Original Article

QJEP

Quarterly Journal of Experimental Psychology
2022, Vol. 75(1) 70–97

DOI: 10.1177/17470218211030825
qjep.sagepub.com
SAGE

Grouping effects in immediate reconstruction of order and the preconditions for long-term learning

Dominic Guitard[1], Jean Saint-Aubin[1] and Nelson Cowan[2]

Abstract

One commonly acknowledged role of working memory is to set up conditions for new learning. Yet, it has long been understood that there is not a perfect correspondence between conditions leading to good immediate recall from working memory and conditions leading to good delayed recall from long-term memory. Here, in six experiments, we investigated the relation between grouping effects in immediate and delayed reconstruction of order for word lists. There has been a striking absence of tests of grouping effects in long-term memory. In the first four experiments, items within groups are presented concurrently, which encourages associations between items in a group. Despite that presumably favourable situation for group learning, in Experiments 1 and 2 we found effects of grouping only in immediate order reconstruction and not in delayed reconstruction. When more processing time was allowed (Experiments 3 and 4), grouping effects in both immediate and delayed order reconstruction were obtained. Experiment 5 showed that, with items presented one at a time, but with roughly the same amount of processing time and spatial separation as the previous two experiments, grouping effects were obtained neither in immediate order reconstruction nor in delayed reconstruction. However, in Experiment 6 with a more salient manipulation of grouping, effects of grouping were obtained in immediate order reconstruction, but not in delayed reconstruction. In sum, we demonstrated for the first time that there are mechanisms of temporal grouping that assist working memory but are relatively ineffective for long-term learning, in contrast to more effective, concurrent presentation.

Keywords

Short-term memory; working memory; long-term memory; learning; grouping effects; reconstruction of order

Received: 4 December 2020; revised: 21 April 2021; accepted: 27 May 2021

Sometimes people need information only in the short term and other times they need it in the long term. There can also be a fuzzy boundary between these two situations. If you are attending to a telephone number to write it down or transfer it manually to your phone, you may only need to do so once, unless your phone battery dies. There are situations in which both short-term memory (STM) and long-term memory (LTM) will definitely be helpful. You may need to know the order in which several metro train stops occur, both to get out at the right station now and to do so on future trips. Learning the order while you have it in mind for your next move may help save you effort, allowing you to think about something else between stations on this trip and facilitating future trips. Yet, the broad literature on STM or working memory still does not clarify which conditions of immediate memory facilitate LTM. We investigated the effects of the grouping of items for the immediate and long-term reconstruction of order, as grouping is one of the best aids to STM (e.g., Ryan, 1969a, 1969b) and it has not been sufficiently investigated in the case of LTM.

Potential relation of grouping in STM to chunking in LTM

The term grouping refers to the perceived organisation of list items into groups, (as in the six-digit, two-group identification number *983-265*) and should not be confused

[1]Université de Moncton, Moncton, New Brunswick, Canada
[2]University of Missouri, Columbia, MO, USA

Corresponding authors:
Dominic Guitard, École de Psychologie, Université de Moncton, 18 avenue Antonine-Maillet, Moncton, New Brunswick E1A 3E9, Canada.
Email: edg2851@umoncton.ca

Nelson Cowan, University of Missouri, Columbia, MO, USA.
Email: cowann@missouri.edu

with the term chunking, which refers to groups that map onto known, integrated memory representations (as in the six-letter, two-chunk series *tag-box*) (Ryan, 1969a). We are interested in studying the process by which perceived groups in working memory help produce more retrievable LTM representations, possibly in a process whereby temporary groups are converted to newly learned chunks, or groups for which the parts have become strongly associated so that they are retrieved together (Miller, 1956). We achieved grouping through spatiotemporal means in four experiments to maximise the opportunity for learning, with all items of a group presented concurrently; and in the final two experiments we used a sequential presentation method of grouping to assess learning without any concurrence of items in a group, while maintaining the same spatial separation of items. Concurrent presentation of items in a group is common in daily life (e.g., when one examines a printed telephone number with a hyphen between groups) but rarely has been examined in the laboratory and conceivably might allow stronger associations to form compared to grouping with sequential presentation.

Rationale for the present set of experiments

Below, we explain the rationale behind the procedural decisions of our experiments, summarised in Table 1. We start by explaining the theoretical issues involved and move on to the task factors determined by what is needed to help distinguish between possible theories.

Theoretical issues. The theoretical motivation for this study was to examine ramifications of the assumption that the human focus of attention can consider only a few items at once (about 3 items according to Cowan, 2019; but for alternative capacity estimates, see Oberauer & Bialkova, 2009; Öztekin et al., 2010; Sutterer et al., 2019). Cowan (2019) further supposed that items in the focus of attention concurrently give rise to inter-item associations, sometimes so quickly that the newly learned groups can contribute to performance on the same trial in which they occur. From this point of view, the reason that identification numbers and telephone numbers are typically presented in groups of three to four items is not only to assist working memory of these items, but also because they facilitate long-term learning of the numbers. To the extent that this long-term learning is contributing to working memory, a delayed test should show persistence of the memory.

The expectation that working memory tasks make use of rapid new long-term memorisation is not unique to the present research; it occurs, for example, in our understanding of complex span tasks that are followed by a delayed test and show evidence of learning (e.g., McCabe, 2008; Souza & Oberauer, 2017; Unsworth & Engle, 2007). Various other theories and data also seem amenable to the

Table 1. Conditions of all experiments.

Experimental details	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5	Experiment 6
Within-list grouping structures (and how distributed)	Ungrouped, pairs, triads (between-participants)	Ungrouped, pairs, triads (within-test-cycle)	Ungrouped, pairs, triads (within-test-cycle)	Ungrouped, pairs, triads (one test type per list)	Ungrouped, temporal, spatiotemporal (within-test-cycle)	Ungrouped, temporal, spatiotemporal (within-test-cycle)
Session organisation (trials and STM-LTM test cycles)	3 test cycles, each with 8 STM trials followed by 60-s math, then 8 LTM trials on same lists.	3 test cycles, each with 9 STM trials followed by 60-s math, then 9 LTM trials on same lists.	3 test cycles, each with 6 STM trials followed by 60-s math, then 6 LTM trials on same lists.	Each list followed by just one test, after no delay, 20-s unfilled delay, or 20-s math-filled delay	3 test cycles, each with 6 STM trials followed by 60-s math, then 6 LTM trials on same lists.	3 test cycles, each with 6 STM trials followed by 60-s math, then 6 LTM trials on same lists.
Presentation method by grouping condition	Words in a group concurrent (1, 2, or 3 words at a time)	Words in a group concurrent (1, 2, or 3 words at a time)	Words in a group concurrent (1, 2, or 3 words at a time)	Words in a group concurrent (1, 2, or 3 words at a time)	Words one at a time; temporal gaps between groups in triad conditions	Words one at a time; temporal gaps between groups in triad conditions
Presentation durations	6 s per list regardless of grouping condition	6 s per list regardless of grouping condition	12 s per list regardless of grouping condition	12 s per list regardless of grouping condition	12 s per list plus 0.4 s after each word (ungrouped) or 2 s between Words 3 and 4 (triads)	1.8 s per list plus 0.3 s after each word (ungrouped) or 0.15 s after each word and 0.9 s between Words 3 and 4 (triads)

There were six words in each list. Immediate and delayed testing entailed order reconstruction. Word usage was counterbalanced across grouping conditions. Math refers to mathematical equation verification.

notion that grouping facilitates binding of items within a group to their group-node context in working memory (e.g., Broadbent & Broadbent, 1981; Cowan et al., 2002; Frankish, 1985, 1989; Hartley et al., 2016; Hitch et al., 1996; Lee & Estes, 1981; Liu & Caplan, 2020; Oberauer, 2019a; Ryan, 1969a, 1969b; Severin & Rigby, 1963; Wickelgren, 1964). Yet, few of these theorists have addressed long-term retention of the benefits of grouping, and many might subscribe to the alternative hypothesis that grouping benefits reflect temporary bindings specific to STM or working memory, not extending to long-term recall.

There are few relevant data. In the only study we can think of to examine long-term effects of short-term grouping, Sukegawa et al. (2019) used the Hebb repetition paradigm with immediate serial recall of spatial locations and manipulated temporal grouping of nine dot locations. The dots were either presented at a constant timing or temporally grouped by inserting a pause after the third and sixth serial positions. They observed the typical advantage for temporal grouping in immediate recall, but no effect of this grouping on the Hebb effect or repetition advantage.

The expectation of Cowan (2019) that concurrent existence of items in a group will lead to long-term learning of the group, facilitating LTM, is at this point only a broad principle and not a precise prediction. We do not know whether all kinds of tests are equally sensitive to a long-term effect despite the dramatic change in testing context from immediate to delayed testing. We do not know how long the presentation has to be to allow time for learning; whether items share the focus of attention only if they are presented concurrently, or also if they are presented sequentially (perhaps so, if groups function in a manner similar to semantically uniform word triads, as examined by McElree, 1998); and we do not know whether the learning is obligatory or based on an optional strategy that can be adopted.

