psychology: Learning, Memory, and Cognition

Reconstruction of Temporal and Spatial Order Information

Madison D. Paron, Alice F. Healy, and Michael J. Kahana

Online First Publication, January 15, 2026. https://dx.doi.org/10.1037/xlm0001575

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Learning, Memory, and Cognition

ISSN: 0278-7393

https://doi.org/10.1037/xlm0001575

Reconstruction of Temporal and Spatial Order Information

Madison D. Paron[1], Alice F. Healy[2], and Michael J. Kahana[1]
[1] Department of Psychology, University of Pennsylvania
[2] Department of Psychology and Neuroscience, University of Colorado Boulder

A reconstruction-of-order task illuminated the dynamics and strategies that underlie serial order recall. An initial benchmark experiment, either with no variation in spatial positions or with spatial positions coinciding with temporal positions, yielded bowed symmetrical serial position functions in each case, consistent with both simple chaining and simple positional coding models. In contrast, these simple models were challenged by two additional experiments, which orthogonally varied temporal and spatial positions. These experiments yielded large performance differences between recalling temporal and spatial information. In the *temporal* condition, participants attempted to reconstruct the temporal order of words that were positioned alphabetically within a vertical array. In the *spatial* condition, participants attempted to reconstruct the spatial positions of words presented in a temporal sequence based on their alphabetical order. After viewing each list, all the words appeared alphabetically, and participants reconstructed the order of the words according to their instructed condition. Compared to temporal recall, spatial recall exhibited superior performance and a more bowed symmetrical serial position function. Analyses showed the effects of temporal contiguity in the spatial condition and spatial contiguity in the temporal condition. These findings suggest the theoretical conclusion that participants do not focus on the words' identities but rather on the temporal–spatial pattern in which the words occur during the study display (i.e., the temporal sequence of the spatial locations in which the words are shown).

Keywords: serial position effects, reconstruction of order, temporal order, spatial order, serial learning

A crucial function of the human memory system is to encode and retrieve information in serial order. The purpose of the present study is to illuminate the procedures and strategies individuals use to retain serial order information. Toward this end, we analyze the temporal dynamics in a novel reconstruction task by examining the precise order and timing of responses, which has not been done previously.

There has been much empirical research on serial order memory (see the reviews by Healy & Bonk, 2008; Hurlstone, 2024; Kahana, 2012, Chapters 8 and 9); however, this research has been primarily limited to a few experimental paradigms. The most common of these paradigms is the standard serial recall task. In this task, participants encode a list of items presented sequentially on a computer screen (or through auditory presentation). Subsequently, either immediately or after a delay, participants attempt to recall the items in the order they were presented. However, this task has one major limitation: Participants often terminate their recall after

Ronaldo Vigo served as action editor.
Alice F. Healy https://orcid.org/0000-0002-1312-9757
The data, materials, and code for this study are available at https://memory.psych.upenn.edu/Data_Archive, and the experiments were not preregistered. These experiments were approved by the institutional review board of the University of Pennsylvania (Approval No. 828063). Prior to participating, participants completed a consent form that was approved by the University of Pennsylvania Institutional Review Board. No identifying information was kept concerning the participants, and all participants consented to have their response data used for research purposes. Preliminary initial versions of this research were presented as a poster at the 2022 Meeting of the Context and Episodic Memory Symposium and as an oral presentation at the 2023 Meeting of the Society of Experimental Psychologists, both in Philadelphia.

The authors declare that there are no conflicts of interest or competing interests. This work was partially funded the National Institutes of Health, Grant MH55687 to the University of Pennsylvania (Michael J. Kahana, principal investigator).

The authors thank Connor Keane for help with the experimental design, programming, and testing of participants in Experiment 1. The authors also appreciate the extremely helpful comments and data analyses from Geoff Ward in response to previous versions of this article. Preparation of this article took place while Alice F. Healy was a visiting scholar in the computational memory laboratory of Michael J. Kahana at the University of Pennsylvania. The authors are grateful for the valuable feedback from the members of the computational memory laboratory concerning this research.

Alice F. Healy and Madison D. Paron drafted the article, and Michael J. Kahana edited it. Madison D. Paron and Alice F. Healy conducted the data analyses. Alice F. Healy and Michael J. Kahana designed the experiments. Madison D. Paron programed and tested the participants in Experiments 2 and 3; in addition, she created all of the figures. All authors approved the scientific content of the final article.

Madison D. Paron played a lead role in data curation, investigation, and visualization, a supporting role in methodology, and an equal role in formal analysis and writing–original draft. Alice F. Healy played a lead role in conceptualization, supervision, and writing–original draft and an equal role in formal analysis and methodology. Michael J. Kahana played a lead role in funding acquisition, a supporting role in supervision and writing–review and editing, and an equal role in conceptualization and methodology.

Correspondence concerning this article should be addressed to Alice F. Healy, Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO 80309-0345, United States. Email: alice.healy@colorado.edu

making one or two errors, resulting in missing information regarding their knowledge of later list items (or the earlier list items in the case of backward recall; Dougherty et al., 2023).

The reconstruction-of-order task is an alternative method for studying serial recall that addresses the aforementioned limitation. In this task, participants also study a list of individually presented items. During the test phase, all the items in the list are revealed to the participants in a random or predictable order, such as alphabetical order. The participants' objective is to reorder the items, putting them into their correct list order. This task enables researchers to assess participants' memory for order without necessitating the recall of the items themselves (cf. Neath, 1997, who proposed that the reconstruction-of-order task does not provide a functionally pure measure of order memory).

A related issue is the crucial distinction between item and order memory. In terms of item memory, for example, participants might remember that the last word in the list was related to clothing but struggle to recall whether it was specifically a shirt or a jacket, resulting in a failure to recall the word itself. However, if provided with the word, they might accurately recall its position in the list.

The reconstruction-of-order task effectively addresses these related issues. However, investigators who used this task typically examined just the final placements of the words, leading to a lack of detailed information regarding the precise order and timing of participants' responses. Without this information, the investigators did not know the cues participants used to retrieve the memory and did not have an understanding of the dynamics of the task (refer to Table 1 for a summary and examples of these problems).

Hence, a primary goal of the present study was to combine the best of serial recall and reconstruction-of-order methods by examining the precise order and timing of responses in the reconstruction-of-order task. The complexity of the reconstruction-of-order task necessitated a more intricate method for recording participants' responses. Leveraging current computing technology, we developed a computer-gamelike version of the reconstruction-of-order task. Participants utilized a computer mouse to rearrange the words, and the movement of each word was recorded.

Comparisons between reconstruction of order and free recall tasks provide a useful analogy for appreciating task dynamics in reconstruction of order. In free recall, one can decompose the serial position curve into a series of item selections from memory.

Table 1
Methods for Studying Serial Order Memory, With Problems, Advantages, and Examples

Method
• Serial recall Study: Cat, tree, book, rose, hat, chair, peach, shoe Recall: Cat, tree, rose, book (stop) Problem 1: Requires both item and order memory Problem 2: Stop recalling before the end of the list
• Reconstruction of order Study: Cat, tree, book, rose, hat, chair, peach, shoe Given at test: Book, cat, chair, hat, peach, rose, shoe, tree Recall: Cat, tree, rose, book, shoe, chair, peach, hat Advantage 1: Never stop recalling before the end of the list Advantage 2: Isolates order information Problem 1: Most prior studies only consider final placements Problem 2: Most prior studies lack timing information

Initially, the participant selects the first response, leading to a distribution of first responses (the probability of first recall, PFR, function). Then, the participant makes a series of subsequent responses, each heavily influenced by the preceding responses, highlighting the role of contiguity. Investigators assess this dependence by employing conditional response probability measures as a function of the lag between items (lag-CRP; Solway et al., 2012).

In a reconstruction-of-order task, the first-word placement is akin to the PFR in free recall. However, this initial placement possesses an additional dimension—the judged position, which can be either exactly correct or from a nearby or more distant location. Subsequent selections are likely to bear some relationship to the just-selected item, for example, coming from neighboring items along the to-be-remembered dimension or along another dimension.

Consequently, it would be intriguing, in line with the research conducted by Ward et al. (2010, exploring the connection between free recall and serial recall), to investigate the extent to which reconstruction of order exhibits organizational principles similar to those uncovered in free recall and serial recall tasks. The advantage of this method is that it enables the observation of both the sequence of selections and, for each one, the accuracy of the retrieved information.

The Present Study

To capitalize on the benefits of the reconstruction-of-order task and overcome its limitations, we developed a novel computerized version of the task. This new version allows for the examination of the temporal dynamics involved in the reconstruction of order.

Research on serial learning yielded asymmetrical bow-shaped serial position functions, such as those Murdock (1960) reported from Bugelski (1950), in which participants did a serial learning task on a list of eight nonsense syllables. Murdock's distinctiveness model, which is based on the relative distinctiveness of the serial positions and involves log functions resulting in asymmetry, nicely fits these serial position functions, and it was subsequently successfully applied to the retention of spatial as well as temporal positions (Healy et al., 2008). Healy (1974) developed procedures to disentangle item and order information with four-item lists (but again, see Neath, 1997). Healy's findings indicated symmetric bow-shaped serial position functions in order-only recall, which used a reconstruction-of-order paradigm. In contrast, she observed flatter functions in item-only recall.

The present study examines the characteristics of serial position functions with longer lists in the reconstruction-of-order task. Specifically, the study aims to determine whether the serial position functions are bow-shaped and whether they are symmetrical or asymmetrical. Furthermore, the study compares the recall of temporal and spatial order and examines whether the serial position functions differ for the two types of order information. If distinct serial position functions occur for temporal and spatial order recall, this result suggests that a simple positional coding model for serial recall is not appropriate (see the classic models discussed by Healy & Bonk, 2008; Hurlstone, 2024; and Kahana, 2012, which refer to paired associate links between the to-be-remembered items and their ordinal positions in the list). In contrast, if the serial position functions are bowed for temporal order recall but not for spatial order recall, this finding suggests that simple temporal chaining of associations (i.e., associations between neighboring to-be-remembered items) plays a role in shaping the serial position functions due to the temporal contiguity of adjacent positions in temporal order recall but not in spatial order recall. Previous studies examining serial position functions with longer lists demonstrated nearly symmetrical serial position functions for both temporal order recall and spatial order recall, using lists of 18 or 20 nouns (Bowles & Healy, 2003; Sinclair et al., 1997), although these studies involved learning a single serial list through multiple exposures.

The present study also considers whether output order in general, or the PFR more specifically, explains the serial position functions (Lewandowsky et al., 2009; Ward et al., 2010). In addition, the present study explores whether participants' confidence influences the serial position functions by comparing the accuracy and response time (RT) functions. Finally, this study asks whether contiguity (outputting adjacent items together) influences the serial position functions (Solway et al., 2012).

We also include a comparison of conditions in which participants vocalize the stimulus words aloud or read them silently (see Murray, 1966, for an early examination of this comparison, demonstrating an advantage for vocalization, and see Neath, 1997, who demonstrated this modality effect in reconstruction of order, but only on the final list item). This manipulation has been used to study the production effect (e.g., Saint-Aubin et al., 2021), whereby reading aloud shows a memory advantage relative to reading silently. However, it is employed in the present study instead as a window to the use of phonological coding, which has been shown in studies of short-term memory to differentiate the recall of item and order information as well as, more specifically, the recall of temporal and spatial order information (e.g., Healy, 1975a; Healy et al., 1991). We simply make the intuitive assumption that recoding visually presented words to phonological representations is necessary for reading words aloud but not for reading them silently (for similar reasoning, see Kole et al., 2005, although the findings in that study showed that vocalizing to-be-remembered items rather than reading them silently retarded processing of initially presented items but enhanced processing of subsequently presented items).

The present study consists of three experiments that focus on the recall of temporal and spatial order information. Experiment 1 serves as a benchmark study, in which either no spatial information is provided or temporal and spatial information coincide. In contrast, Experiments 2 and 3 are the primary experiments, which independently vary temporal and spatial information and compare recall functions for the two types of information. Like early studies, we examine the serial position functions, and like subsequent studies by Lewandowsky et al. (2009) and Ward et al. (2010), we also examine response output functions, but we go beyond those studies by adding examinations of RT functions (see, e.g., Osth & Farrell, 2019, for examination of RT functions in free recall) and lag-CRP functions (Solway et al., 2012) to help us appreciate the task dynamics and give us a window into the processes and strategies employed by participants for recalling temporal and spatial information.

Experiment 1

Experiment 1 comprised two experimental conditions, *central* and *top-to-bottom*. The central condition was the standard reconstruction of temporal order condition, in which all to-be-remembered words occurred in a fixed central location in the middle of the screen. The top-to-bottom condition introduced a variation in spatial locations and intentionally confounded temporal and spatial orders. In this condition, the to-be-remembered words occurred vertically arrayed in a single column, with the first word in the top location, the second word in the location second from the top, and so on. The final word in the list was in either the eighth, 12th, or 16th location, depending on the length of the list.

Method

Participants

The experiment employed 795 participants tested using Amazon Mechanical Turk (MTurk) and paid $7.50 for their participation. The MTurk configurations ensured that only participants with IP addresses in the United States could participate. In order to exclude participants who cheated, the experimenters discarded all the data from participants whose total accuracy exceeded 95% correct or who admitted in an optional end-of-experiment questionnaire to writing down notes. In addition, the data from one participant were not part of the data analyses due to missing cells for the length 16 lists. No participants were excluded for low accuracy, and the data from participants who did not complete the experiment were never received from Amazon Turk. The high rate of exclusions due essentially to cheating is most likely caused by the use of online participants with no experimenter present to oversee participant conduct. As a result of the exclusions made by the experimenters, data from 466 participants remained for the analysis of recall performance.

Materials and Procedure

The to-be-remembered words were randomly selected nouns from the Toronto Word Pool (Friendly et al., 1982), with a different randomization used for each participant, so every participant had a unique set of lists. For a given participant, the words were selected without replacement for a particular list, ensuring that there were no repeated words in a list.