The predominant purpose of the present work is therefore not to select one theory out of many possible ones. Instead, it is to distinguish between three possibilities that will help to strengthen and constrain any relevant theory. (1) It is possible that any manipulation sufficient to show strong grouping effects in working memory will also result in long-term learning benefits of the grouping. (2) Another alternative is that there are some kinds of grouping (e.g., perhaps with concurrent presentation of a group) that are successful for both short- and long-term retention, whereas other kinds of grouping (e.g., perhaps with sequential presentation of items and a gap between groups) may help short- but not long-term retention. This dissociation could occur because of the great shift in context between short- and long-term retrieval, perhaps making deeper encoding necessary for longer term retrieval (e.g., Craik, 2020). Either of these possibilities is compatible with versions of the theoretical framework of Cowan (2019) and will help to sharpen that framework. Against the framework, (3) it is possible that grouping does not help long-term retrieval under any circumstances, that it is based on temporary bindings that are lost quickly.

The present work focuses on reconstruction of order to determine specifically whether list organisation properties stored in the STM and in LTM are influenced by grouping. The view as stated by Cowan (2019) would suggest that if grouping assists in the immediate reconstruction of order, it should also assist in delayed reconstruction under at least some circumstances. The next section shows how the theoretical aims helped to determine manipulations within and between experiments in the current study.

Task factors

Reconstruction of order task. We chose to examine reconstruction of order in STM and LTM for lists of randomly arranged words, because it purely reflects newly encountered information, providing what we viewed as a sound basis to examine effects of grouping. It is already well known that the phonological sequence of information encountered in STM strengthens LTM for lexical items (e.g., Baddeley et al., 1998), but this learning rests to some extent on previous learning, including knowledge about the phonotactic combinations allowed in the language and knowledge of words related to the ones being learned or reinforced in STM. What is less clear is what aspects of a single STM trial strengthen newly encountered spatial and temporal associations between elements, and between the elements and serial positions or serial structure (Lee & Estes, 1981), thereby improving LTM for order. Theoretically, Cowan (2019) has argued that new associative information should be quickly learned and reconstruction of order tasks provides a way to concentrate on that type of information.

Reconstruction of order has been used successfully in various studies of immediate memory (e.g., Soemer & Saito, 2016) and delayed memory (e.g., Healy et al., 2000), but the relation between the two is still in question. Lists of items that are presented with a non-memory task can improve knowledge about list membership (Jiang & Cowan, 2020), but this is rather gross associative information. Nairne and Neath (2001) presented spoken word lists of various lengths for pleasantness rating and then, after a 5-min geometric filler task, presented a reconstruction of order task and found that correct list order memory depended on the list length and occurred about half the time with lists of 3.75 words on average. These studies provide reasons to expect that reconstruction of order can benefit from learning during an STM task.

Spatiotemporal and temporal grouping in lists. The first four of six experiments used concurrent presentation of items individually or in a group, with blank periods between

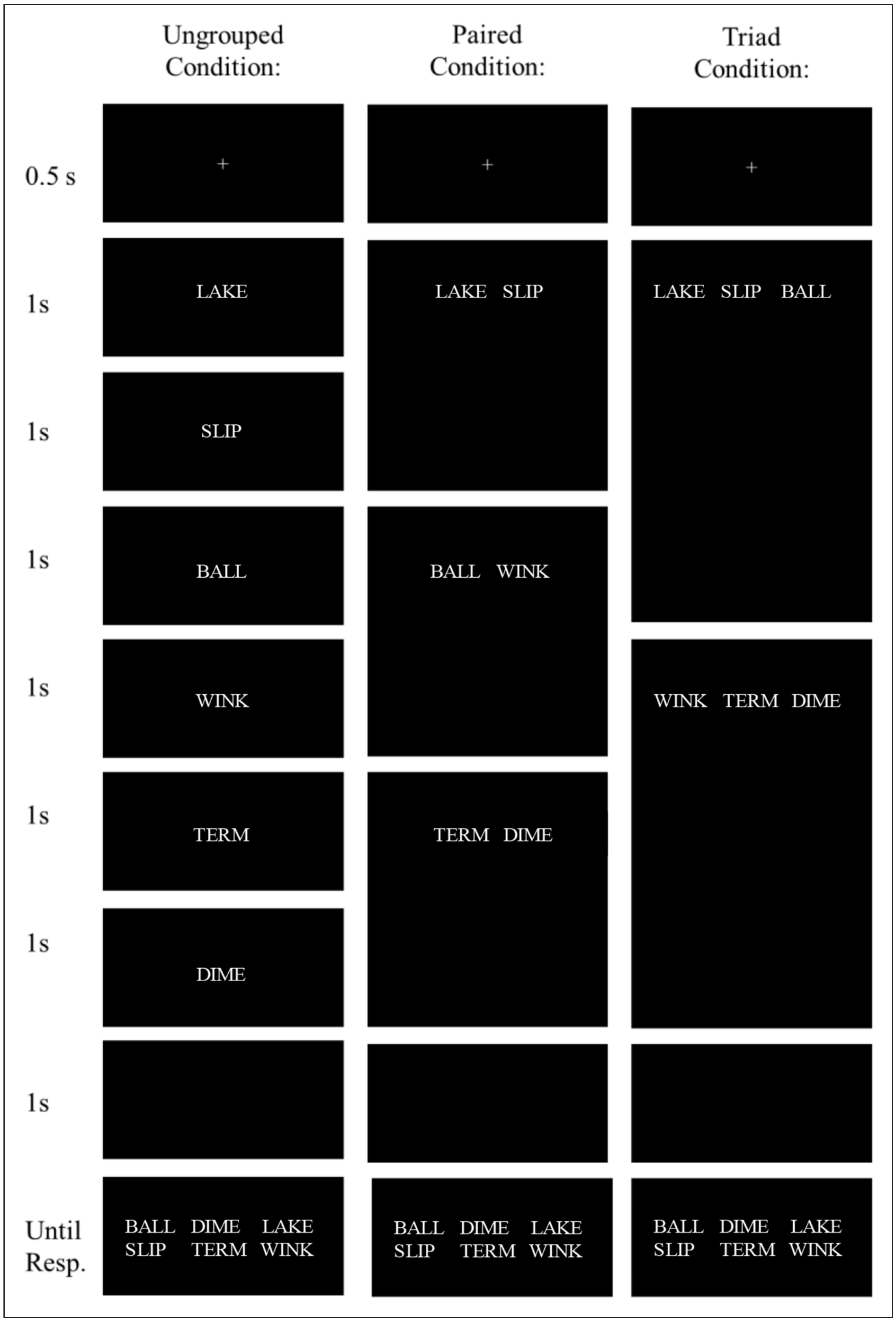

Figure 1. Illustration of a trial of the immediate reconstruction procedure for three different grouping conditions.
The vertical placement of words represents when the words first appeared and the length of the black box containing the words represents the duration for which those words remained on the screen—1 s for the ungrouped condition, 2 s for the paired condition, and 3 s for the triad condition. These three trials using the same words would not be presented to a participant; the words were randomly chosen in the experiment for each trial and are the same here across grouping conditions for the sake of comparison in the figure. For the response stimuli in the bottom row, the task was to click on a hand icon under the words, not represented here for illustration purposes, in the order of their presentation in the preceding list.

items or groups (Figure 1). In contrast, the last two experiments used one-by-one presentation of items with additional blank periods between groups. Whereas there is solid evidence that temporal grouping can assist STM (e.g., Cowan et al., 2002; Hitch et al., 1996; Ryan, 1969a, 1969b; Severin & Rigby, 1963; Wickelgren, 1964), there is less information on the effects of spatiotemporal grouping. For both of these forms of grouping, there is insufficient information to know whether they should result in improved LTM in a reconstruction of order task.

The spatiotemporal form of grouping illustrated in Figure 1 includes both strengths and weaknesses. On the plus side, it strengthens the associations and possibly the order memory for items within a group. In making items within a group concurrent, however, it may also weaken the memory the order of groups because there is a transition between one group and another; this is in contrast to temporal grouping, in which there is a transition between the last item in one group and the first item in the next group. By making the order of items in a group more distinct and certain in memory, spatiotemporal grouping may make the order of groups less distinct and uncertain (cf. Glenberg & Swanson, 1986).

Cowan (2019) and Jiang and Cowan (2020) asserted that items concurrently in the focus of attention are associated with one another but the theoretical framework leaves open the question of whether items presented in sequence in an STM task are typically in the focus of attention at once or not, as would be expected for concurrently presented items. According to some theorists, it should be possible for multiple items to be considered as part of the same group and therefore to occupy the focus of attention at once (McElree, 1998), but that has not been examined when the only grouping factor is a physical one (a temporal gap) rather than a semantic similarity between items. Thus, the theoretical concept of the focus of attention as a learning device will be sharpened by knowledge of whether LTM effects of grouping apply not only to spatiotemporal grouping but also to temporal grouping.

Between- and within-participants design. In might be expected that changing the grouping between participants could be important if participants get used to a certain grouping and impose their own organisation on a list. In Experiment 1 we tried between-participants manipulation on grouping but in doing so, we allowed variation due to individual differences in the level of performance. Given that effects on LTM were not observed in that experiment, for the remainder of the study we switched to a within-participant manipulation of grouping.