The experiment consisted of three practice lists and 24 experimental lists. The experimental lists were composed of three different list lengths: 8, 12, or 16 words. These list lengths were presented in a pseudorandom order, with one of each list length in the practice lists and one in each block of three successive trials in the experimental lists. We randomized the order of list lengths for each participant so that participants were uncertain of list length on a given trial.

The presentation rate was 1.4 s per word, with a .4–.8 jittered interstimulus interval (i.e., a 2-s average stimulus onset asynchrony; the stimulus jittering was done because it made use of an experimental code base for electroencephalogram studies in our laboratory. Experiments in the laboratory generally jitter interstimulus intervals to decouple stimulus-evoked neural responses from successive item presentations.)

The instructions said,

> During the course of this experiment, you will see lists of words, which you should try to remember. Following a 10-second countdown period [which was included in order to give participants a short period of time to mentally prepare for the upcoming list], each list of words will be presented visually on the screen, one word at a time for about 2 seconds each.

In the central condition, the instructions specified that each word would appear "in the same central location on the screen." In contrast, in the top-to-bottom condition, the instructions said instead that each word would appear "in a different location on the screen in a vertical column, progressing from the top to the bottom of the column." In this condition, the to-be-remembered words occurred in a single vertical column, with the first word in the top location, the second word in the location second from the top, and so forth. The purpose of comparing the central and top-to-bottom conditions in the present experiment is to examine whether the introduction of spatial information in the top-to-bottom condition influences participants' strategies and performance levels. This comparison serves as a preparation for the investigation of temporal and spatial recall conditions in Experiments 2 and 3.

In the experiment, instructions to half of the participants in each condition stated, "read each word aloud as it appears on the screen" (*aloud*), whereas instructions to the other half of the participants stated, "read each word silently as it appears on the screen" (*silent*). Comparing aloud and silent subconditions helped to determine whether participants use verbal (i.e., phonological) coding of the to-be-remembered words (although we had no way to check on whether participants followed these instructions beyond examining the results of the manipulation). Furthermore, within each subcondition, there were two subgroups of participants: *dual task* and *single task*. The dual-task instructions stated, "As the words are being presented, you should also use your two index fingers to hold down two keys (A with your left hand and L with your right hand)." These instructions aimed to prevent participants from writing down the words. In contrast, the single-task instructions did not include such a requirement. The comparison of single- and dual-tasks subgroups aimed to determine whether preventing participants from writing down words in the dual-task condition would influence participants' strategies and performance levels compared to those in the single-task condition.

During the reconstruction-of-order task (see Figure 1), each word in a given list appeared in a vertical column on the left side of the screen, arranged in alphabetical order. On the right side of the screen, there were eight, 12, or 16 empty boxes, corresponding to the number of words in the list. For list Length 8, only the first eight boxes appeared, and for list Length 12, only the first 12 boxes appeared.

To complete the task, participants used their mouse to select a word on the left side of the screen and drag it to one of the boxes on the right, with the computer emitting a clicking sound when a word was released into a position. The instructions explicitly explained the task as follows:

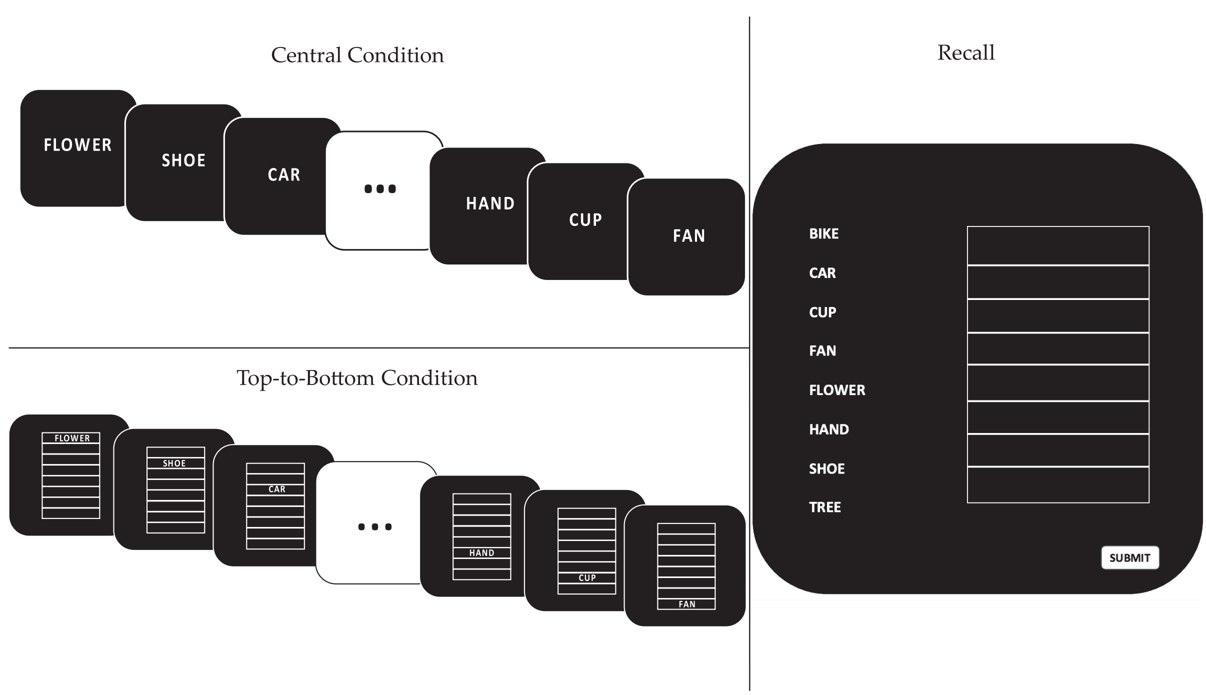
> After each list of words ends, you will see on the left side of the screen, a vertically arranged alphabetical ordering of the words from the list you just saw, and you will see on the right side of the screen a vertically arranged set of empty boxes, one box for each position in the list. Your task is to reconstruct the order of the words in the list you just saw by using your mouse to drag each word on the left to a different box on the right depending on its position in the list. You should place the first word of the list you just saw into the top box, the second word of the list into the second box, and so on, with the last word of the list into the bottom box. Thus, the words in the boxes should end up being in the same order as they appeared in the list you just saw.

The task allowed participants to move a word from one position to another during the response period. In particular, the instructions were,

> You can fill in the boxes in any order you want, not necessarily from top to bottom. Also, you can reorder the words by moving a word from one

Figure 1

Sample Screens Like Those Shown in Experiment 1



box to another box or by temporarily moving a word that you put in one of the boxes on the right back to its location in the alphabetical ordering on the left and later moving it to a new box on the right. After you have filled in all of the boxes on the right and are satisfied with the order of the words in the boxes, you should press the SUBMIT button at the bottom of the screen to submit your ordering and proceed to a new list of words.

Design

The study utilized a $2 \times 2 \times 2 \times 3$ mixed factorial design, in which the first three factors varied between subjects and the last factor varied within subjects. The first factor was order type (central, top-to-bottom), the second factor was word reading (aloud, silent), and the third factor was task number (dual task, single task). The last factor was word list length (8, 12, 16). We assigned the participants to one of the eight between-subjects combinations of order type, word reading, and task number in a pseudorandom order, with every eight participants in a different combination. Although the statistical data analyses included a breakdown by task number (as well as breakdowns by order type and word reading), no interesting results involving task number were revealed, so the results and figures here do not include the breakdown by task number.

Results and Discussion

In considering the results, we examine several different measures of performance, including accuracy, response initiation times (RITs), PFR, and lag-CRPs. These various measures provide insight into different aspects of the participants' memory and the procedures and strategies they used for the reconstruction of order.

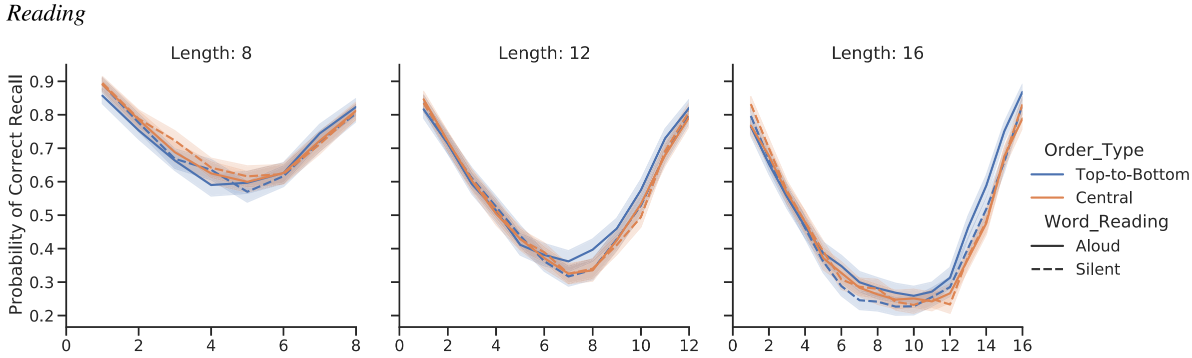
Accuracy

The proportion of correct responses at each serial position served as our measure of accuracy. We scored each word separately and absolutely, rather than using relative scoring, and we considered a word correct if placed into the correct serial position. The resulting scoring yielded serial position functions that were bowed symmetrically, in which performance was best at the beginning (primacy) and end (recency) of the lists of each length (see Figure 2). These bow-shaped functions were generally consistent with prior serial recall studies and both the classic simple chaining and simple positional coding models (see again Healy & Bonk, 2008; Hurlstone, 2024; Kahana, 2012, Chapters 8 and 9, for summaries), although the symmetry is inconsistent with Murdock's (1960) distinctiveness model. The two reading types (aloud, silent) did not differ in accuracy. Likewise, the two order types (central, top-to-bottom) did not differ in accuracy, although a difference might be predicted from the findings of Fischer-Baum and Benjamin (2014), who showed that recall of temporal information was influenced by the orientation of spatial information; specifically, temporal recall was better with the more common left-to-right spatial arrangement of the stimuli than with the less common right-to-left spatial presentation; see also, e.g., Guida & Campitelli, 2019, for a discussion of related effects involving "spatialization" in other paradigms. Finally, the two task numbers (dual task, single task) did not differ in overall accuracy, suggesting that holding down the keys did not prohibit note taking, although there was some evidence that it might depress the magnitude of the recency advantage.

Because there were necessarily different numbers of serial positions for the three list lengths, to compare the serial position functions across list lengths statistically, a combined analysis of variance (ANOVA) on the proportion of correct responses included only six serial positions from each list length, dividing them into beginning, middle, and end positions. Specifically, two serial positions contributed to each level: beginning = Positions 1 and 2 for all list lengths; middle = Positions 4 and 5 for Length 8, Positions 6 and 7 for Length 12, and Positions 8 and 9 for Length 16; and end = the last two positions for all list lengths (Positions 7 and 8 for Length 8, Positions 11 and 12 for Length 12, and Positions 15 and 16 for Length 16). This ANOVA revealed significant main effects of list length, $F(2, 916) = 455.998$, mean squared error (MSE) = 0.019, $p < .0001$, and of beginning-middle-end positions, $F(2, 916) = 734.117$, MSE = 0.084, $p < .0001$, along with a significant interaction of those two variables, $F(4, 1832) = 252.893$, MSE = 0.016, $p < .0001$, reflecting more pronounced effects of beginning-middle-end positions as list length increased from 8 to 12 to 16 positions. There were also significant higher order interactions, which we will

Figure 2

Probability of Correct Recall Responses in Experiment 1 as a Function of List Length, Serial Position, Order Type, and Word Reading



Note. Error bars (shaded regions) represent between-subjects standard errors of the mean (here and elsewhere) to facilitate the comparison of the between-subjects conditions. See the online article for the color version of this figure.

ignore for simplicity, including a significant three-way interaction involving task number, list length, and beginning-middle-end positions, $F(4, 1832) = 2.677$, MSE = 0.016, $p = .0304$, as well as the component two-way interaction of task number and beginning-middle-end positions, $F(2, 916) = 4.826$, MSE = 0.084, $p = .0082$. Another significant effect was the four-way interaction of list length and the three between-subjects variables (order type, word reading, and task number), $F(2, 916) = 4.460$, MSE = 0.019, $p = .0118$.

Response Initiation Times

To compare the times at which participants recalled words across serial positions, we computed a cumulative index of RIT for each response. Similar to the accuracy index, we scored RITs separately for each word that appeared on a given trial, including error responses as well as correct responses, to avoid eliminating participants who made too many errors. RIT is a cumulative time and thus highly correlated with output order. RIT indicates when a participant chose a given word from the alphabetical column (i.e., the time between when the reconstruction task screen appeared and when the participant selected a given word for movement from the left side of the screen). Because participants can change the order of their responses, only the RIT for the first movement of a given word is part of the analysis (this measure should not be confused with interresponse times.)

The RIT graphs show inverse bowed serial position functions (see Figure 3), and they can be viewed as complementary to the functions for accuracy. This pattern suggests that participants tended to output words in order of response confidence, with faster response initiation corresponding to greater accuracy.

For the ANOVA restricted to the beginning, middle, and end positions for all three list lengths (and excluding the factor of task number), the main effects of list length, $F(2, 926) = 374.301$, MSE = 280.139, $p < .0001$, and of beginning-middle-end positions, $F(2, 926) = 278.601$, MSE = 291.653, $p < .0001$, were both significant along with the interaction of those two variables, $F(4, 1852) = 142.899$, MSE = 86.741, $p < .0001$, reflecting more pronounced inverse bow-shaped functions for beginning-middle-end positions as list length and overall RIT increased from 8 to 12 to 16 words.