Schedule of group presentation. Factors of grouping that might affect its efficacy include (1) the absolute amount of time available to make use of any grouping cues and convert them to LTM traces (see, for example, Atkinson & Shiffrin, 1968) and (2) the relative amount of time within versus between groups, which can affect the temporal distinctiveness of groups. As Table 1 shows, we slowed down the presentation between Experiments 1–2 and Experiments 3–5, and in Experiment 6 we tried a presentation schedule that was rapid, but with quite salient time gaps between-groups of items. A slow enough presentation did prove relevant to whether grouping could enhance LTM, though the spatiotemporal versus temporal form of grouping also proved to be critical.

STM retrieval experience preceding LTM testing. Finally, our experiments differed in whether an STM test of the materials was administered before the information was retested in LTM, which was the case except in Experiment 4. Our assumption in this practice in most of the experiments was that if an LTM trace failed to form, it would be despite STM testing, not because of it; we later provide evidence backing up that assumption. When we did get an effect on LTM, it occurred in one experiment with preceding testing of the material in STM (Experiment 3) and in the experiment without that preceding testing (Experiment 4). Thus, as an ensemble, the set of experiments sets out the boundary conditions for when grouping of materials affects LTM in reconstruction of order.

Summary of experiments

Table 1 summarises the six experiments conducted to examine the factors explained above. The experiments include a between- and within-participant manipulation of spatiotemporal grouping (Experiments 1 and 2) with concurrent presentation of items in a group and a fast relatively pace of presentation; within-participant manipulation of grouping with a relatively slow pace of grouping, both with immediate tests preceding long-term testing as in the previous experiments (Experiment 3) and without immediate tests preceding long-term testing (Experiment 4); and tests of temporal-only grouping, with two different paces of presentation (Experiments 5 and 6).

Experiment 1

Experiment 1 was designed to explore the effect of spatiotemporal grouping in an immediate and delayed reconstruction of order task. Participants saw six items in one of the following conditions, which are illustrated in Figure 1: ungrouped (presented one-by-one), paired (presented two-by-two), and triad (presented three-by-three). In the immediate reconstruction condition, only one item or group of items at a time appeared on the screen as shown in Figure 1. The overall plan of our first experiment is shown in Table 1. A participant received test cycles in which eight immediate reconstruction of order trials were followed by a 60-s distracting period filled with arithmetic, and then eight delayed reconstruction of order trials for the same lists. Each participant received three such cycles with different items in each of them. Figure 1 shows the details of immediate reconstruction of order. Different participants in this experiment received six-item lists with the items ungrouped, grouped into pairs, or grouped into triads. We assume that one or both kinds of grouping should assist immediate reconstruction, with that effect more likely for triads based on previous grouping and chunking experiments in the verbal domain (Norris et al., 2020; Ryan, 1969a, 1969b; Severin & Rigby, 1963; Wickelgren, 1964).

Method

Participants

Sixty volunteers from an online data collection agency, Prolific (https://www.prolific.co/), participated and were paid £2. Inclusion criteria for this and all other experiments were as follows: (1) the participant must be a native speaker of English; (2) their nationality must be British, American, or Canadian; (3) they must have normal or corrected-to-normal vision; (4) they must have no cognitive impairment or dementia; (5) they must have no language-related disorders; (6) their age must be between 18 and 30 years; and (7) they must have an approval rating of at least 90% on prior submissions at Prolific. Inclusion criteria from 1 to 6 were self-reported and the approval rate is objectively computed by Prolific. The latter correspond to the percentage of studies for which the participant has been approved. The mean age was 24.28 ($SD=3.64$, range = 18–30); 40 self-identified as female and 20 as male.

Materials

This and all subsequent experiments were programmed with PsyToolKit (Stoet, 2010, 2017). The stimuli were monosyllabic, four-letter English words selected from the MRC Psycholinguistic Database (Coltheart, 1981). Words had a mean CELEX frequency of 38.78 and ranged from 1.01 to 199.30, according to Medler and Binder (2005). The 144 words were randomly assigned to 24 lists, each with 6 words, with the constraints that a word could only be used in one list and no words within a list could rhyme. All words and texts, unless otherwise mentioned, were presented in white, uppercase, 40 points Times New Roman font, at the centre of the computer screen on a black background. The stimuli are presented in online Supplementary Material A.

Design

Table 1 indicates the overall layout of the procedure for a participant, and Figure 1 illustrates the immediate-reconstruction phase for each of the three grouping conditions. A $3 \times 2 \times 6$ mixed design was implemented with the grouping of words within a trial (ungrouped, paired, triad) as the between-participants factor and with two repeated-measure factors: memory task (immediate reconstruction, delayed reconstruction) and serial position (1–6). There were 20 participants in each grouping condition and 48 trials (24 trials of immediate reconstruction of a six-word list, and 24 corresponding trials of delayed reconstruction of the same lists). The experiment was divided into three cycles. A cycle included (1) the immediate reconstruction task for eight trials, each involving a six-word list; (2) the distractor task, and (3) the delayed reconstruction for these same eight lists. The order of the words within a list and the lists within a cycle of eight trials were identical for each participant (see online Supplementary Material A). However, the order of the eight lists within a cycle was randomised for each participant, separately for immediate and delayed reconstruction.

Procedure

All participants were tested in one experimental session lasting approximately 25 min, or approximately 8 min for each cycle of immediate reconstruction, distractor task, and delayed reconstruction described in the "Design" section. Participants were informed that there would be three experimental cycles. Participants were further informed that the same lists used in immediate reconstruction would be represented at the end of each cycle for a delayed order reconstruction task.

Immediate reconstruction. The immediate reconstruction phase consisted of eight trials at the beginning of each of three test cycles. The participants initiated the immediate reconstruction phase after reading the instruction by pressing the "space bar" key or the phase was automatically initiated after the maximum delay of 60 s. Participants in the ungrouped condition, first saw a fixation cross "+" for 500 ms on the centre of the screen. Immediately after the fixation cross, the six to-be-remembered words were presented at a rate of one word per second (1,000 ms on, 0 ms off) at the centre of the screen. Following the presentation of the last word, there was a short retention interval of 1,000 ms. After the short retention interval, all words reappeared simultaneously in alphabetical order on two lines on the centre of the screen. Three questions marks, "???," were presented on the upper part of the screen in red as a reconstruction cue.

For participants in the paired condition, the procedure was exactly as described above, except that the six words were presented two-by-two at a rate of two words per 2 s (2,000 ms on, 0 ms off), with the pair centred on the screen with a gap between words as wide as approximately two letters. For participants in the triad condition, the procedure was exactly as described above, except that the six words were presented three-by-three at a rate of three words per 3 s (3,000 ms on, 0 ms off) at the centre of the screen with the gaps between words as wide as in the paired condition (see Figure 1).

In all three grouping conditions, participants were expected to reconstruct the order from the first to the last word presented. More exactly, participants were instructed to report items both in temporal order and in order from left to right when words were presented simultaneously. Participants were instructed to reconstruct the order by clicking on the green hand icons under the words. When participants clicked on the hand icon, the colour of the hand icon changed from green to red and simultaneously the colour of the word on top of the hand icon turn from white to grey to indicate that the word has now been selected. Participants were not allowed to backtrack in order to modify a previous response. However, the programme did let participants click more than once on an item to reuse the same item. This occurred rarely and, when it occurred, contributed to errors. Participants initiated the next trial by pressing the "space bar" key.

Distractor phase. The distractor phase immediately followed the immediate reconstruction phase. Participants initiated the distractor phase after reading the instructions by pressing the "space bar" key, or the phase was automatically initiated after the maximum delay of 60 s. Participants were instructed to go as fast as they could without sacrificing the accuracy of their responses. During this phase, participants verified math equations of the form $a \times b + c = d$, where *a, b*, and *c* were integers from 1 to 9 and *d* was equal to $a \times b + c$ or differed from that expression by ± 1. The integers for *a, b*, and *c* were drawn randomly with replacement from the integers 1 to 9, inclusive. For example, with $a = 7$, $b = 6$, and $c = 9$, a true equation would be $7 \times 6 + 9 = 51$. The results of the equations *d* had a 50% chance of being correct. Participants pressed the "z" key if they thought the equation was correct and the "m" key if they thought the equation was incorrect. To engage participants in the distractor phase they received accuracy feedback for each equation. More specifically, a counter was displayed at the bottom of the screen that indicated the number of correct responses and the number of responses they had given until that point in the cycle. The counter was reinitialised for each cycle. Participants in the distractor phase answered math equations for 60 s.

Delayed reconstruction. The delayed reconstruction phase immediately followed the distractor phase and consisted of eight trials. Like the previous phase, participants initiated the delayed reconstruction phase after reading the instruction by pressing the "space bar" key, or the phase was automatically initiated after the maximum delay of 60 s. The delayed reconstruction phase consisted of response displays identical to those in the immediate reconstruction phase, with the words from a list shown in alphabetical order to be clicked in the presented order. Participants first saw a fixation crossed "+" during 500 ms on the centre of the screen and, immediately after the fixation crossed, all the words from one of the eight lists of six to-be-remembered words that were presented in the immediate reconstruction phase reappeared simultaneously in alphabetical order on two lines on the centre of the screen. The lists were therefore represented in the same order as used for the immediate order reconstruction task. Participants were instructed to reconstruct the order in which the items had been presented in the immediate reconstruction phase they had just completed prior to the distraction phase: 500 ms after the participant clicked on the last word from a response list, they received accuracy feedback, percentage correct, for 1,500 ms followed by a blank delay of 500 ms. This was done to engage the participants in the delayed reconstruction phase. Like in the immediate reconstruction phase, the participants were not allowed to backtrack in order to modify a previous response. Participants initiated the next trial by pressing the "space bar" key. All other details of the delayed reconstruction phase were identical to the immediate reconstruction phase.

Data analysis

A strict serial reconstruction criterion was used. With this criterion, words must be reconstructed in their presentation position to be considered correct. For all analyses, the proportion of correct reconstruction was assessed as a function of serial position (1–6), grouping condition (ungrouped vs. paired vs. triad), and memory task (immediate reconstruction vs. delayed reconstruction).

In all experiments our data were analysed using both frequentist and Bayesian statistics approach using R (Version 3.6.1; R Core Team, 2019). The frequentist approach was used to generate useful descriptive statistics, but our conclusions are based on the Bayesian analyses, which allow an assessment of the reliability of null findings. For the frequentist approach, the "ez" package was used (Version 4.4-0, Lawrence, 2016) and for the BF approach the "BayesFactor" package and the default priors were used (Version 0.9.12-4.2; see Morey & Rouder, 2018; Rouder et al., 2009, 2012). For BF ANOVAs, which used Monte Carlo simulation to estimated BF, we selected an initial 100,000 iterations followed by 10,000 additional iterations which were repeated until the proportional error was below 5%. Also, for BF ANOVAs, main effects and interaction models were tested by omitting these effects one at a time from the full model and comparing them to the full model. In all BF ANOVAs, participants were included as a random effect and other factors were included as fixed effect. For Bayesian statistics, BF_{10} corresponds to the BF for the presence of an effect and BF_{01} corresponds to an absence of an effect, where $BF_{01} = 1/BF_{10}$. Post hoc comparisons were conducted with Bayesian *t*-tests.

In each experiment, participants clicked on a word for more than one serial position (a repetition error) on 3% of the total trials. We scored each serial position based on what was clicked in that position without considering the number of times the item had been selected.

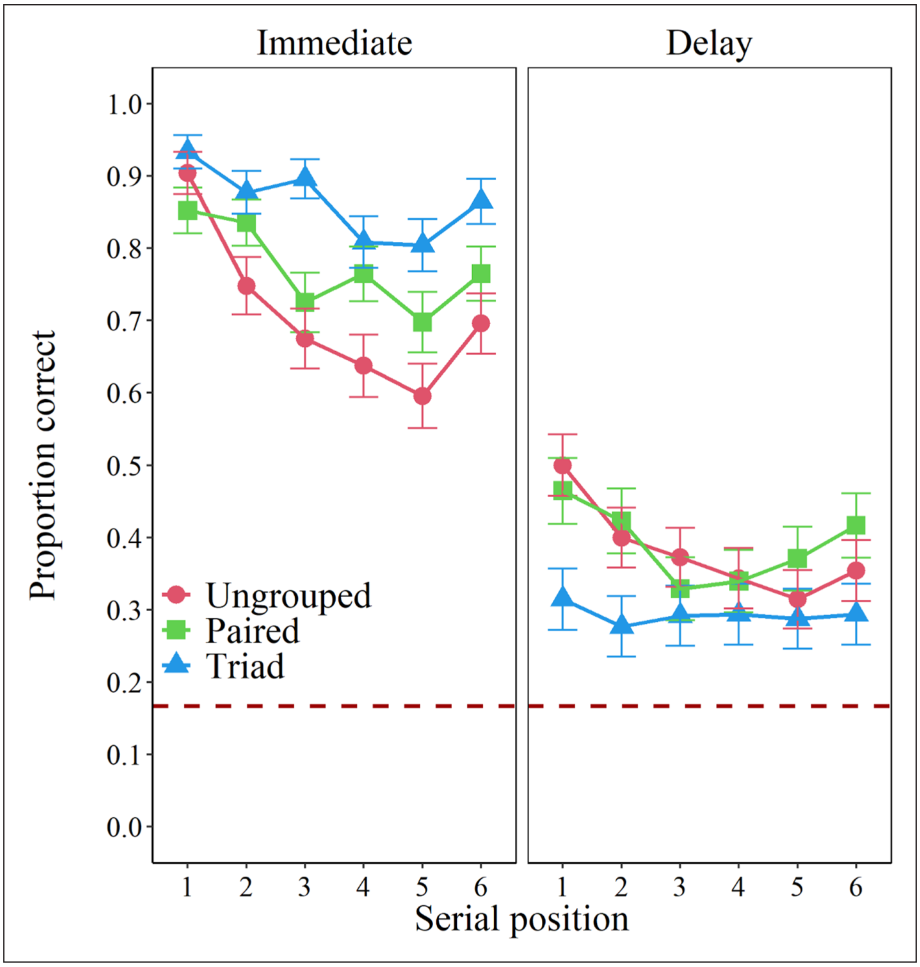

Figure 2. Proportion of correct reconstruction for Experiment 1 as a function of serial position (1–6), grouping condition (ungrouped vs. paired vs. triad), and memory task (immediate reconstruction vs. delayed reconstruction).
Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure. The dashed line represents chance performance levels.

Results

The detailed results of this experiment are shown in Figure 2. The results of all experiments are summarised for comparison in Figure 9.

Overall ANOVA. Participants were better at reconstructing the order of the words in the immediate (M=0.78, SD=0.29) than in the delayed reconstruction condition (M=0.35, SD=0.31).

Reflecting those trends, there was a main effect of memory task, F (1, 57)=488.37, η_p^2 =.90, BF_{10} >1,000, a main effect of position, F (5, 285)=47.40, η_p^2=.45, BF_{10} >1,000, but no main effect of grouping, F<1, η_p^2 =.02, BF_{01} =32.06. Most importantly, there was an interaction between memory task and grouping, F (2, 57)=14.43, η_p^2 =.34, BF_{10} >1,000, as one would expect from more task grouping in immediate reconstruction than in delayed reconstruction (Figure 2).

There was also an interaction between grouping and position, F (10, 285)=7.16, η_p^2 =.20, BF_{10} >1,000. The two-way interaction between memory task and position, F (5, 285)=5.25, η_p^2 =.08, did not contribute importantly to the full model, BF_{01} =6.04. There was no three-way interaction, F (100, 285)=1.73, η_p^2 =.06, BF_{01} >1,000.

Effects of grouping examined separately in immediate and delayed reconstruction. Given the theoretical importance of the interaction between grouping condition and memory task, we further explored the latter by running separate one-way BF ANOVAs and one-way ANOVAs for each memory task (immediate reconstruction, delayed reconstruction) with the grouping condition (ungrouped vs. paired vs. triad) as the only fixed factor. The ANOVAs revealed a main effect of grouping condition for the immediate reconstruction condition, F (2, 57)=6.71, η_p^2 =.19, BF_{10} =9.76, but no main effect of grouping condition for delayed reconstruction condition, F (2, 57)=2.30, η_p^2 =.08, BF_{01} =2.52. Post hoc testing for the immediate reconstruction condition shows that there was no reliable difference between the ungrouped and the paired condition (BF_{01}

= 1.51). However, the performance in the triad condition was superior to the ungrouped condition (BF_{10} = 69.50). The performance in the triad condition was superior to the paired condition, but only superficial evidence was found for the latter statement (BF_{01} = 1.96).

Pattern of errors in delayed reconstruction. In the absence of an advantage of grouping for delayed reconstruction, we wondered whether grouping into triads, which enhanced immediate reconstruction, had any effect on the pattern of errors in delayed reconstruction. We considered the possibility that the grouped presentation might have facilitated the clustering of items into triad groups but still, in LTM, the response could have occurred (1) without the correct order of items within a triad group, or (2) without the correct order of the two triad groups.

For the first of these questions, the six delayed-reconstruction responses on a trial were coded in terms of their intra-triad positions [1, 2, 3, 1, 2, 3], so that the intra-triad position of each response could be examined independent of whether it was placed within the correct triad. With this coding scheme, if an item should be recalled in Position 1 it is counted as correct if it is placed in Positions 1 or 4, that is, the first position within a triad; and so on. The within-triad misplacement rate for items from ungrouped lists (M = 0.10, SD = 0.06) was lower than the within-triad misplacement rate for lists grouped into triads (M = 0.14, SD = 0.04), Welch's independent samples t(33.87) = −2.15, Cohen's d = 0.68, BF_{10} = 1.83. This finding suggests that lists grouped into triads were recalled more often at the wrong within-triad position. However, only superficial evidence was found for the latter statement.