PFR

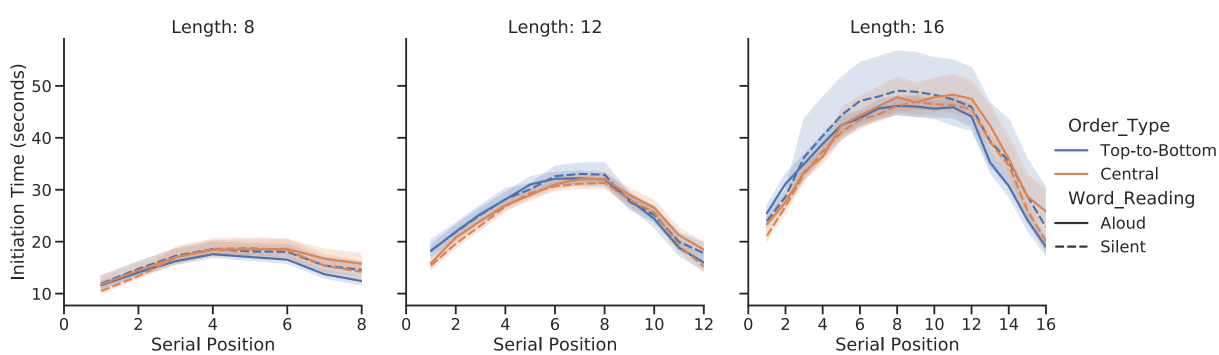
For a second index of output order, we examined each word separately as a function of its serial position to determine whether or not it was the first-word recalled. Like the probability of correct recall function, the PFR function is also bowed and roughly symmetrical (see Figure 4). However, there appears to be some preference for the primacy positions relative to the recency positions for Length 8, with the opposite pattern (preferences for the recency positions relative to the primacy positions) for the longer list lengths. Also, it appears that overall participants either initiated recall with the word in the first serial position or with one of the last four words (cf. Ward et al., 2010). The shape of the PFR function does not depend on order type or word reading (again, see Figure 4), in accordance with the findings for probability of correct recall. Thus, the bowed serial position functions for response accuracy can be explained largely in terms of output order, as might be predicted from the studies by Lewandowsky et al. (2009) and Ward et al. (2010), although Ward et al. found that both primacy and recency effects remain even after conditionalizing the data based on the first recall being from the primacy part of the curve or conditionalizing the data based on the first recall being from the recency part of the curve.

Lag-CRPs

Lag-CRP analyses reflect the probability of transitioning between item i and item i + lag conditional on the possibility that item i + lag could be recalled at that point (Kahana, 2012). Because participants saw all words at all times in the reconstruction task, the participants could recall each word at any time from the current word, item i, either by moving it directly from the list of words into a box or moving it from one box to another. Nevertheless, as with earlier studies using tasks in which participants have different opportunities for recalling words at various lags (e.g., Solway et al., 2012), conditionalizing on the number of opportunities is desirable for this measure because when a word at a given position has already been

Figure 3

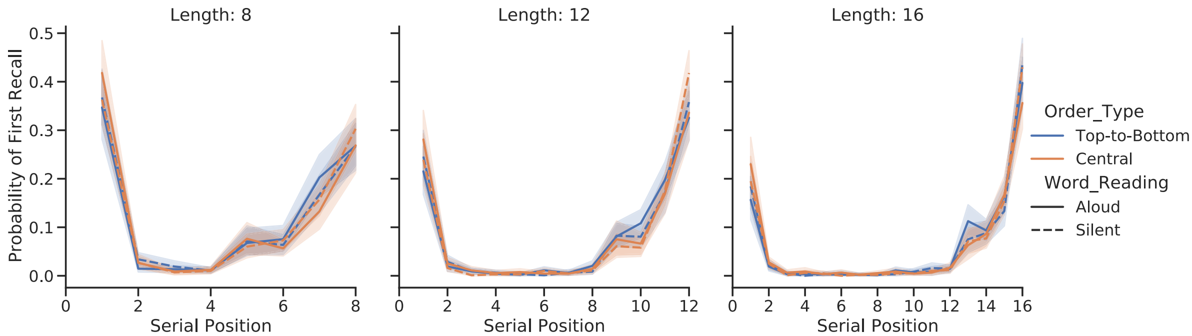
Response Initiation Time (in Seconds) in Experiment 1 as a Function of List Length, Serial Position, Order Type, and Word Reading



Note. Error bars (shaded regions) represent between-subjects standard errors of the mean. See the online article for the color version of this figure.

Figure 4

Probability of First Recall in Experiment 1 as a Function of List Length, Serial Position, Order Type, and Word Reading



Note. Error bars (shaded regions) represent between-subjects standard errors of the mean. See the online article for the color version of this figure.

recalled that position is essentially no longer available for subsequent recalls. The list of words on a given trial occurred in a central location on the screen in the central condition, and the list of words occurred from top to bottom spatially as well as from first to last temporally in the top-to-bottom condition (so the temporal and spatial locations were the same in the top-to-bottom condition). In both conditions, participants showed strong contiguity effects (lag-CRP; see, e.g., Solway et al., 2012) for the temporal lag (i.e., they responded with words that appeared close together in time), suggesting that they responded primarily with words adjacent in the given list and the closer a word was to another, the more likely the two words followed each other in the responses. Note, however, that, contrary to pure associative chaining, sometimes participants who started with words at the end of a list moved not to an adjacent word but rather to a word at the beginning of the list (an extreme negative transition; see Ward et al., 2010, pp. 1213–1215, for a similar finding, which they attribute to a primacy effect). Also, some preference was evident for forward responses (+1 lag from the previous response) rather than backward responses (−1 lag from the previous response; see the top panel of Figure 5). Note that the functions are almost identical in the central and top-to-bottom conditions.

In the bottom panel of Figure 5 is the lag defined in terms of the alphabetical order of the words, their order given on the reconstruction-of-order test. That is, instead of examining the transitions between item i and item i + lag in the list order, we examine the transitions between item i and item i + lag in the alphabetical order. These less bowed functions showed that some, but many fewer, participants used instead the method of responding with words close together in the alphabetical order given on the reconstruction test.

Summary and Conclusions for Experiment 1

At the outset of Experiment 1, there were four primary questions:

1. Are the serial position functions bow shaped? Are they symmetrical or asymmetrical?

There were bowed serial position functions, as in serial learning and recall, but they were symmetrical, unlike in serial learning.

2. Does output order, in general, and the PFR, more specifically, explain the serial position functions?

The serial position functions for the PFR were similar to those for correct recall, being bow-shaped and fairly symmetrical at each list length. Thus, output order can largely explain the serial position functions.

3. Does participants' confidence influence the serial position function (compare accuracy and RIT)?

Participants appeared to output words in order of response confidence (more accurate and faster at the beginning and end of the list).

4. Does contiguity, lag-CRP influence the serial position function?

Participants output adjacent words in sequence, with a preference for outputting the words in a forward (rather than backward) direction, as in studies of forward serial recall (e.g., Dougherty et al., 2023). However, a striking feature of the lag-CRP is that sometimes participants who are starting at the end of the list of words make a transition to the first word in the list (i.e., an extreme negative transition) rather than starting from the first word of the list and making a transition to the last word in the list (an extreme positive transition).

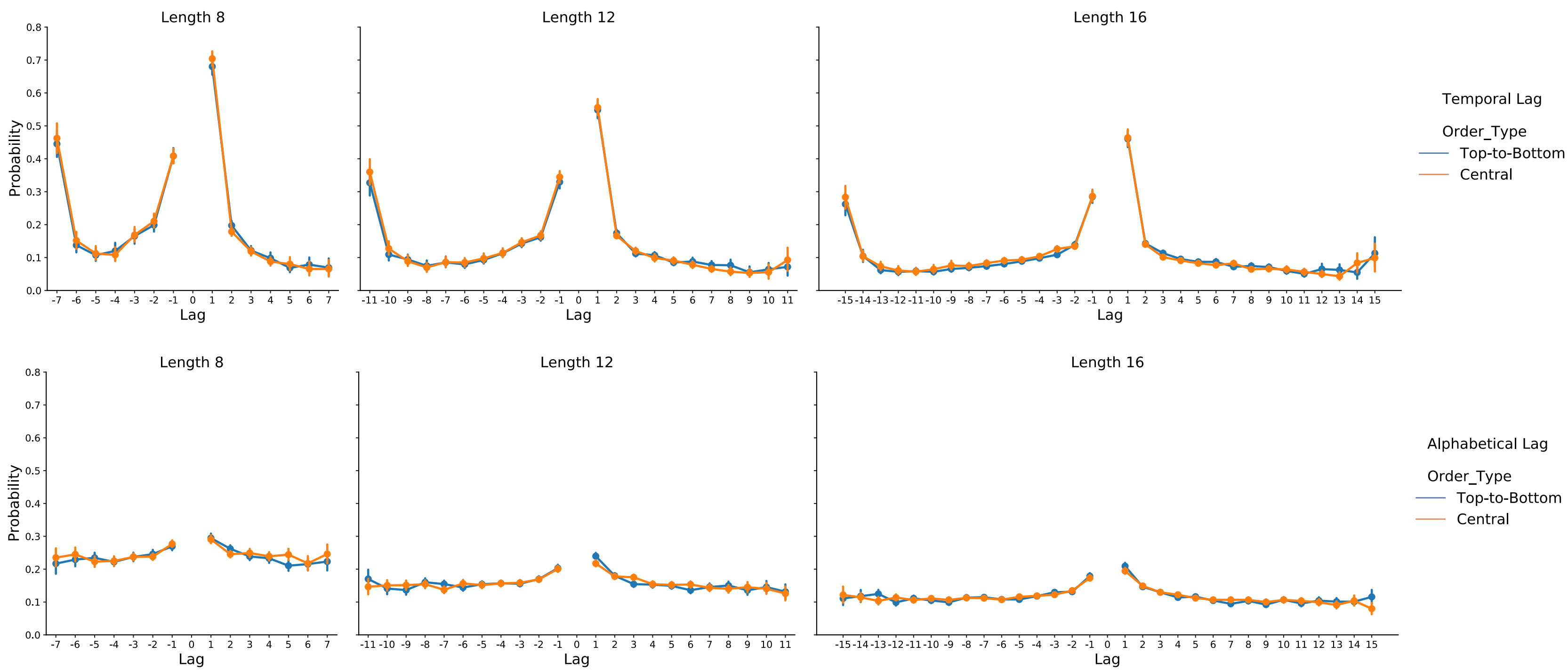
Experiment 2

Most serial recall and serial learning studies are based on the temporal sequence of events, as in Experiment 1. However, there are really two types of order that can be distinguished: temporal sequence and spatial arrangement. In the central condition of Experiment 1, every word appeared in the same central location, so there was no spatial arrangement of the words. In contrast, in the top-to-bottom condition, the temporal sequence coincided exactly with the spatial arrangement, so the first word temporally was also the first word spatially, and so on.

However, in earlier studies, Healy (1975a, 1975b, 1977, 1978, 1982) compared short-term memory for temporal and spatial order information; she varied temporal and spatial orders independently,

Figure 5

Lag-Conditional Response Probabilities in Experiment 1 as a Function of List Length, Order Type, and Type of Lag (Temporal and Alphabetical)



Note. Error bars represent between-subjects standard errors of the mean. However, some error bars were too small to be visible. See the online article for the color version of this figure.

and she found very different results for the two types of order information. For example, she found phonological coding for temporal but not spatial order, more rapid forgetting for temporal than for spatial information, and more bowed serial position functions for temporal than for spatial information. Healy examined very short lists of letters (usually only four letters long), so these differences might be a direct consequence of comparing temporal and spatial information for immediate recall with short lists of letters (see Mandler & Anderson, 1971; Murdock, 1969; Shiffrin & Cook, 1978; and Slamecka, 1967, for other early studies comparing temporal and spatial order recall, and see Bowles & Healy, 2003; Fischer-Baum & Benjamin, 2014; Healy et al., 1991; Nairne & Dutta, 1992; and Sinclair et al., 1997, for more recent comparisons.)

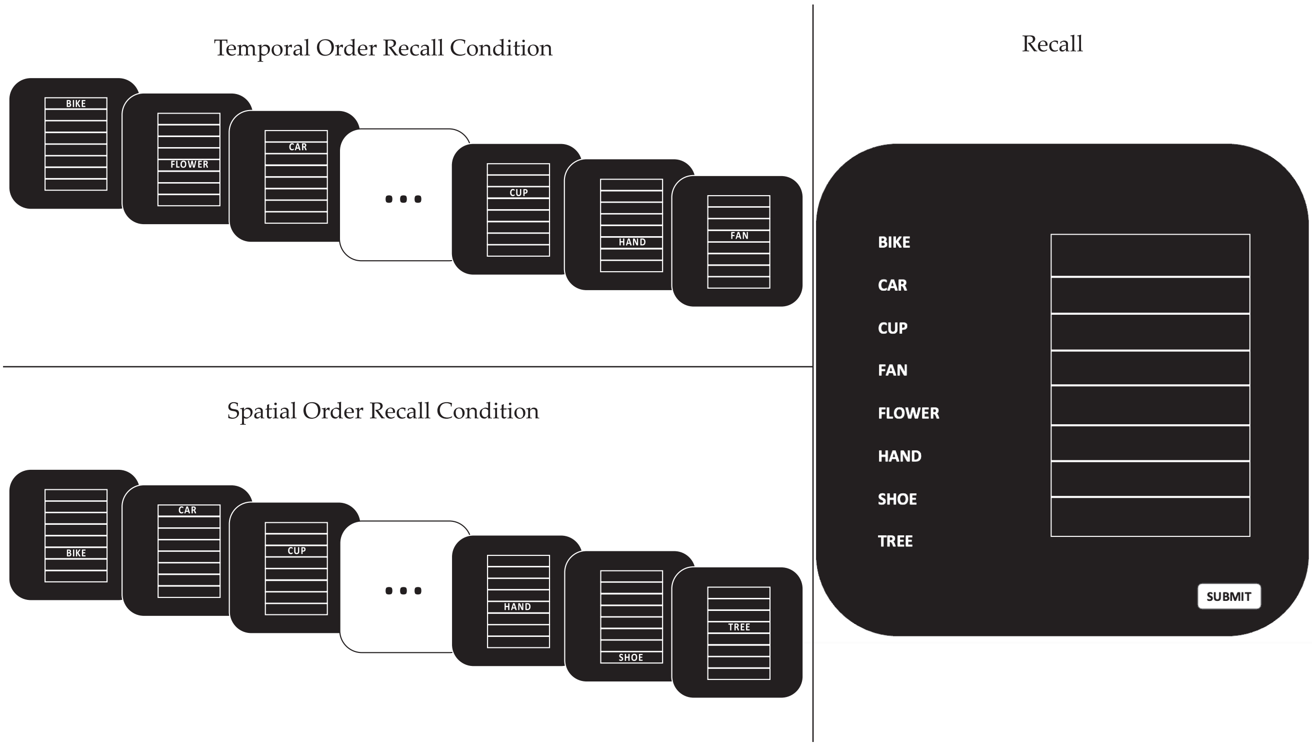
The present study aims to investigate whether the differences observed in previous research between temporal and spatial order recall for short lists of letters extend to longer lists of words: Specifically, the question asked is whether there are bowed serial position functions for temporal but not spatial order recall and whether there is evidence for phonological coding for temporal but not spatial order recall. To assess phonological coding, previous studies have often examined substitution errors, when a to-be-remembered item is replaced by a phonologically similar item (see, e.g., Healy, 1975a). Such substitution errors would be difficult to tabulate in the present study in which each participant received different lists of nouns, and extralist intrusions are impossible in the reconstruction-of-order task. Instead, comparing the aloud and silent subconditions should provide insight into the participants' potentially using a phonological coding strategy.

In addition, the recall dynamics should provide insight into the recall strategies used for the two types of information. Specifically, following Lewandowsky et al. (2009) and Ward et al. (2010), the PFR for each serial position reveals whether the serial position functions can be explained by output order, such that the participants output first the best-recalled positions. In addition, as for Experiment 1, response timing reveals whether accuracy and RT are congruent or whether there is a speed–accuracy trade-off.

Experiment 2 included two different experimental conditions, *temporal* and *spatial*. In both conditions, participants saw words one at a time in a vertical array of spatial locations like that in the top-to-bottom condition of Experiment 1. In the temporal condition, the to-be-remembered order was the temporal order. The spatial order of the words in the temporal condition followed an alphabetical arrangement, which served as an alternate dimension. In contrast, in the spatial condition, the to-be-remembered order was the spatial order, and the temporal order was alphabetical and served as an alternate dimension. Thus, temporal and spatial positions varied orthogonally in both conditions. This experimental design allowed us to investigate the independent effects of temporal and spatial information on participants' ability to remember and reproduce the order of the presented words. To understand this distinction between temporal and spatial conditions, consider a simplified list of three letters, A, B, and C, and three spatial locations arrayed horizontally - - -, with the alphabetical order ABC and the to-be-remembered order BCA. For the temporal condition, the sequence shown would be - B -, - - C, A - -, so the temporal order would be BCA, and the spatial order would be ABC. In contrast, for the spatial condition, the sequence shown would be - - A, B - -, - C -, so the temporal order would be ABC, and the spatial order would be BCA. Experiment 2 used analogous procedures with longer lists of words, where temporal and spatial orders varied independently and where words occurred in alphabetical spatial order in the temporal condition and in alphabetical temporal order in the spatial condition. See Figure 6 for sample screens like those that participants saw. Note that in the recall phase in both temporal and spatial conditions, as in both order-type conditions of Experiment 1, for the reconstruction task, each word in a list appeared in a vertical column on the left, ordered alphabetically. On the right were empty boxes, with the number of boxes matching the list length. Participants selected a word on the left and dragged it to an available box on the right, with reordering allowed.

In this experiment, as in the ABC example, the alternate order was alphabetical. This constraint was imposed as a way to overcome the asymmetry between temporal and spatial order recall noted by Healy (1975a, 1975b), who pointed out that without such a constraint when the two orders are varied orthogonally in a set of experimental trials, temporal order recall has an advantage because the sequence of temporal serial positions necessarily always occurs in a fixed order 1234 (first, second, third, fourth) in a four-item list, whereas the sequence of spatial serial positions occurs in different orders across lists. Presumably, as a consequence of this asymmetrical presentation order of the serial positions (constant in the temporal condition and changing in the spatial condition), order reconstruction was much higher for the temporal than for the spatial condition in Healy (1975b). To eliminate that asymmetry, Healy (1975a) imposed the constraint that the alternate order (spatial in the temporal condition and temporal in the spatial condition) was fixed and known in advance by the participants, so that across all of the experimental trials the temporal serial positions were still given in a constant 1234 order in the temporal condition, but the to-be-remembered items were now given in a constant temporal order in the spatial condition. That particular constraint could not be imposed in the present experiment because of the long list lengths and the fact that the stimulus words varied across experimental trials, but the alphabetical ordering constraint could be imposed, following Sinclair et al. (1997) and Bowles and Healy (2003), both of whom also used longer lists of words as stimuli, so that participants in the present experiment did not know the words in advance but did know that the alternate order of them for all lists in both the temporal and spatial conditions was alphabetical. To see how the alphabetical ordering constraint diminishes or removes the asymmetry between temporal and spatial conditions, let us continue with the ABC example. In the temporal condition, the sequence of the three temporal serial positions is 123 (because the letter in the first temporal serial position is shown first, then the letter in the second temporal serial position, and last the letter in the third temporal serial position), and that sequence of temporal serial positions necessarily occurs on every experimental trial. Using the example given earlier for the spatial condition (- - A, B - -, - C -, with the temporal order ABC and the spatial order BCA), the sequence of the three spatial serial positions is 312, because the letter in the third spatial serial position is shown first, then the letter in the first spatial serial position, and finally the letter in the second spatial serial position. Thus, in the spatial condition, there is a constant temporal order of the letters (ABC, the alphabetical order), which occurs on every trial, whereas both the spatial order of the letters and the sequence of the spatial serial positions vary from trial to trial. With this design, any encoding (intentional or incidental) of the alternate (alphabetical) order during the trial should not affect the learning and retention of the

Figure 6
Sample Screens Like Those Shown in Experiment 2



critical to-be-remembered order (temporal in the temporal condition and spatial in the spatial condition). This control method maintained stimulus attribute independence and, thus, assured that only the stimulus attributes associated with the critical dimension (not those associated with the alternate dimension) could be used as relevant cues for learning the critical order (see Sinclair et al., 1997, for a discussion of this issue).

Method

Participants

The experiment employed 592 participants tested by way of MTurk and paid $7.50 for their participation. As in Experiment 1, the MTurk configurations restricted participants to IP addresses within the United States. For the reasons discussed in the method of Experiment 1, the experimenters excluded the data from an additional 613 participants either for an accuracy level above 95% or for admitting that they wrote notes on the now required end-of-experiment questionnaire. Despite the large number of exclusions in Experiment 1, no attempt was made to reduce the number in Experiment 2 because a comparison of the results of the two experiments was desired, and such a comparison would be misleading if different exclusion criteria were used in the two experiments.

Materials and Procedure

The materials and procedure (e.g., the lists of words and their rate of presentation) were the same as in Experiment 1 except for the differences in the experimental conditions, with the change in experimental conditions from central and top-to-bottom to temporal and spatial, and with all participants in the single-task subgroup (i.e., there were no instructions to hold keys as there were in the dual-task subgroup of Experiment 1). Specifically, the instructions for all participants stated,

> During the course of this experiment, you will see lists of words, which you should try to remember. Following a 10-second countdown period, each list of words will be presented visually on the screen, one word at a time for about 2 seconds each, with each word in a different location on the screen in a vertical column.

In the temporal condition, the instructions stated,

> The spatial arrangement of the words will always be alphabetical, with the word at the top of the column always first in alphabetical order and the word at the bottom of the column always last in alphabetical order, but the temporal sequence in which the words occur will vary across trials, and your primary goal is to learn and reconstruct the temporal sequence.

In contrast, in the spatial condition, the instructions stated,

> The temporal sequence of the words will always be alphabetical, with the first word shown always first in alphabetical order and the last word shown always last in alphabetical order, but the spatial arrangement in which the words occur will vary across trials, and your primary goal is to learn and reconstruct the spatial arrangement.

It is important to note that although the instructions for the two conditions were as similar as possible, in each case conveying the fact that participants were to learn the positions of words in a list, the instructions made the tasks specific to the respective conditions, with the alternate order clearly specified as alphabetical.

Note that, as in the top-to-bottom condition of Experiment 1, each word occurred within a box in Experiment 2, with empty boxes continually shown during the word presentation phase to facilitate the participants' determining the specific location of each word in the vertical array (i.e., its spatial position in the list).

Note also that, unlike Experiment 1, in both conditions of the present experiment, there are two conflicting sources of order information, involving the temporal and spatial dimensions, with one source identified as primary and varying across trials (temporal in the temporal condition and spatial in the spatial condition) and the alternate source identified as secondary and fixed across trials (spatial in the temporal condition and temporal in the spatial condition). The to-be-remembered order is always random because the to-be-remembered words were randomly selected nouns from the Toronto Word Pool, whereas the alternate order is always alphabetical, and the alphabetical order is always provided at test, as in Experiment 1, in a vertical column on the left side of the screen. Thus, the designs of the two experimental conditions are parallel and complementary with respect to the number of dimensions and their properties. In both cases, the addition of a conflicting source of information was expected to lower overall performance levels relative to those obtained in Experiment 1, which was without any conflicting source of information.

Design

The design of Experiment 2 was the same as that of Experiment 1 except for the change in experimental conditions, eliminating the factor of task number, and changing the order-type factor from central and top-to-bottom to temporal and spatial. As in Experiment 1, Experiment 2 also contained *aloud* and *silent* subconditions that varied between subjects. By maintaining the basic experimental design while altering the specific conditions, Experiment 2 aimed to provide insights into the effects of temporal and spatial order information as well as the influence of verbalization on recall performance.

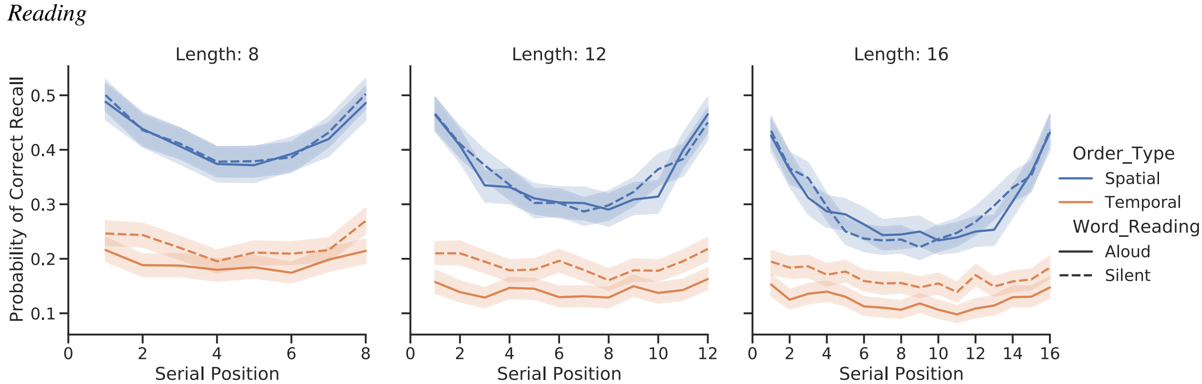
Results and Discussion

Accuracy

As in Experiment 1, the serial position functions indicate the proportions of correct responses for each list length (see Figure 7). The overall recall rate was relatively low, and the spatial condition was more accurate than the temporal condition, although the temporal condition was above chance ($1/n$, with n = list length). This poor performance likely reflects interference caused by competing spatial and temporal information. Contrary to predictions based on Healy's (1975a, 1977, 1978, 1982) short-term memory studies, the spatial condition showed a perfectly bowed symmetrical serial position function. In contrast, the recall function for the temporal condition was flatter. These findings challenge the classic simple positional coding models, which do not differentiate between temporal and spatial positions so seem to assume that positional coding should be equivalent for recall in the temporal and spatial conditions. However, on the assumption that chaining associations for serial recall (i.e., linking each word to the word following it) is more likely with temporally contiguous words, finding a flatter serial position function for the temporal than for the spatial condition suggests that chaining is also not the crucial mechanism underlying the bowed functions. A combined ANOVA on the proportion of correct responses restricted to six serial positions from each list length, dividing them into beginning, middle, and end positions, as in Experiment 1, yielded a substantial main effect of order type, $F(1, 588) = 95.195$, MSE = 0.597, $p < .0001$, showing much higher accuracy for recall in the spatial than in the temporal condition, as well as significant main effects of list length, $F(2, 1176) = 200.691$, MSE = 0.013, $p < .0001$, and of beginning-middle-end positions, $F(2, 1176) = 123.741$, MSE = 0.027, $p < .0001$, along with the interaction of list length and beginning-middle-end positions, $F(4, 2352) = 7.066$, MSE = 0.008, $p < .0001$, reflecting more pronounced effects of beginning-middle-end positions as list length increased from 8 to 12 to 16 positions. There was also a significant three-way interaction of order type, list length, and beginning-middle-end positions, $F(4, 2352) = 9.485$, MSE = 0.008, $p < .0001$, because the interaction of beginning-middle-end positions and list length was more pronounced for recall in the spatial condition than

Figure 7

Probability of Correct Recall Responses in Experiment 2 as a Function of List Length, Serial Position, Order Type, and Word Reading



Note. Error bars (shaded regions) represent between-subjects standard errors of the mean. See the online article for the color version of this figure.

in the temporal condition. In addition, there were significant component two-way interactions involving order type and beginning-middle-end positions, $F(2, 1176) = 52.440$, MSE = 0.027, $p < .0001$, and involving order type and list length, $F(2, 1176) = 7.107$, MSE = 0.013, $p = .0009$.