For the second question, about the order of the response triads, the six responses were coded in terms of their triad, i.e., [1, 1, 1, 2, 2, 2], to determine whether each item response was placed within the correctly located half of the response sequence. Within this coding, if an item belonged in Position 1 it was counted as correct if it was recalled in Positions 1, 2, or 3; and so on. This measure showed no difference between errors in the ungrouped (M = 0.31, SD = 0.08) and triad-grouped (M = 0.32, SD = 0.04) conditions, Welch's independent samples t(28.73) = −0.65, Cohen's d = 0.21, BF_{01} = 2.74. Putting these results together, there is reason to believe that between the short and long term representations, information about the exact serial position within a triad tended to be lost slightly more often in the triad-grouped condition, counteracting any advantage of the triad condition that existed in STM.

In the absence of an advantage of grouping for delayed reconstruction, we also wondered for the grouping condition into triads in the delayed reconstruction task after the immediate reconstruction task, if the participants might try to recall the order that they answered during the immediate reconstruction task, rather than the order of the presented items at encoding phase. To answer that question, we calculated conditional order errors. More exactly, for each delayed reconstruction trial we calculated the number of times each participant made an error by recalling the order that they answered during the immediate reconstruction task rather than the order of presented items at encoding phase divided by the total number of errors made on that trial. This measure showed that there were more errors based on the previously reconstructed order in the ungrouped (M = 0.14, SD = 0.08) compared to the triad-grouped (M = 0.06, SD = 0.05) conditions, Welch's independent samples t(30.73) = 3.87, Cohen's d = 1.22, BF_{10} = 64.09. It therefore seems unlikely that participants trying to recall the order that they answered during the immediate reconstruction task, rather than the order of the presented items at encoding phase, obscure the impact of the grouping in the delayed reconstruction task.

Discussion

Overall, memory performance was lower in the delayed than in the immediate reconstruction condition. Performance was low in the delayed reconstruction condition as expected based on the complexity of the task. However, as can be seen in Figures 2 and 9, the performance was above chance.

In the immediate reconstruction condition, participants were better in the triad condition than in the other conditions. However, this benefit was abolished in delayed reconstruction. Indeed, the mean for the triad condition in delayed reconstruction was lower than the other two conditions. Although we have no definite account of this finding, there are other cases in which temporal parameters of the stimuli seem to give rise to especially good immediate recall of some items and poor delayed recall of those same items (Kuhn et al., 2018). In our case, the triad arrangement may be helpful for encouraging organisation of the list in the short term, but it might conceivably lead to poor temporal distinctiveness of the items in a triad in the long term, as our error analysis suggests. We explored further to determine if there are boundary conditions to this result.

Experiment 2

Experiment 2 was a direct replication of Experiment 1 except that we used a within-participants manipulation and we increased the number of trials to counterbalance the grouping conditions within-participants. The change to a within-participants manipulation was made to ensure that sampling differences between conditions in Experiment 1 could not explain the pattern of results before concluding that the benefit of grouping in immediate reconstruction indicates a temporary memory that does not become a prominent part of LTM. The changes from a between-participants to a within-participants also increase the likelihood of detecting a possible effect (Thompson & Campbell, 2004).

Method

Participants. Sixty different participants volunteered from Prolific (https://www.prolific.co/). The mean age was 24.45 (SD=3.76, range=18–30); 34 self-identified as female and 26 participants self-identified as male.

Materials. The material was identical to Experiment 1 except for the following changes. Three additional lists of six words were created. The additional 18 words were monosyllabic four-letter English words selected from the MRC Psycholinguistic Database (Coltheart, 1981). The 162 words had a mean CELEX frequency of 37.90 and ranged from 1.01 to 199.30 according to Medler and Binder (2005). The stimuli are presented in the online Supplementary Material B.

Design. In Experiment 2, 3×2×6 repeated measure design was implemented with three repeated-measure factors: grouping of words within a trial (ungrouped, paired, triad), memory task (immediate reconstruction, delayed reconstruction), and serial position (1–6). There were 54 trials (27 trials of immediate reconstruction of a six-word list, and 27 corresponding trials of delayed reconstruction of the same lists). Comparable to what is shown in Table 1, this experiment was divided into three cycles. A cycle included (1) the immediate reconstruction task for nine trials, each involving a six-word list and three trials of each grouping condition (three trials ungrouped, three trials paired, three trials triad); (2) the distractor task, and (3) the delayed reconstruction for these same nine lists. The order of the words within a list and the lists within a cycle of nine trials were identical for each participant (see online Supplementary Material B). However, the order of the nine lists within a cycle was randomised for each participant, separately for immediate and delayed reconstruction. The lists that were presented for each grouping condition were counterbalanced across participants so that all the lists were encountered equally often in each grouping condition. The order of the grouping conditions within a cycle was also randomised.

Procedure. The procedure was the same as in Experiment 1 except that participants in the immediate reconstruction task and delayed reconstruction task had to complete nine trials, three trials of each grouping condition (three trials ungrouped, three trials pair, three trials triad).

Results

Overall ANOVA. As can be seen in Figures 3 and 9, participants were better at reconstructing the order of the words when they were presented in triads (M=0.57, SD=0.38) compared to the paired (M=0.55, SD=0.36) and the ungrouped condition (M=0.51, SD=0.34). Like in Experiment 1, participants were better in the immediate (M=0.75, SD=0.29) than in the delayed reconstruction condition (M=0.34, SD=0.30).

Results from the analysis of variance reveal the presence of a main effect of memory task, F (1, 59)=507.90, η_p^2 =.90, BF_{10} > 1,000, a main effect of grouping, F (2, 118)=11.92, η_p^2 =.17, BF_{10} > 1,000, and a main effect of position, F (5, 295)=72.27, η_p^2 =.55, BF_{10} > 1,000. Importantly, like Experiment 1, there was an interaction between memory task and grouping, F (2, 118)=25.02, η_p^2 =.30, BF_{10} > 1,000, indicating more grouping in immediate than in delayed reconstruction (Figure 3).

There was also an interaction between memory task and position, F (5, 295)=19.49, η_p^2 =.25, BF_{10} > 1,000, and between grouping and position, F (10, 590)=6.40, η_p^2 =.10, BF_{10} =34.36. There was also overwhelming evidence against the three-way interaction, F < 1, η_p^2 =.01, BF_{01} > 1,000.

Effects of grouping examined separately in immediate and delayed reconstruction. In this section, we further explored the two-way interaction between grouping condition and memory task by running separate one-way BF ANOVAs and one-way ANOVAs for each memory task (immediate reconstruction, delayed reconstruction) with the grouping condition (ungrouped vs. paired vs. triad) as the only fixed factor. As observed in Experiment 1, there was a main effect of grouping condition for the immediate reconstruction condition, F (2, 118)=30.45, η_p^2 =.34, BF_{10} > 1,000, but not for the delayed reconstruction condition, F (2, 118)=3.24, η_p^2 =.05, BF_{01} =7.17. In effect, for the delayed reconstruction condition there was more evidence in favour of the null model. Post hoc comparison of the delayed reconstruction conditions confirms that none of the grouping conditions differed from one another. However, post hoc testing for the immediate reconstruction condition reveals that participants were better in the triad than in the paired condition (BF_{10} > 1,000), which was in turn better than the ungrouped condition (BF_{10} =4.27).

Pattern of errors in delayed reconstruction. Unlike Experiment 1, the within-triad misplacement rate in the ungrouped condition (M=0.12, SD=0.06) and triad-grouped condition (M=0.13, SD=0.06), did not differ, paired samples t(59)=−0.71, Cohen's d=0.09, BF_{01} =5.55. Triad-placement errors (an item belonging in the first half of responses being placed in the second half or vice versa) also did not differ between the ungrouped (M=0.31, SD=0.08) and triad-grouped (M=0.31, SD=0.09) conditions, paired sample t(59)=0.62, Cohen's d=0.08, BF_{01} =5.88. The difference in within-triad placement in Experiment 1 could have been a sampling error inasmuch as ungrouped and triad-grouped presentations were seen by different participant groups. Therefore, the benefit of grouping seen in STM appears to have come from aspects of the memory representation that were inaccessible from LTM by the time of the delayed test.