Based on the previous literature (e.g., Healy, 1975a), recall in the temporal condition should be higher with the aloud instructions if they lead to enhanced phonological encoding. However, recall in the spatial condition should not be higher with the aloud instructions if no phonological encoding occurs for the spatial condition. Contrary to these predictions, the temporal condition showed better recall in the silent subcondition than in the aloud subcondition, although there was little difference for the spatial condition (again, see Figure 7). The beginning-middle-end ANOVA supported this finding with a significant three-way interaction of list length, order type, and word reading, $F(2, 1176) = 3.381$, MSE = 0.013, $p = .0344$. Although this interaction was unexpected, we have a reasonable post hoc explanation for it, namely that the requirement to read aloud serves as a type of articulatory suppression (Baddeley, 2002; Baddeley et al., 1975) because vocalizing the words in the aloud subcondition might conflict with the participants' subvocalizing strategy. A conflicting subvocalizing strategy might involve subvocalizing (i.e., saying to oneself) the spatial positions of the words rather than the words themselves, in accordance with the coding of temporal–spatial patterns, as discussed later. (See Healy et al., 1991, for the use of a condition with explicit instructions to vocalize spatial positions, rather than the to-be-remembered items, which were letters in that case, and that condition performed at higher levels for spatial order recall than did a condition with instructions instead to vocalize the letters.) In any event, it should be noted that the disadvantage for reading words aloud relative to reading them silently in the temporal condition runs counter to the production effect, whereby words read aloud typically (but not always) have a recall advantage relative to words read silently (e.g., Saint-Aubin et al., 2021).

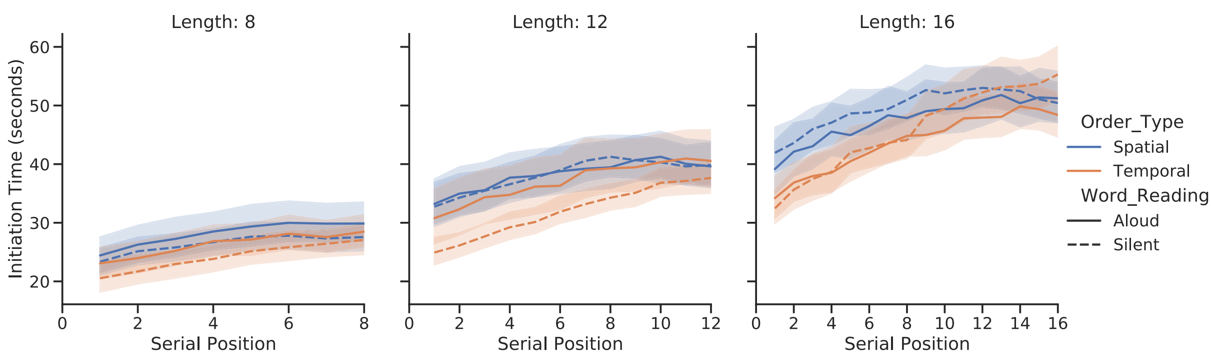
Because performance in the temporal condition was quite low, we wanted to determine whether it was at the floor. Hence, we conducted separate ANOVAs for the temporal and spatial conditions. The ANOVA for the spatial condition yielded significant main effects of list length, $F(2, 512) = 96.176$, MSE = 0.016, $p < .0001$, and of beginning-middle-end positions, $F(2, 512) = 107.622$, MSE = 0.038, $p < .0001$, along with a significant interaction of those two variables, $F(4, 1024) = 11.656$, MSE = 0.010, $p < .0001$. In contrast, the ANOVA for the temporal condition yielded significant main effects of word reading, $F(1, 332) = 4.062$, MSE = 0.361, $p = .0447$; of list length, $F(2, 664) = 100.892$, MSE = 0.010, $p < .0001$; and of beginning-middle-end positions, $F(2, 664) = 12.340$, MSE = 0.019, $p < .0001$, as well as a significant interaction of word reading and list length, $F(2, 664) = 3.797$, MSE = 0.010, $p = .0229$. Thus, the patterns for the two conditions differ to some extent, but there is no evidence of a floor effect in the temporal condition. Specifically, the effect of the beginning-middle-end positions in the temporal condition reveals a significant bow-shaped serial position function with recall in the beginning (.190) and end (.187) higher than recall in the middle (.162). That bowed function is clear even though it is less dramatic than in the spatial condition, where again recall is higher in the beginning (.434) and end (.426) than in the middle (.304).

Response Initiation Times

As in Experiment 1, we averaged RITs across correct and error responses to avoid eliminating participants who made too many errors. Again, RIT is the cumulative time it took for the participant to select a word from the alphabetical list (on the left-hand side of the screen) for movement to a box on the right-hand side of the screen. As in Experiment 1, because participants can change the order of their responses, only the RIT for the first movement of a given word is part of the analysis. As shown in Figure 8, participants are generally slower for recall in the spatial than in the temporal condition (although they are more accurate for recall in the spatial than in the temporal condition; again, see Figure 7). Thus, participants trade speed for accuracy. This finding contrasts with that of Experiment 1, in which participants were more accurate and faster in the primacy and recency sections than in the middle section of the serial position function. These findings from Experiment 1 had suggested that participants output words in order of their response confidence.

Figure 8

Response Initiation Time (in Seconds) in Experiment 2 as a Function of List Length, Serial Position, Order Type, and Word Reading



Note. Error bars (shaded regions) represent between-subjects standard errors of the mean. See the online article for the color version of this figure.

The functions for the temporal condition are generally increasing across serial positions, presumably reflecting a strategy of participants responding from the first to the last serial position. The functions for the spatial condition are increasing to some extent but also show a small inverse bow. However, the inverse bow shape is less prominent than in Experiment 1 (see, again, Figure 3).

For the ANOVA restricted to the beginning, middle, and end positions for all three list lengths, the main effects of list length, $F(2, 1176) = 101.203$, MSE = 1776.192, $p < .0001$, and of beginning-middle-end positions, $F(2, 1176) = 69.676$, MSE = 292.953, $p < .0001$, were both significant along with the interaction of those two variables, $F(4, 2352) = 31.653$, MSE = 61.334, $p < .0001$, reflecting more pronounced increasing functions for beginning-middle-end positions as list length and overall RIT increased from 8 to 12 to 16 positions.

These effects also depended to some extent on order type and word reading: Specifically, there was a significant two-way interaction of order type and beginning-middle-end positions, $F(2, 1176) = 17.421$, MSE = 292.953, $p < .0001$, reflecting a slightly bowed function for the spatial condition but an increasing function for the temporal condition, a significant three-way interaction of order type, word reading, and beginning-middle-end positions, $F(2, 1176) = 3.740$, MSE = 292.953, $p = .0240$, a significant three-way interaction of order type, list length, and beginning-middle-end positions, $F(4, 2352) = 13.316$, MSE = 61.334, $p < .0001$, a significant three-way interaction of word reading, list length, and beginning-middle-end positions, $F(4, 2352) = 2.413$, MSE = 61.334, $p = .0470$, and a significant four-way interaction of order type, word reading, list length, and beginning-middle-end positions, $F(4, 2352) = 4.996$, MSE = 61.334, $p = .0005$.

PFR

As mentioned for Experiment 1, we examined each word separately as a function of serial position to determine whether or not it was the first-word recalled.

The PFR function is also bowed, but, unlike the probability of correct recall, the bowing is more pronounced for recall in the temporal than in the spatial condition in both the primacy and the recency positions (see Figure 9). Thus, contrary to the conclusion reached based on the PFR functions for Experiment 1 and to what might be predicted from the studies by Lewandowsky et al. (2009) and Ward et al. (2010), the bowed serial position functions for response accuracy cannot be fully explained in terms of output order.

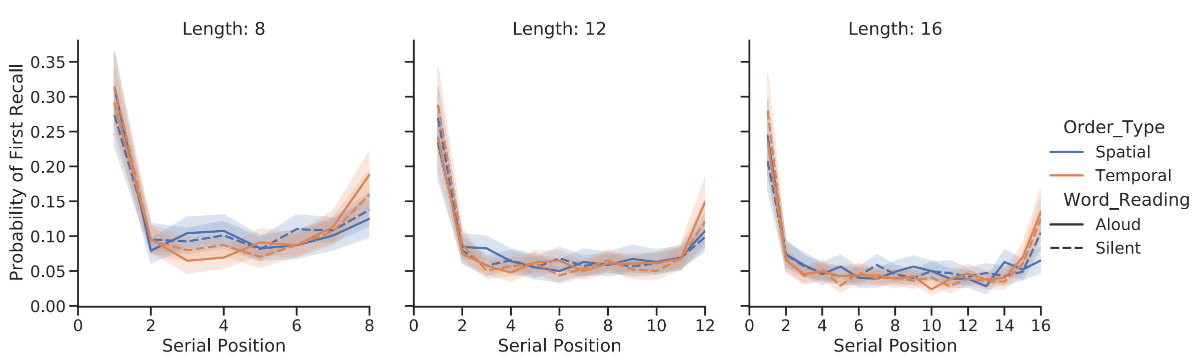
Lag-CRP

Analyses of the lag-CRPs showed effects of spatial and temporal contiguity (lag) in both the spatial and temporal conditions (see Figure 10). Also, as in Experiment 1, there is some evidence that participants preferred forward responses (+1 lag) rather than backward responses (−1 lag), especially along the temporal dimension. The contiguity effects of the alternate order on the to-be-remembered order (i.e., temporal contiguity in the spatial condition and spatial contiguity in the temporal condition) are steeper than those for the to-be-remembered order. A potential explanation for these effects might be the strategy in which participants are guided in their reconstruction responses by the alphabetically ordered list, going sequentially down the list, because the alternate order is always alphabetical and always corresponds to the list given on the left-hand side of the screen at test.

Using this strategy, participants might not concentrate on the words' identities but rather on the temporal–spatial pattern in which the words occurred during the study display (i.e., the temporal sequence of spatial locations; e.g., first word in fourth position, second word in first position, third word in eighth position, etc.). Although this strategy is available to participants in both the temporal and spatial conditions, participants might perform better in the spatial than in the temporal condition because it is easier to use the temporal–spatial patterns for recall when the alphabetical list reflects the words' temporal sequence, as it does in the spatial condition but not in the temporal condition. Thus, following the same example, in the spatial condition, participants simply have to remember the temporal sequence of spatial locations (just recall fourth, then first, then eighth, etc.). This theoretical speculation is consistent with evidence from Healy (1977, 1978, 1982) that participants base short-term spatial order recall on temporal–spatial patterns with lists of four letters. With four-letter sequences, there are only 24 possible temporal–spatial patterns, and these patterns are

Figure 9

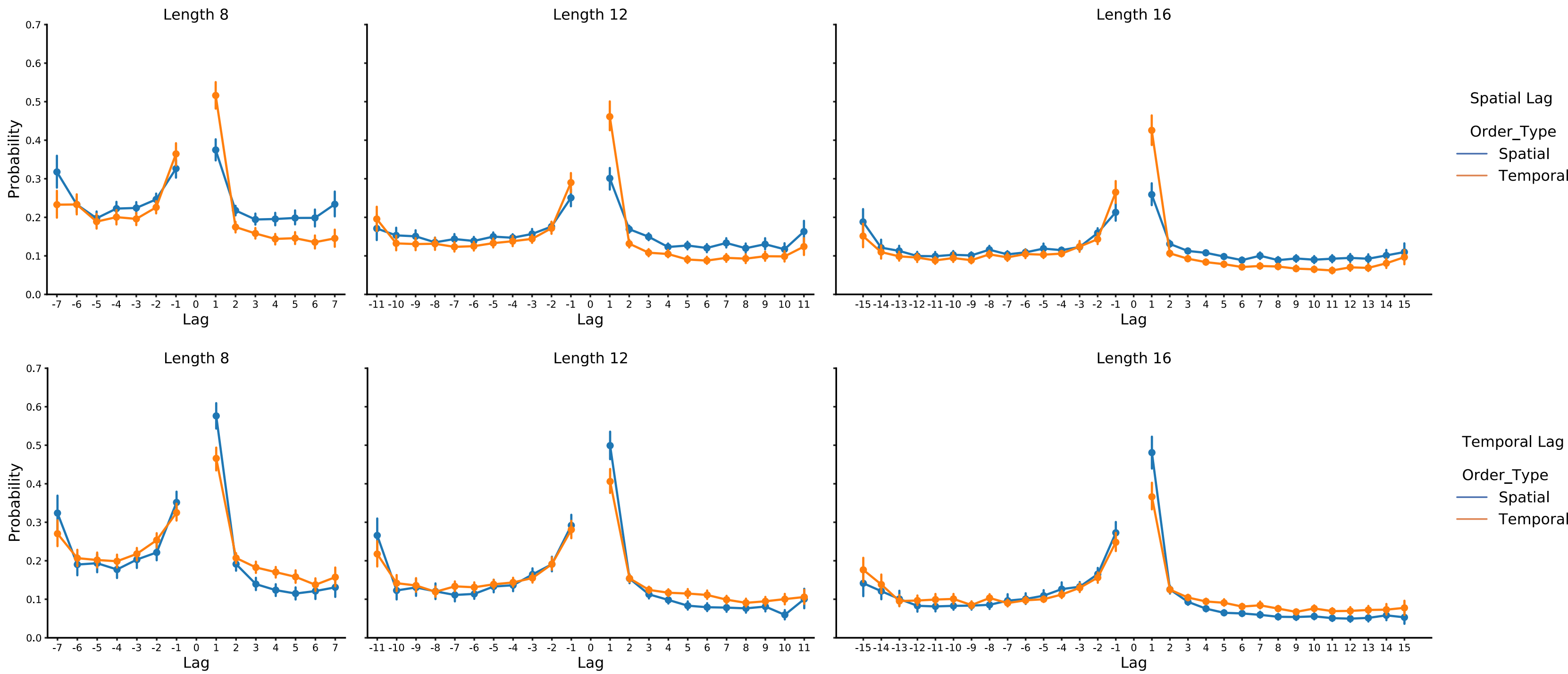
Probability of First Recall in Experiment 2 as a Function of List Length, Serial Position, Order Type, and Word Reading



Note. Error bars (shaded regions) represent between-subjects standard errors of the mean. See the online article for the color version of this figure.