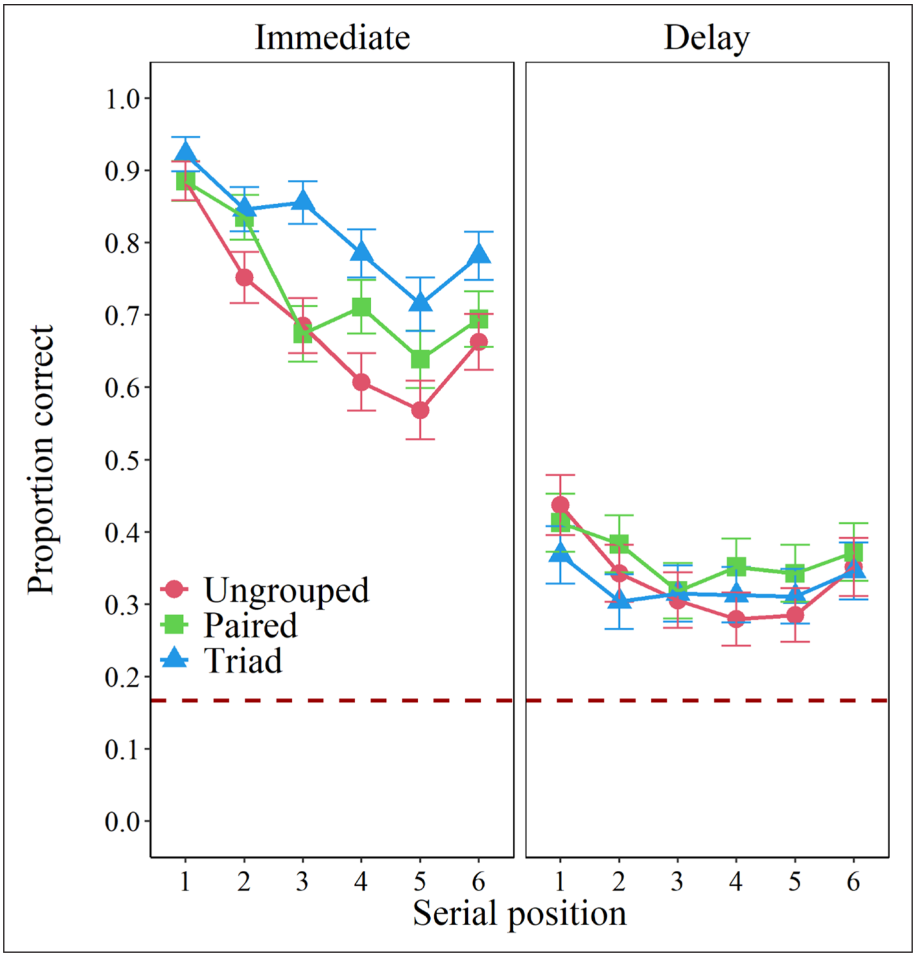

Figure 3. Proportion of correct reconstruction for Experiment 2 as a function of serial position (1–6), grouping condition (ungrouped vs. paired vs. triad), and memory task (immediate reconstruction vs. delayed reconstruction).
Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure. The dashed line represents chance performance levels.

Like Experiment 1, we explored when participants make errors in the delayed reconstruction tasks if the participants might try to recall the order that they answered during the immediate reconstruction task, rather than the order of the presented items at encoding phase. Like Experiment 1, this measure showed that there was more errors based on the immediate reconstruction order in the ungrouped (M=0.13, SD=0.09) compared to the triad-grouped (M=0.06, SD=0.06) conditions, paired samples $t(59)$=5.31, Cohen's d=0.69, BF_{10} >1,000. It is therefore unlikely again that this effect of immediate reconstruction obscures the impact of grouping in the delayed reconstruction task.

Discussion

As expected, participants were better at the immediate than the delayed reconstruction task. Participants were better in the triad condition, but only in the immediate reconstruction. As in Experiment 1, the advantage of the triad condition was abolished in the delayed reconstruction condition. This was observed despite the change to a completely within-participant design in the present experiment. This experiment thus provides further evidence that the benefit of grouping in immediate reconstruction may indicate a temporary memory that does not become a prominent part of LTM. As in Experiment 1, the triad condition improves the preservation of the words organised into triads and the orders of the triads in immediate but not in delayed reconstruction.

One possible factor that might be responsible for the absence of the benefit of grouping in delayed reconstruction is encoding time. For instance, Atkinson and Shiffrin's (1968) theory of information processing suggest that during the time that information is present in a short-term store, transfer of information from the short-term store to a long-term store will occur. The same could be true in other theories postulating separate STM and LTM stores. In theories with more integrated STM and LTM, including the unitary and embedded-processes theories, more encoding time could be used to strengthen and increase the learned associations between items in a group, forming a stronger newly learned chunk and assisting memory in a way that should be most important for LTM retrieval after intervening trials have occurred. One way to account for the absence of the benefit of grouping in delayed reconstruction in Experiments 1 and 2 might be that the encoding time was insufficient for the information to adequately transfer from the short-term store to a long-term store or to form a rich enough LTM representation. Experiment 3 was designed to explore this possibility.

Experiment 3

Experiment 3 was designed to explore the abolition of the benefit of the triad condition in delayed reconstruction. One possibility for the absence of LTM grouping effect, according to the embedded processes view (Cowan, 2019), is that items might need to be in the focus of attention concurrently for a longer period of time to give rise to inter-item associations that would transfer to LTM in a manner resulting in a retrievable trace (consistent with many other findings, e.g., Souza & Oberauer, 2017). This possibility is also in line with other well-known views in which making longer-lasting or deeper encoding might benefit from additional encoding time (e.g., Atkinson & Shiffrin, 1968; Craik, 2020). According to many views but especially a temporal distinctiveness view (e.g., Glenberg & Swanson, 1986), to observe learning in a delayed reconstruction task, one must strengthen the trace enough to overcome the interfering effects of intervening information from other trials and the resultant change in retrieval context. Therefore, we considered that the time for encoding the list in the first two experiments may have been sufficient for immediate reconstruction but insufficient to allow retrievable long-term reconstruction. If so, increasing the encoding time could extend the triad benefit to the delayed reconstruction condition. Experiment 3 was like Experiment 2 except that encoding time was doubled.

Method

Participants. Sixty different participants volunteered from Prolific (https://www.prolific.co/). The mean age was 24.22 (*SD* = 3.45, range = 18–30); 39 self-identified as female and 21 participants self-identified as male.

Materials. The material was identical to Experiment 2 except for the following changes. Three lists of six-words were removed for each cycle to accommodate the increase encoding time. The 108 words had a mean CELEX frequency of 39.14 and ranged from 1.01 to 199.30 according to Medler and Binder (2005). The stimuli are presented in the online Supplementary Material C.

Design and procedure. The design and the procedure were as described in Experiment 2, except for the following changes. There were 36 trials (18 trials of immediate reconstruction of a six-word list, and 18 corresponding trials of delayed reconstruction of the same lists). In the immediate reconstruction task participants completed six trials, each involving a six-word list and two trials of each grouping condition (two trials ungrouped, two trials paired, two trials triad).

Most critically, the encoding time of the six-word list in the immediate reconstruction task was doubled. More specifically, in the ungrouped condition the items were presented at a rate of one word per 2 s (2,000 ms on, 0 ms off), in the paired condition the six words were presented two-by-two at a rate of two words per 4 s (4,000 ms on, 0 ms off), and in the triad condition, the six words were presented three-by-three at a rate of three words per 6 s (6,000 ms on, 0 ms off).

Results

As shown in Figures 4 and 9, participants were better at reconstructing the order of the words when they were presented in triads (M = 0.69, SD = 0.36) compared to the paired (M = 0.63, SD = 0.37) and the ungrouped condition (M = 0.63, SD = 0.37). As observed in Experiments 1 and 2, participants' performance was superior in the immediate (M = 0.84, SD = 0.26) compared to the delayed (M = 0.46, SD = 0.36) reconstruction condition.

Consistent with the descriptive statistics, there was a main effect of memory task, F (1, 59) = 212.71, $\eta_p^2 = .78$, $BF_{10} > 1{,}000$, a main effect of grouping, F (2, 118) = 15.37, $\eta_p^2 = .21$, $BF_{10} > 1{,}000$, and a main effect of position, F (5, 295) = 31.74, $\eta_p^2 = .35$, $BF_{10} > 1{,}000$. Importantly, unlike Experiments 1 and 2, there was no interaction between memory task and grouping, F (2, 118) = 2.20, $\eta_p^2 = .04$, $BF_{01} = 8.90$, indicating equivalent amounts the grouping into triads in immediate and delayed reconstruction (Figure 4). Post hoc testing combining immediate and delayed conditions confirmed that there was no reliable difference between the ungrouped and the paired condition ($BF_{01} = 7.06$). However, the performance in the triad condition was superior to the ungrouped condition ($BF_{10} = 820.12$) and the paired condition ($BF_{10} = 782.63$).

The interaction between memory task and position, F (5, 295) = 4.08, $\eta_p^2 = .06$, $BF_{01} = 308.89$, and between grouping and position, F (10, 590) = 5.76, $\eta_p^2 = .09$, $BF_{01} = 2.15$, favoured the null hypothesis. Again, there was overwhelming evidence against the three-way interaction, F (10, 590) = 1.34, $\eta_p^2 = .02$, $BF_{01} > 1{,}000$.

Discussion

As found in the previous experiments, participants were better in the immediate than the delayed reconstruction condition. In this experiment, however, participants were better in the triad condition in both immediate and delayed reconstruction, not only immediate reconstruction as in

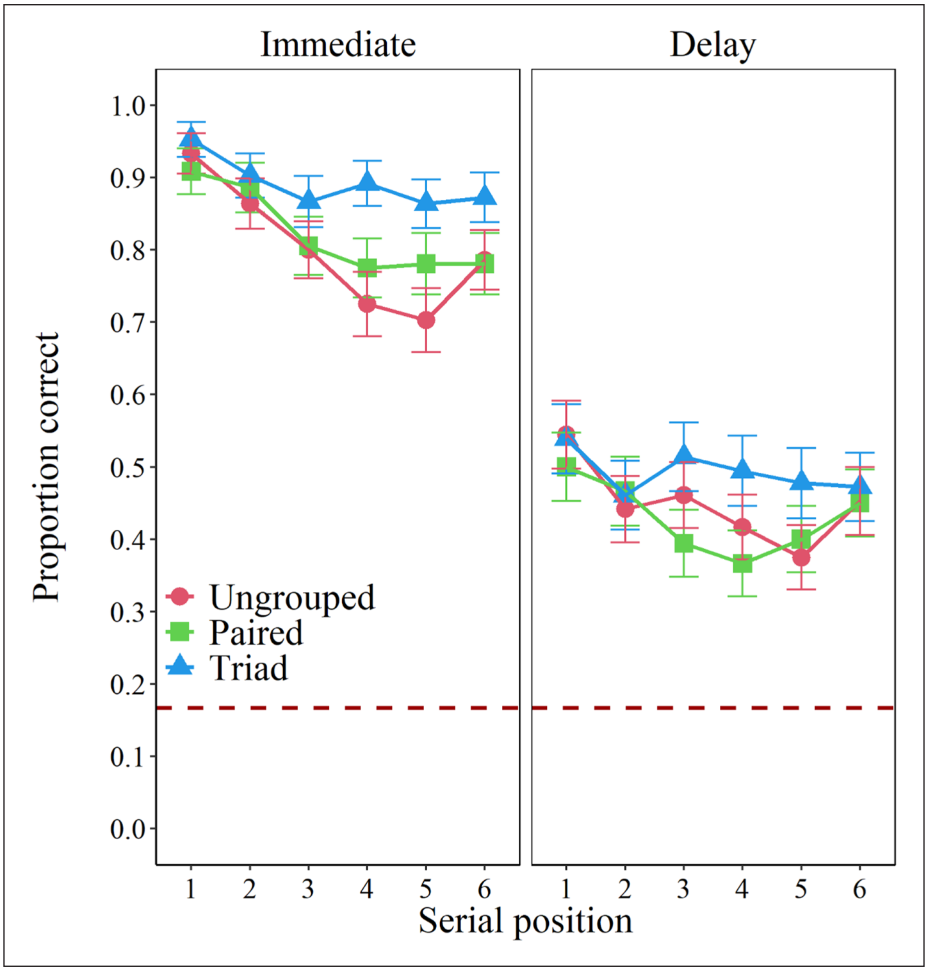

Figure 4. Proportion of correct reconstruction for Experiment 3 as a function of serial position (1–6), grouping condition (ungrouped vs. paired vs. triad), and memory task (immediate reconstruction vs. delayed reconstruction).
Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure. The dashed line represents chance performance levels.

Experiments 1 and 2. This experiment provides evidence that an STM representation of the list presumably give rise to long-term learning. It also suggests that grouping effects that help in STM might be part of what is transferred to LTM if participants have sufficient encoding time.

Experiment 4

Experiment 4 was designed to provide further information about whether the benefits of grouping items into triads observed in STM transfers to LTM. In Experiments 1–3, the information to be remembered was tested in the short term and then reinstated with a delayed test. The finding of long-term effects of grouping in Experiment 3 could be attributed in part to effects of the immediate test in strengthening learning, and participants had to overcome the change in context to delayed recall. In Experiment 4, we kept the presentation context relatively intact for the delayed test by eliminating any immediate test or intervening lists. In one condition, as shown in Table 1, distraction was used during a 20-s delay (filled retention interval) to eliminate STM representations. Participants were tested only once on each list, either immediately or after a delayed of 20 s that was either unfilled, or filled with math equations as a distraction to eliminate STM representations.

Method

Participants. Sixty different participants volunteered from Prolific (https://www.prolific.co/). The mean age was 24.88 (SD=3.50, range=18–30); 38 self-identified as female, 21 participants self-identified as male, and 1 participant did not answer the question.

Materials. The material was identical to Experiment 3 except for the following changes. A total of 144 words were used with a mean CELEX frequency of 36.22, with a range from 0.95 to 141.24 according to Medler and Binder

(2005). The stimuli are presented in the online Supplementary Material D.

Design and procedure. The design and the procedure were as described in Experiment 3, except for the following changes. There was only one cycle of 24 trials (8 trials of immediate reconstruction, 8 trials of delayed reconstruction with an unfilled delay, and 8 trials of delayed reconstruction with a filled delay, in a random order). Participants were only tested once for each list, inasmuch as there was no delayed reconstruction of all lists in the cycle as in the previous experiments. Participants had to reconstruct the six-word list on a trial either immediately after the presentation, or after 20 s in which they were instructed to either watch a fixation cross (in the unfilled delay) or solve math equations (in the filled delay). The encoding time was identical to Experiment 3. The participants only learn the condition (immediate, unfilled delay, filled delay) at the end of the trial after the presentation of the six-word list by the presentation of the words (immediate), or a fixation crossed (unfilled delay) or the arithmetic (filled delay).

Given that the main interest of this experiment was to further understand the triad benefit, the paired grouping condition was removed to save the necessary testing time for the extended retention delays on two-thirds of the trials. Participants either studied the sequence in the ungrouped condition in which the items were presented at a rate of one word per 2 s (2,000 ms on, 0 ms off) or in the triad condition, in which the six words were presented three-by-three at a rate of three words per 6 s (6,000 ms on, 0 ms off).

Results

As shown in Figures 5 and 9, participants were better at reconstructing the order of the words when they were presented in triads (M=0.77, SD=0.32) compared to the ungrouped condition (M=0.70, SD=0.33). Participants' performance was superior in the immediate (M=0.81, SD=0.28), compared to the unfilled delay (M=0.76, SD=0.32) and the filled delay condition (M=0.63, SD=0.36). The latter statement was supported by post hoc testing (all $BFs_{10} > 4.29$).

Results from the analysis of variance reveal a main effect of memory task, F (2, 118)=42.50, $\eta_p^2 = .42$, $BF_{10} > 1{,}000$, a main effect of grouping, F (1, 59)=20.64, $\eta_p^2 = .26$, $BF_{10} > 1{,}000$, and a main effect of position, F (5, 295)=31.35, $\eta_p^2 = .35$, $BF_{10} > 1{,}000$. Most importantly, like Experiment 3 but unlike the prior experiments with a faster stimulus presentation rate, there was no interaction between memory task and grouping, $F < 1$, $\eta_p^2 = .00$, $BF_{01} = 204.00$. The effect of grouping into triads was again comparable in all three tasks (Figure 5).

For the two-way interactions, the interaction between grouping and position F (5, 295)=3.40, $\eta_p^2 = .05$, did not contribute importantly to the full model, $BF_{01} = 478.11$. and between memory task and position F (10, 590)=1.73, $\eta_p^2 = .03$, $BF_{01} > 1{,}000$. The three-way interaction, F (10, 590)=2.49, $\eta_p^2 = .04$, did not contribute importantly to the full model, $BF_{01} > 1{,}000$.

Discussion

In line with the previous experiments, participants were better in the immediate than the delayed (unfilled and filled) reconstruction condition. Participants were also better in the triad condition compared to the ungrouped condition. Most importantly, as in Experiment 3, the grouping effect was persistent across STM and LTM conditions.

As Figure 5 shows, the advantage for triad grouping was manifest in the middle of the list for the immediate and the unfilled delay conditions, but it shifted to the second triad in the filled delay condition. This shift in the serial position curve was unanticipated but it seems reminiscent of previous free recall studies in which a list was followed by a filled delay before recall (e.g., Glanzer & Cunitz, 1966). A filled delay allows forgetting in the recency portion of the curve, and that seems to be the case also with ungrouped lists in our study of order reconstruction, although triad organisation appears to have protected that part of the list from effects of the filled delay.

Experiment 5

Experiment 5 was designed to explore if the benefit of the triad grouping condition in delayed reconstruction would be preserved with a sequential rather than concurrent presentation of items in a group. Spatiotemporal grouping with concurrent presentation of items within a group may encourage the linking together of items through inter-item associations. That kind of process may be less likely with items presented one at a time if, as some investigators have suggested (McElree, 1998), items that are unrelated and presented one at a time are represented only one at a time in the focus of attention. From the embedded-processes view, the focus of attention can zoom in to concentrate on a single item or zoom out as necessary to encompass several items that need to be integrated with one another or compared (e.g., Cowan et al., 2005). Therefore, information on the role of temporal grouping in working memory and LTM is needed to confine the embedded processes model.

If temporal grouping helps immediate reconstruction but not delayed reconstruction, it still will be possible to propose that the grouping effect is caused by rapid learning, and therefore according to the theory, concurrent presence of multiple items in the focus of attention. However, that concurrent presence might only be partial (e.g., Items 1 and 2 of a group; then Items 2 and 3 of a group, which could result in associative traces more susceptible to interference between immediate reconstruction and delayed reconstruction.