Figure 10

Lag-Conditional Response Probabilities in Experiment 2 as a Function of List Length, Order Type, and Type of Lag (Spatial and Temporal)



Note. Error bars represent between-subjects standard errors of the mean. However, some error bars were too small to be visible. See the online article for the color version of this figure.

depicted for readers by Healy (1977, 1978) as 4X4 grids with the four spatial positions shown horizontally and the four temporal positions shown vertically and with an X indicating the location of each of the four letters in the grid. In fact, in an earlier one of the short-term memory studies (Healy, 1982, Experiment 2), the identity of the items was not given to the participants, who instead saw only X's occurring in different locations, and the type of temporal–spatial pattern coding used was shown to be similar to that when letter identity varied.

Recall that, in Experiment 1, we calculated the lag-CRPs as a function of alphabetical list order when there was only a single ordering of words during study. Recall also that the list of words on a given trial in Experiment 1 occurred in a central location on the screen in the central condition and that the list of words occurred from top to bottom spatially as well as from first to last temporally in the top-to-bottom condition (so the temporal and spatial positions were the same in the top-to-bottom condition). Participants showed strong contiguity effects for the temporal lag (in the top panel of Figure 5) and similar, although much weaker, contiguity effects for the alphabetical lag (in the bottom panel of Figure 5), even though the words did not occur alphabetically during study. However, the words did occur alphabetically during test when they appeared on the left-hand part of the screen. These findings are consistent with the theoretical hypothesis that in Experiment 2 the lags between successive responses are determined largely through participants' being guided in their reconstruction responses by the alphabetically ordered test list, going sequentially down the list.

Summary and Conclusions for Experiment 2

Relative to Experiment 1 and previous studies comparing temporal and spatial order recall with longer list lengths (e.g., 18–20; see Bowles & Healy, 2003; Sinclair et al., 1997), the recall rate was extremely low overall in Experiment 2. This poor performance is presumably due in part to the proactive interference (i.e., interference from previous lists) created by presenting a new list of words on every trial, which was not done in the previous comparisons of temporal and spatial order recall by Sinclair et al. (1997) and by Bowles and Healy (2003), who repeated the same list of words occurring in the same order on every trial of their experiments. However, new lists also occurred in Experiment 1, so the low performance is also undoubtedly mainly due to the huge interference created by including two types of serial positions, both spatial and temporal positions, which varied independently and hence presumably competed with each other. The same independent variation of temporal and spatial orders along with alphabetical order as the alternate order in each case occurred in the earlier studies by Sinclair et al. and Bowles and Healy, but, as just mentioned, participants in those experiments learned only a single list in a fixed order across multiple presentations and tests. Hence, performance was naturally much higher in those cases.

The two types of order (temporal and spatial) varied in complementary ways in the two experimental conditions (e.g., the alternate order in each case was always alphabetical). Nevertheless, there was a large and surprising difference between them, with recall accuracy in the temporal condition much lower than in the spatial condition. Also, the serial position functions were quite different, with the function for the spatial condition, surprisingly, symmetrically bow-shaped but the function for the temporal condition flatter.

Despite the very low performance on recall in the temporal condition, the PFR function was also bowed but more pronounced for the temporal than for the spatial condition in primacy and recency positions. As mentioned earlier, this finding is contrary to the conclusion based on the PFR functions from Experiment 1 and based on the findings from studies by Lewandowsky et al. (2009) and Ward et al. (2010). Hence, output order can no longer fully explain the bowed serial position functions for recall accuracy.

Also surprising were the observed effects of the word reading manipulation on the two order types because phonological coding was evident for temporal order recall but not for spatial order recall in previous comparisons involving short-term memory (e.g., Healy, 1975a), but reading aloud hurt performance in the current temporal condition relative to reading silently. However, word reading did not affect the current spatial condition. The present results are also puzzling because in the aloud subcondition participants were required to vocalize the words as they appeared on the screen, which corresponded to the temporal sequence of words in the temporal condition but the alternate alphabetical sequence in the spatial condition. Thus, remembering the order of the spoken words (i.e., phonological coding) could be helpful in learning the critical order in the temporal condition but not in the spatial condition (see Bowles & Healy, 2003, for a discussion of this issue).

What is responsible for these large, surprising differences in performance level, in the shapes of the recall serial position functions and PFR functions, and in the effects of word reading for the two types of order information? Based on the observed lag-CRP functions for both temporal and spatial dimensions in each of the temporal and spatial conditions, it is suggested that participants in both the temporal and spatial conditions do not always concentrate on the words' identities (e.g., their phonological, visual, and/or semantic features). Instead, they sometimes concentrate on the temporal–spatial pattern in which the words occurred during the study display (the temporal sequence of spatial locations, e.g., first word in fourth position, second word in first position, third word in eighth position, etc.), as demonstrated in short-term memory studies (Healy, 1977, 1978, 1982). Participants might perform better in the spatial than in the temporal condition because, although temporal–spatial pattern coding can be used for temporal order recall as well as spatial order recall, it is easier to use the temporal–spatial patterns for recall when the alphabetical list shown on the left-hand side of the test display reflects the words' temporal sequence. Thus, in the spatial condition, participants simply have to remember the temporal sequence of spatial locations (e.g., fourth, then first, then eighth, etc.) and can do so without any reference to the specific words that occurred at those times in those places. Participants' use of temporal–spatial pattern coding in the temporal condition could help explain why recall performance was higher in the silent than in the aloud subcondition for the temporal condition. Perhaps participants use the strategy of subvocalizing the spatial positions of words rather than the words themselves. Participants speaking the words aloud (as required in the aloud condition) would, then, interfere with that subvocalization strategy by serving as a type of articulatory suppression.

As mentioned earlier, the theoretical conclusion that participants use temporal–spatial pattern coding, especially for the spatial condition, is best supported in the present results by the more pronounced contiguity effects (lag-CRP functions) of the alternate order on the to-be-remembered order (i.e., temporal contiguity for the spatial condition and spatial contiguity for the temporal condition), which are consistent with a strategy in which participants at test recall words by sequentially going down the given alphabetical list (which corresponds to the alternate order). Steep lag-CRP functions imply that participants typically responded with words that appeared close together along the given dimension. Thus, in both conditions, they tended to respond with words closer together in the alternate order (alphabetical) than in the to-be-remembered order (temporal in the temporal condition and spatial in the spatial condition). This is a surprising and unexpected finding but can easily be explained by participants' use of the strategy of recalling words in their alphabetical order, which is the alternate order for both conditions. This strategy reflects recalling the temporal–spatial pattern of word presentations in both conditions, even in the temporal condition, where the strategy is difficult to apply.

Experiment 3

Healy (1977, 1978, 1982), under various experimental manipulations, found participants using temporal–spatial pattern coding to recall short lists of letters in short-term memory. To confirm the hypothesis that participants are also using temporal–spatial pattern coding in the present task tapping long-term memory with longer lists of words, Experiment 3 was designed to create a simpler task by repeating the exact words on every trial (along with their alternate alphabetical order), so that just their to-be-remembered order (temporal in the temporal condition and spatial in the spatial condition) was varied. The participants were told in advance that the same words would occur on each list.

By simplifying the recall task in this way, the overall recall level should be raised (see Neath, 1997), including that for the temporal condition, so the serial position functions should occur when recall in the temporal condition is higher and not close to the floor. In addition, this design is closer to that of Healy's (1975a) short-term memory study, so the proposed experiment should enable a better understanding of why there is such a different pattern of results in the present study compared to those in the earlier short-term memory studies.

Only a single list length (length 12) was used because, by definition, different numbers of words, and hence different specific words, must be shown in the different list length conditions. Consequently there were 24 experimental trials total, all at the list 12 length (as there were 24 experimental trials in Experiments 1 and 2). Because 12 is the average of 8 and 16, the same number of words were shown to participants in Experiment 3 as in Experiments 1 and 2.

Instead of just examining recall in the simplified task, we also replicated the Experiment 2 conditions so that recall under the new simplified list type (with item information fixed across trials) was compared to that in the earlier list type (with item information varied across trials), both using a single list length, within a single experiment. Finding the same pattern of results in the replicated list type as in the reported Experiment 2 version should increase confidence in the Experiment 2 results. Finding better performance accuracy in the new list type (with item information fixed) than in the earlier list type (with item information varied) would enable us to view the serial position functions when accuracy is higher (and away from the performance floor; again see Neath, 1997). In contrast, finding little difference in accuracy between the two list types would provide evidence that participants are ignoring the identities of the words and instead focusing on coding the temporal–spatial patterns of word presentations, as we have speculated here.

For participants in the earlier list type (with item information varied), as in Experiment 2, for each of the 24 trials, each participant received a different random selection of words (without replacement) from the Toronto Word Pool, so each list contained different words. In contrast, for participants in the new list type (with item information fixed), for all 24 trials, each subject received a single random selection of words (without replacement) from the Toronto Word Pool, and that same selection of words occurred for all 24 trials, so all 24 lists contained the same words.

Method

Participants

The experiment employed 109 participants tested by way of MTurk and, as in Experiments 1 and 2, paid $7.50 for their participation. Also, as in Experiments 1 and 2, the MTurk configurations restricted participants to IP addresses within the United States. For the reasons discussed in the method of Experiment 1, the experimenters excluded the data from an additional 116 participants who also completed the experiment but showed some evidence of cheating (i.e., overall accuracy over 95% or admitted to writing notes). Admitting to writing notes excluded 75 participants, and of the remaining participants, 41 were excluded for having accuracy over 95% correct. (A similar breakdown for Experiments 1 and 2 was not available to the authors at the time of writing the article.) Even though there was a large number of exclusions in Experiments 1 and 2, we used the same method in Experiment 3, to facilitate a comparison of the results of the three experiments. In any event, the fact that Experiment 3 employed 109 participants but excluded data from an additional 116 participants suggests that there was a lot of cheating in these experiments, and interpretation of the present results needs to consider this caveat, although we are confident that the data from the participants who were employed were trustworthy.

Materials and Procedure

The materials and procedure (e.g., the lists of words and their rate of presentation) were the equivalent to those in Experiment 2 except for the removal of the list length manipulation and the addition of the manipulation of list type (fixed item information, varied item information). Specifically, instructions to the participants in the fixed item information condition stated, "The same words will appear in each list." In contrast, in the varied item information condition, the instructions stated, "The words that appear will vary across each list."

Design

The experimental design was a 2 × 2 × 2 factorial, in which all three factors were varied between subjects. The first factor was order type (temporal, spatial), the second factor was word reading (aloud, silent), and the third factor was list type (fixed item information, varied item information). Participants were assigned to one of the eight between-subjects combinations of order type, word reading, and list type in a pseudorandom order, with every eight participants in a different combination, although the data from some participants were excluded due to evidence of cheating (see the Participants section above). We used the stopping rule for testing participants that we would end after we successfully tested 96 participants and tested all of those participants recruited at the same time or earlier (i.e., they were already committed to being tested because of signing up for a slot).

Results and Discussion

Accuracy

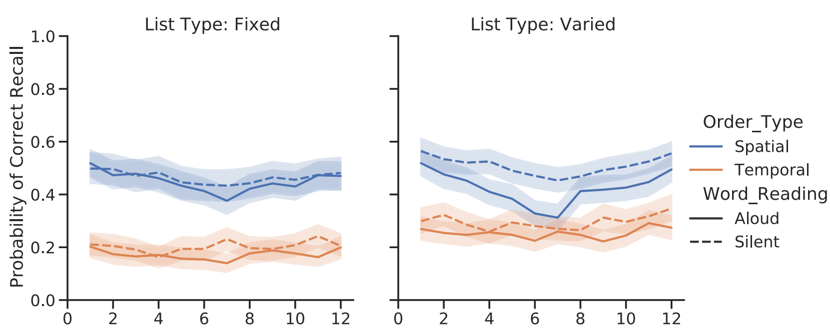
As in Experiments 1 and 2, the serial position functions show correct response proportions (see Figure 11). The spatial condition in the present experiment (.470) was more accurate than was the temporal condition (.241), even though accuracy was numerically higher in the present experiment than it was for both the spatial (.388) and the temporal (.180) conditions in Experiment 2. Contrary to predictions based on Healy's (1975a, 1977, 1978, 1982) short-term memory studies, but in agreement with the observations in Experiment 2, the spatial condition showed a perfectly bowed symmetrical serial position function. In contrast, the recall function for the temporal condition was flatter. As in Experiment 2, these findings challenge the classic simple positional coding models, which seem to assume that positional coding should be equivalent for recall in the temporal and spatial conditions because temporal and spatial positions should be equivalent. However, on the assumption that chaining associations for serial recall (i.e., linking each word to the word following it) is more likely with words that are temporally contiguous, observing flatter serial position functions for the temporal than for the spatial condition suggests that chaining is also not the crucial mechanism underlying the bowed functions. A combined ANOVA on the proportion of correct responses restricted to six serial positions, dividing them into beginning, middle, and end positions, as in Experiments 1 and 2, yielded a substantial main effect of order type, $F(1, 101) = 13.827$, MSE $= 0.304$, $p = .0003$, showing much higher accuracy for recall in the spatial than in the temporal condition, as well as a significant main effect of beginning-middle-end positions, $F(2, 202) = 26.240$, MSE $= 0.005$, $p < .0001$. Importantly, as in Experiment 2, there was also a significant interaction between order type and beginning-middle-end positions, $F(2, 202) = 8.734$, MSE $= 0.005$, $p = .0002$, reflecting a larger bow for the spatial than for the temporal condition. There were no significant effects involving the new manipulation of list type (fixed item information, varied item information) or those involving word reading (aloud, silent). In fact, there was a small and unexpected nonsignificant advantage for the varied list type (.384) relative to the fixed list type (.319), $F(1, 101) < 1$ (cf. Neath, 1997).

Unlike Experiment 2 and contrary to our initial predictions, there appeared to be somewhat better recall in the silent subcondition than in the aloud subcondition in both the temporal and spatial conditions (again, see Figure 11). However, the beginning-middle-end ANOVA did not support this finding because there was no main effect of word reading, $F(1, 101) < 1$, nor a significant two-way interaction of order type and word reading, $F(1, 101) < 1$. Hence, the significant interaction in Experiment 2 can probably be disregarded, and so it is probably unwise to accept the earlier post hoc explanation that reading aloud serves as a type of articulatory suppression because vocalizing the words in the aloud subcondition might conflict with the participants' subvocalizing strategy.

As for Experiment 2, we wanted to determine whether performance in the temporal condition was at the floor. Hence, we conducted separate ANOVAs for the temporal and spatial conditions. The ANOVA for the spatial condition yielded a significant main effect of beginning-middle-end positions, $F(2, 100) = 23.565$, MSE $= 0.007$, $p < .0001$. Likewise, the ANOVA for the temporal condition yielded a significant main effect of beginning-middle-end positions, $F(2, 102) = 4.798$, MSE $= 0.004$, $p = .0102$. Thus, as in Experiment 2, the patterns for the temporal and spatial conditions differ to some extent, but there is no evidence of a floor effect in the temporal condition. Specifically, the effect of the beginning-middle-end positions in the temporal condition shows a bow-shaped serial position function with recall in the beginning (.244) and end (.257) higher than recall in

Figure 11

Probability of Correct Recall Responses in Experiment 3 as a Function of List Type, Serial Position, Order Type, and Word Reading



Note. Error bars (shaded regions) represent between-subjects standard errors of the mean. See the online article for the color version of this figure.

the middle (.221), as in the spatial condition, where recall is also higher in the beginning (.513) and end (.494) than in the middle (.404).

Response Initiation Times

As in Experiments 1 and 2, we averaged RITs across correct and error responses, and RIT is the cumulative time to select a word from the alphabetical list (on the left) for movement to a box on the right. As in Experiments 1 and 2, only the RIT for the first movement of a given word is part of the analysis. As shown in Figure 12, as in Experiment 2, the functions for the temporal condition are generally increasing across serial positions, whereas the functions for the spatial condition are increasing to some extent but also show a small inverse bow.

For the ANOVA restricted to the beginning, middle, and end positions, the main effect of beginning-middle-end positions, $F(2, 202) = 34.207$, MSE = 107.862, $p < .0001$, was significant. This effect also depended on order type: Specifically, there was a significant two-way interaction of order type and beginning-middle-end positions, $F(2, 202) = 17.211$, MSE = 107.862, $p < .0001$, reflecting a somewhat bowed function for spatial recall but an increasing function for temporal recall. For RITs, there was also a significant main effect of the new variable of list type, reflecting shorter times for the fixed condition (38.770 s) than for the varied condition (58.689 s), $F(1, 101) = 5.882$, MSE = 5595.975, $p = .0171$, even though there were no effects of this variable on accuracy. Thus, using the same words on every trial did not improve recall accuracy but did improve recall speed.

PFR

As mentioned for Experiments 1 and 2, each word was separately examined as a function of serial position to determine whether or not it was the first word recalled.

As in Experiment 2, the PFR function is bowed, with the bowing more pronounced for temporal than for spatial recall in primacy and recency positions (see Figure 13). Thus, as we concluded on the basis of the similar findings in Experiment 2, the bowed serial position functions for response accuracy cannot be fully explained in terms of output order.

Lag-CRP

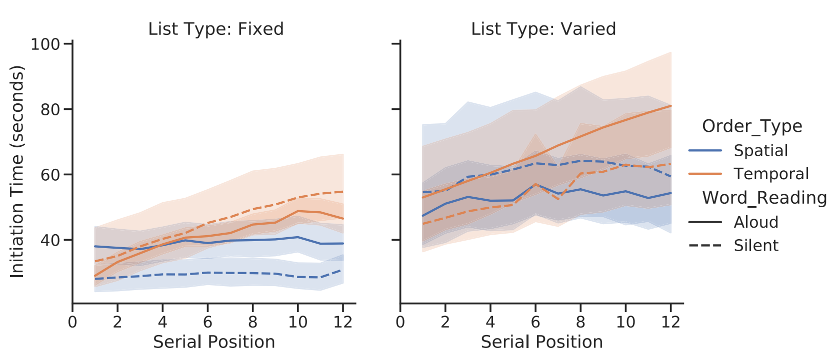
As in Experiment 2, analyses of the lag-CRPs showed spatial and temporal contiguity effects (lag) in the spatial and temporal conditions (see Figure 14). Also, as in Experiments 1 and 2, there is some evidence that participants preferred forward (+1 lag) rather than backward (−1 lag) responses, especially for temporal contiguity. Importantly, as in Experiment 2, the contiguity effects for the alternate order (alphabetical order) on the to-be-remembered order (temporal in the temporal condition and spatial in the spatial condition) are generally steeper than those for the to-be-remembered order, which suggests that participants are guided in their reconstruction responses by the alphabetically ordered list, progressing down the list.

Summary and Conclusions for Experiment 3

List type (fixed item information, varied item information) had a significant effect on RT, with the fixed item information yielding shorter responses than the varied item information, as expected because the exact same words were repeated on every trial in the fixed list type (Neath, 1997). There was, however, no significant difference between the two list types in accuracy. We were, though, successful at raising the overall numerical accuracy level, including the level for the temporal condition. In any event, most of the novel and originally unexpected results of Experiment 2 were replicated for the present experiment, including the recall advantage for the spatial relative to the temporal condition, the more bowed serial position function for the spatial than for the temporal condition, and the steeper contiguity effects for the alternate (alphabetical) order than for the to-be-remembered order (temporal for the temporal condition and spatial for the spatial condition) in both temporal and spatial order recall. Together these findings, along with the equivalent accuracy levels for the two list

Figure 12

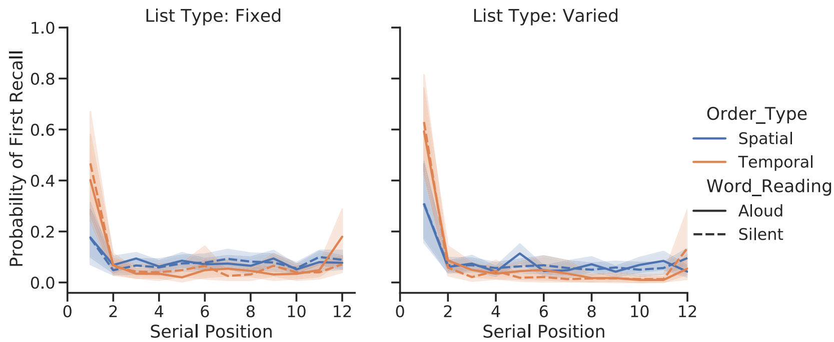
Response Initiation Time (in Seconds) in Experiment 3 as a Function of List Type, Serial Position, Order Type, and Word Reading



Note. Error bars (shaded regions) represent between-subjects standard errors of the mean. See the online article for the color version of this figure.

Figure 13

Probability of First Recall in Experiment 3 as a Function of List Type, Serial Position, Order Type, and Word Reading

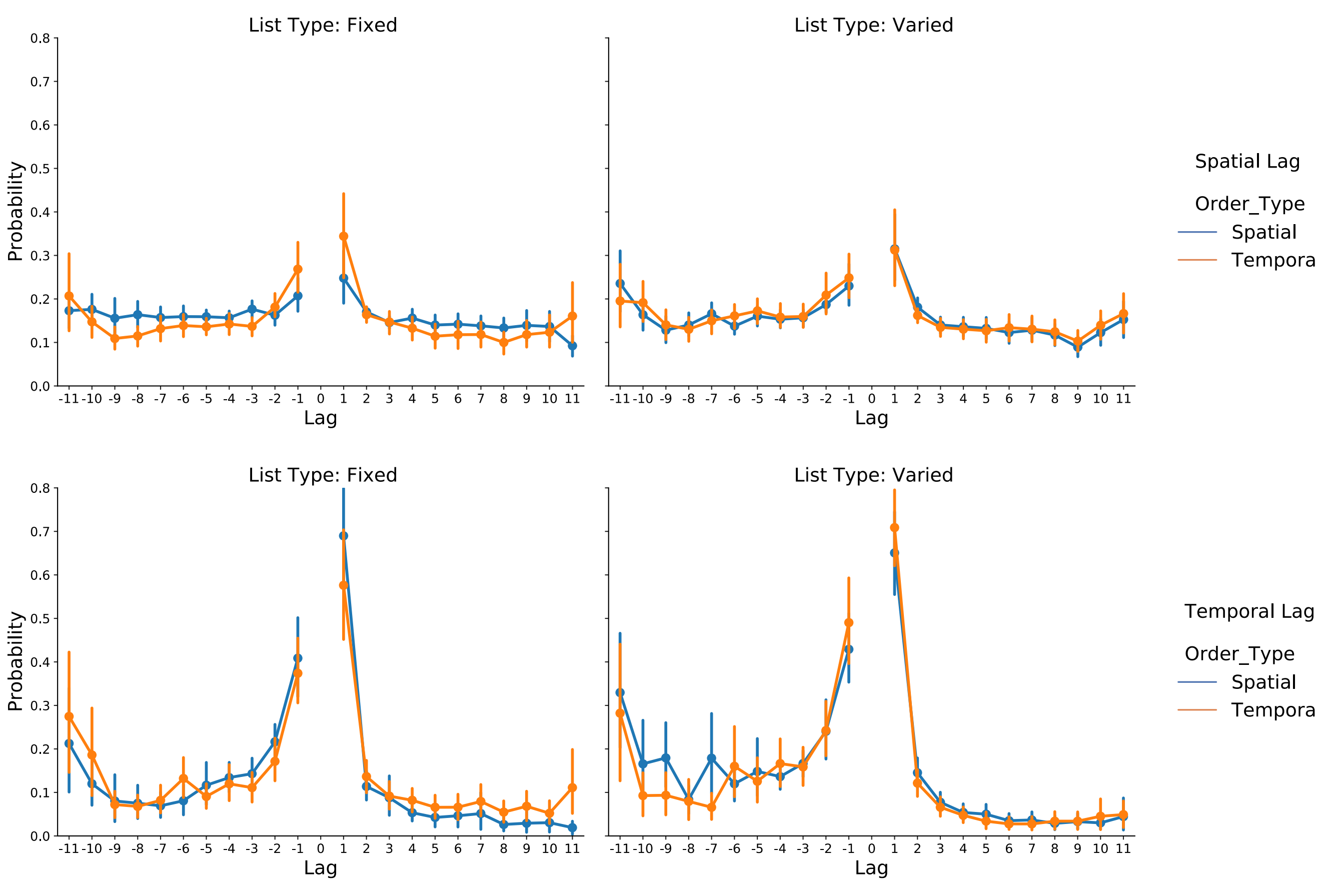


Note. Error bars (shaded regions) represent between-subjects standard errors of the mean. See the online article for the color version of this figure.

types (fixed and varied item information), are consistent with the hypothesis that instead of paying attention to the to-be-remembered position of each word, participants essentially ignore the word identity information and instead respond on the basis of the temporal–spatial pattern in which the words are presented (i.e., the temporal sequence of spatial locations), as they do when immediately reconstructing the temporal and spatial orders of short lists of four letters (e.g., Healy, 1975a).

Figure 14

Lag-Conditional Response Probabilities in Experiment 3 as a Function of List Type, Order Type, and Type of Lag (Spatial and Temporal)



Note. Error bars represent between-subjects standard errors of the mean. However, some error bars were too small to be visible. See the online article for the color version of this figure.

Transparency and Openness

The experimental materials are available at https://memory.psych.upenn.edu/Data_Archive. The data and analyses of variance are available at https://memory.psych.upenn.edu/Data_Archive. The analyses of variance were conducted using StatView (SAS Institute, 1999). The study's design and its analysis were not preregistered.

General Discussion

Experiment 1 involved only a single temporal dimension, either on its own in the central condition or coinciding with a spatial dimension in the top-to-bottom condition, whereas Experiments 2 and 3 involved two conflicting orthogonal dimensions, temporal and spatial. Interference caused by the competing spatial and temporal information in Experiments 2 and 3 is likely to be responsible for the lower overall accuracy in Experiment 2 (temporal = .180, spatial = .388) relative to Experiment 1 (central = .651, top-to-bottom = .649). This interference was naturally expected, but what was surprising is that the temporal condition was more severely harmed by this interference than was the spatial condition. Experiment 3 confirmed this pattern when performance levels were raised in both list types (fixed item information, varied item information).

Despite large differences between the performance in reconstructing temporal and spatial order information in Experiments 2 and 3, the accuracy and RT functions showed an overall similarity in the process dynamics underlying the two conditions. Specifically, the temporal and spatial conditions were similar both in the PFR functions (with the first response coming from both the primacy and recency sections of the serial position function for each type of order) and in the lag-CRP functions (with participants tending to respond with neighboring words, especially those in the forward direction, along each of the two stimulus dimensions and showing steeper contiguity functions for the alternate, alphabetical order than for the to-be-remembered order). This similarity in process dynamics is consistent with the theoretical hypothesis that participants are using the same strategy in both temporal and spatial conditions, a strategy not focused on the words' identities but instead based on the use of temporal–spatial pattern coding, although this pattern coding strategy is more successful in the spatial than in the temporal condition. The theoretical conclusion that participants use temporal–spatial pattern coding for spatial order recall was previously observed for short-term recall of short lists of four letters (Healy, 1977, 1978, 1982); the present study extends this observation to longer lists of words. This extension is important because ignoring word identities in favor of pattern coding seems much more surprising and unintuitive than ignoring meaningless letters.

With respect to the long-standing debate between simple positional coding models and temporal chaining models for serial learning (Healy & Bonk, 2008; Hurlstone, 2024; Kahana, 2012, Chapters 8 and 9), the results of Experiment 1 are consistent with both simple models, but the fact that the spatial condition shows more bowed serial position functions than does the temporal condition in Experiments 2 and 3 challenges both of these classic accounts. Specifically, as mentioned earlier, unlike the predictions based on Healy's (1975a, 1977, 1978, 1982) short-term memory studies, the spatial condition yielded a bowed symmetrical recall accuracy serial position function, and the function for the temporal condition was flatter. Simple positional coding models cannot handle this finding on the assumption that positional coding should be equivalent for temporal order recall and spatial order recall (i.e., the two types of positions should be treated similarly for the two types of orders). However, on the intuitive assumption that chaining associations for serial recall (i.e., linking each word to the word following it) is more likely with temporally contiguous words than with spatially adjacent words, finding a flatter recall accuracy serial position function (and lower accuracy) for the temporal than for the spatial condition suggests that chaining is also not the crucial mechanism underlying the bowed functions.

The shape of the serial position function for RIT (an inverse of the accuracy function) gave clear evidence in Experiment 1 for the hypothesis that participants responded with words in order of their confidence at test. That is, participants responded more quickly on a given word when they were more accurate on that word. This same pattern was not evident in Experiments 2 and 3, which instead showed that participants often responded with words in the order they occurred in the list (as reflected in the increasing functions for RIT) or (as reflected in the lag-CRP functions) in the alphabetical order of the words so that a given word preceded another word adjacent to it in the alphabetical list (and, thus, adjacent to it along the alternate order dimension, the temporal dimension for the spatial condition and the spatial dimension for the temporal condition). This result is presumably due to the experimental design's specific test conditions, in which the words on the screen's left-hand side were presented alphabetically for movement to the right-hand side of the screen. This strategy of responding alphabetically is consistent with the theoretical hypothesis that participants used temporal–spatial pattern coding for both conditions, especially for the spatial condition. There is also evidence (from lag-CRP functions) that even in Experiment 1, the participants output responses to some (although much smaller) extent in alphabetical order (i.e., responded with words close together in the alphabet sequence) when the alphabetical order did not correspond to any order occurring during word presentation.

Thus, the surprising results of Experiments 2 and 3, including unexpected differences between the temporal and spatial conditions, can be largely understood in terms of a strategic choice of using the alphabetical sequence shown in the test display as a way to guide participants' recall instead of responding with words close together along the to-be-remembered dimension or instead of relying on response confidence to guide recall, as suggested for the findings of Experiment 1 involving the complementary serial position functions for accuracy and RITs. As mentioned earlier, outputting responses alphabetically was a possible strategy, but participants used that strategy to a smaller extent in Experiment 1. The varying use of this strategy in the three experiments is presumably due in large part to the increase in the difficulty of Experiments 2 and 3 from having two conflicting stimulus dimensions.

Table 2 provides a summary collection of notes describing the findings and interpretations of them; this summary describes clear-cut but surprising results. The complete set of results can be understood mainly by relying on the theoretical hypothesis that participants use temporal–spatial pattern coding to recall both temporal and spatial order information when independently varied (and hence conflicting). Healy (1977, 1978, 1982), under various experimental manipulations, found participants using temporal–spatial pattern coding to recall short lists of letters in short-term memory.

The fact that both conditions of Experiment 3 (fixed item information, varied item information) showed the same pattern

Table 2

Summary of Results and Interpretations of Experiments 1, 2, and 3

Measure	Experiment 1: Result	Experiment 1: Interpretation	Experiments 2 and 3: Result	Experiments 2 and 3: Interpretation	Figure
Accuracy	Symmetric, bowed serial position functions for central and top-to-bottom.	Consistent with prior studies and both simple chaining and positional coding models. However, inconsistent with Murdock's (1960) distinctiveness model.	• Bowed serial position functions for spatial, flatter for temporal. • Recall quite low, with spatial more accurate than temporal. • In Experiment 2 (but not in Experiment 3), recall better in temporal under silent than aloud; little difference between aloud and silent in spatial.	• Low performance from conflicting temporal and spatial positions. • Requirement to read aloud might serve as a type of articulatory suppression. • Subvocalize spatial positions rather than words.	Experiment 1: Figure 2; Experiment 2: Figure 7; Experiment 3: Figure 11
Response initiation times	• Inverse bowed serial position functions for central and top-to-bottom. • Complementary functions; initiated more quickly when more accurate.	Output words in order of response confidence.	• Increasing serial position functions for temporal, small inverse bow for spatial.	Increasing functions reflect responding from first to last serial position. Speed–accuracy trade-off	Experiment 1: Figure 3; Experiment 2: Figure 8; Experiment 3: Figure 12
Probability of first recall	Bowed functions for both conditions.	Bowed serial position functions for accuracy explained in terms of output order.	Bowed for both conditions but more pronounced for temporal than for spatial.	Bowed serial position functions for accuracy not fully due to output order.	Experiment 1: Figure 4; Experiment 2: Figure 9; Experiment 3: Figure 13
Lag-CRP	• Strong contiguity effects, with some preference for forward relative to backward. • Similar but much weaker contiguity effects for alphabetical lag.		• Temporal contiguity effects for both conditions, with preference for forward. • Spatial contiguity effects for both conditions. • Contiguity effects steeper for alternate than for to-be-remembered order.	• Responses guided by alphabetically ordered list. • Not concentrate on words' identities but rather on temporal–spatial patterns. • Temporal–spatial pattern coding easier for spatial than for temporal order recall.	Experiment 1: Figure 5; Experiment 2: Figure 10; Experiment 3: Figure 14

Note. Lag-CRP = lag-conditional response probability.

of results implicating the use of temporal–spatial pattern coding provides some evidence that the unique design of Experiment 2 cannot be fully responsible for the present evidence of pattern coding. Furthermore, the fact that our findings concerning temporal–spatial pattern coding in Experiments 2 and 3 are consistent with the findings in earlier studies involving short-term memory (Healy, 1977, 1978, 1982) provides an additional source of evidence because of the many differences in method between those earlier studies and the present study.

To confirm further, the hypothesis that participants are also using temporal–spatial pattern coding in the present task tapping long-term memory with longer lists of words, follow-up experiments could be conducted, by increasing task difficulty to create a more challenging task, as opposed to decreasing task difficulty to create an easier task, as done in Experiment 3.

For a more challenging task, the alternate orders could be varied across trials, using a different random order instead of a fixed alphabetical order but still using the alternate order on the left-hand side of the test response screen. Such an experiment might lead to lower recall levels than those in the present study because of the variation in alternate order if the alternate order variation hinders temporal–spatial pattern coding. However, even with alternate order variation, temporal–spatial pattern coding should still be possible. Finding similar performance levels with and without the variation in the alternate orders could thus provide additional evidence for our theoretical conclusion that participants are ignoring the identities of the words and instead focusing on coding the temporal–spatial patterns of word presentations.

Another possible test of whether participants use pattern coding instead of coding the features and properties of the words would be to examine participants' knowledge of the word features and properties. Are participants better able to remember the spatial attributes of the words than their other properties (e.g., word length or initial letter)? This new question about word features (along with the questions already considered about serial position functions, output orders, confidence, and contiguity) should help to delineate further the specific processes and strategies used in reconstruction of order for lists of words and, more generally, should illuminate how individuals can learn and recall both temporal and spatial serial order information.

We have considered here classical, simple models of serial order memory (Healy & Bonk, 2008; Hurlstone, 2024; Kahana, 2012, Chapters 8 and 9). There are also more complex and complete models that have been very influential (e.g., Brown et al., 2007; Henson, 1998). However, it is beyond the scope of the present study to evaluate or compare these models in terms of their ability to account for the procedures and strategies we have identified, including the crucial temporal–spatial coding strategy. As far as we know, none of the existing models has yet addressed the use of temporal–spatial pattern coding in tasks requiring serial recall or reconstruction of order. Nevertheless, a full consideration of the influential models would certainly be another important direction for follow-up research.

References

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, *7*(2), 85–97. https://doi.org/10.1027//1016-9040.7.2.85

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

Bowles, A. R., & Healy, A. F. (2003). The effects of grouping on the learning and long-term retention of spatial and temporal information. *Journal of Memory and Language*, *48*(1), 92–102. https://doi.org/10.1016/S0749-596X(02)00520-X

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539–576. https://doi.org/10.1037/0033-295X.114.3.539

Bugelski, B. R. (1950). A remote association explanation of the relative difficulty of learning nonsense syllables in a serial list. *Journal of Experimental Psychology*, *40*(3), 336–348. https://doi.org/10.1037/h0057259

Dougherty, M. R., Halpern, D., & Kahana, M. J. (2023). Forward and backward recall dynamics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(11), 1752–1772. https://doi.org/10.1037/xlm0001254

Fischer-Baum, S., & Benjamin, A. S. (2014). Time, space, and memory for order. *Psychonomic Bulletin & Review*, *21*(5), 1263–1271. https://doi.org/10.3758/s13423-014-0604-7

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, *14*(4), 375–399. https://doi.org/10.3758/BF03203275

Guida, A., & Campitelli, G. (2019). Explaining the SPoARC and SNARC effects with knowledge structures: An expertise account. *Psychonomic Bulletin & Review*, *26*(2), 434–451. https://doi.org/10.3758/s13423-019-01582-0

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

Healy, A. F. (1975a). Coding of temporal–spatial patterns in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(5), 481–495. https://doi.org/10.1016/S0022-5371(75)80026-0

Healy, A. F. (1975b). Short-term retention of temporal and spatial order. *Bulletin of the Psychonomic Society*, *5*(1), 57–58. https://doi.org/10.3758/BF03336703

Healy, A. F. (1977). Pattern coding of spatial order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *16*(4), 419–437. https://doi.org/10.1016/S0022-5371(77)80037-6

Healy, A. F. (1978). A Markov model for the short-term retention of spatial location information. *Journal of Verbal Learning and Verbal Behavior*, *17*(3), 295–308. https://doi.org/10.1016/S0022-5371(78)90192-5

Healy, A. F. (1982). Short-term memory for order information. *Psychology of Learning and Motivation*, *16*, 191–238. https://doi.org/10.1016/S0079-7421(08)60550-2

Healy, A. F., & Bonk, W. J. (2008). Serial learning. In H. L. Roediger, III (Ed.), *Cognitive psychology of memory* (Vol. 2, pp. 53–64). Elsevier.

Healy, A. F., Cunningham, T. F., Gesi, A. T., Till, R. E., & Bourne, L. E., Jr. (1991). Comparing short-term recall of item, temporal, and spatial information in children and adults. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 127–154). Lawrence Erlbaum.

Healy, A. F., Shea, K. M., Kole, J. A., & Cunningham, T. F. (2008). Position distinctiveness, item familiarity, and presentation frequency affect reconstruction of order in immediate episodic memory. *Journal of Memory and Language*, *58*(3), 746–764. https://doi.org/10.1016/j.jml.2007.06.011

Henson, R. N. A. (1998). Short-term memory for serial order: The Start–End Model. *Cognitive Psychology*, *36*(2), 73–137. https://doi.org/10.1006/cogp.1998.0685

Hurlstone, M. J. (2024). Serial recall. In M. J. Kahana & A. D. Wagner (Eds.), *Oxford handbook of human memory* (pp. 799–830). Oxford University Press.

Kahana, M. J. (2012). *Foundations of human memory*. Oxford University Press.

Kole, J. A., Healy, A. F., & Buck-Gengler, C. J. (2005). Does number data entry rely on the phonological loop? *Memory*, *13*(3–4), 388–394. https://doi.org/10.1080/09658210344000224

Lewandowsky, S., Brown, G. D. A., & Thomas, J. L. (2009). Traveling economically through memory space: Characterizing output order in memory for serial order. *Memory & Cognition*, *37*(2), 181–193. https://doi.org/10.3758/MC.37.2.181

Mandler, G., & Anderson, R. E. (1971). Temporal and spatial cues in seriation. *Journal of Experimental Psychology*, *90*(1), 128–135. https://doi.org/10.1037/h0031342

Murdock, B. B., Jr. (1960). The distinctiveness of stimuli. *Psychological Review*, *67*(1), 16–31. https://doi.org/10.1037/h0042382

Murdock, B. B., Jr. (1969). Where or when: Modality effects as a function of temporal and spatial distribution of information. *Journal of Verbal Learning and Verbal Behavior*, *8*(3), 378–383. https://doi.org/10.1016/S0022-5371(69)80129-5

Murray, D. J. (1966). Vocalization-at-presentation and immediate recall, with varying recall methods. *Quarterly Journal of Experimental Psychology*, *18*(1), 9–18. https://doi.org/10.1080/14640746608400002

Nairne, J. S., & Dutta, A. (1992). Spatial and temporal uncertainty in long-term memory. *Journal of Memory and Language*, *31*(3), 396–407. https://doi.org/10.1016/0749-596X(92)90020-X

Neath, I. (1997). Modality, concreteness, and set-size effects in a free reconstruction of order task. *Memory & Cognition*, *25*(2), 256–263. https://doi.org/10.3758/BF03201116

Osth, A. F., & Farrell, S. (2019). Using response time distributions and race models to characterize primacy and recency effects in free recall initiation. *Psychological Review*, *126*(4), 578–609. https://doi.org/10.1037/rev0000149

Saint-Aubin, J., Yearsley, J. M., Poirier, M., Cyr, V., & Guitard, D. (2021). A model of the production effect over the short-term: The cost of relative distinctiveness. *Journal of Memory and Language*, *118*, Article 104219. https://doi.org/10.1016/j.jml.2021.104219

SAS Institute. (1999). *StatView: StatView reference*. SAS Publishing.

Shiffrin, R. M., & Cook, J. R. (1978). Short-term forgetting of item and order information. *Journal of Verbal Learning and Verbal Behavior*, *17*(2), 189–218. https://doi.org/10.1016/S0022-5371(78)90146-9

Sinclair, G. P., Healy, A. F., & Bourne, L. E., Jr. (1997). The acquisition and long-term retention of temporal, spatial, and item information. *Journal of Memory and Language*, *36*(4), 530–549. https://doi.org/10.1006/jmla.1996.2505

Slamecka, N. J. (1967). Serial learning and order information. *Journal of Experimental Psychology*, *74*(1), 62–66. https://doi.org/10.1037/h0024504

Solway, A., Murdock, B. B., & Kahana, M. J. (2012). Positional and temporal clustering in serial order memory. *Memory & Cognition*, *40*(2), 177–190. https://doi.org/10.3758/s13421-011-0142-8

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. https://doi.org/10.1037/a0020122

Received July 15, 2024
Revision received August 12, 2025
Accepted November 20, 2025 ■