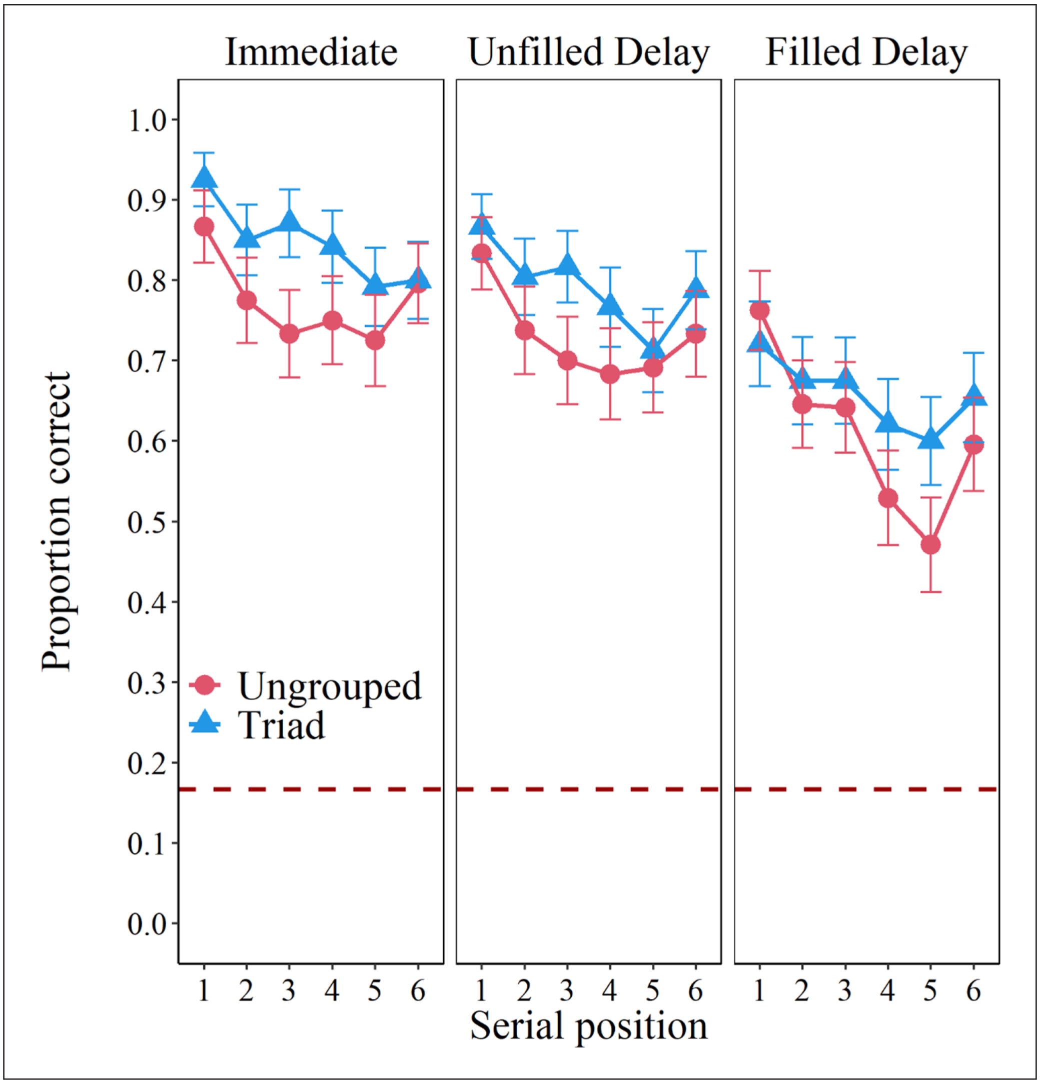

Figure 5. Proportion of correct reconstruction for Experiment 4 as a function of serial position (1–6), grouping condition (ungrouped vs. triad), and memory task (immediate reconstruction vs. unfilled delay, filled delay).
Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure. The dashed line represents chance performance levels.

Thus, in this experiment, triads were defined in terms of temporally grouped items but with items presented one at a time (see Figure 6). In the temporal grouping condition, the six items were presented centrally one-by-one for 2,000 ms each, and a 2,000 ms blank interval was inserted after the presentation of the third item (i.e., after the first triad) to create grouping. In the spatiotemporal condition, items were presented one-by-one for 2,000 ms in different spatial positions identical to the triad item positions in the previous experiments (left, centre, and right in turn for the first three items, with that placement repeated for the remaining three items). To create grouping, once again a 2,000 ms blank interval was introduced after the presentation of the first triad of items.

Both grouping conditions were compared to the ungrouped condition in which a 400 ms blank interval was introduced between the presentation of each word. This introduced blank interval resulted in presentations of equal total duration across all conditions. If previous delayed-memory benefits were due to the simultaneous presentation of the items, the benefit of the temporal or spatiotemporal condition might be observed in the immediate but not in the delayed reconstruction condition in the present experiment.

Method

Participants. Sixty different participants volunteered from Prolific (https://www.prolific.co/). The mean age was 22.53 (*SD* = 3.76, range = 18–30); 26 self-identified as female and 34 participants self-identified as male.

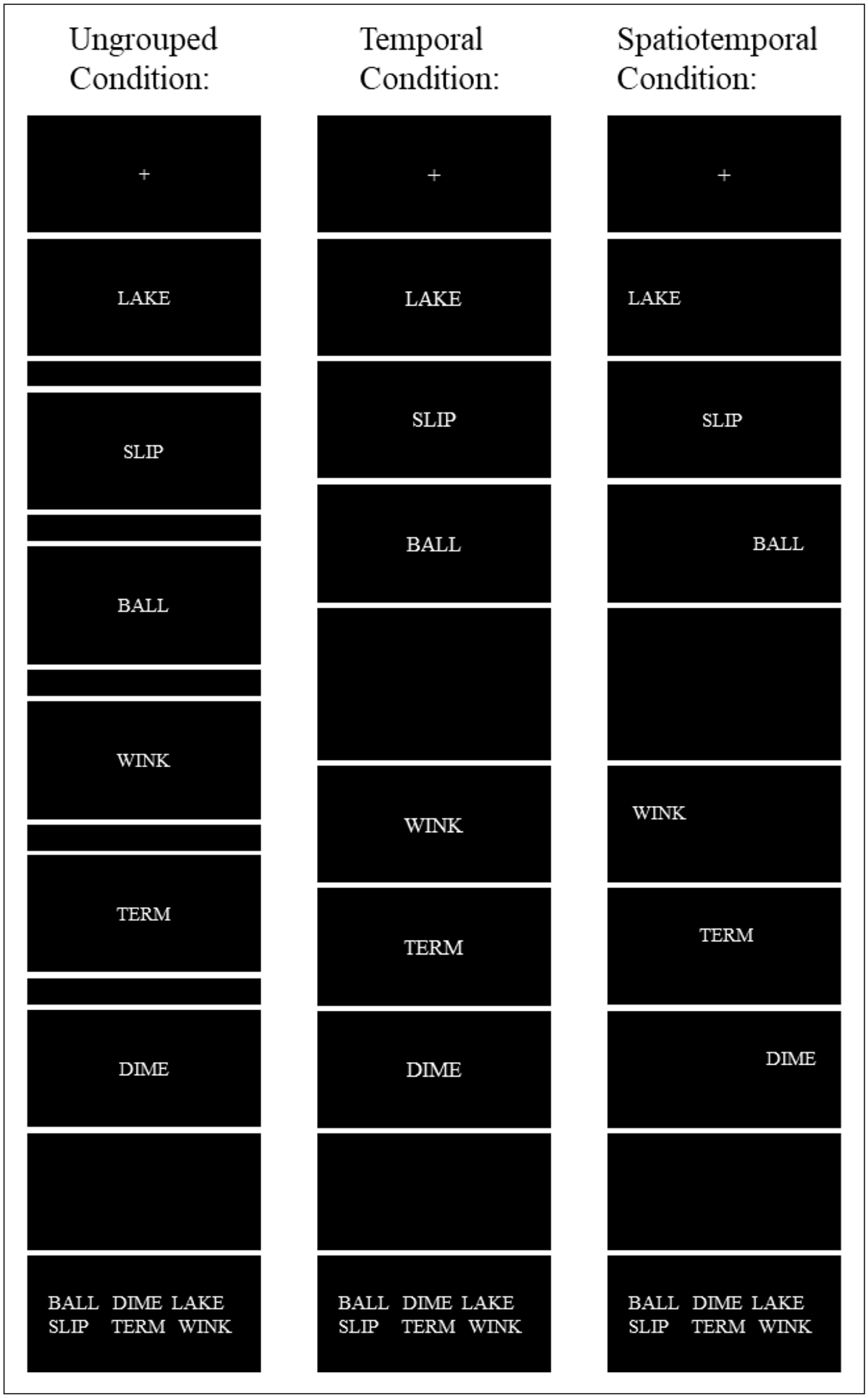

Figure 6. Illustration of a trial from top to bottom of the immediate reconstruction procedure for the three different grouping conditions in Experiments 5 and 6 (see text for details regarding presentation time and inter-stimulus intervals).
These three trials using the same words would not be presented to a participant; the words were randomly chosen in the experiment for each trial and are the same here across grouping conditions for the sake of comparison in the figure. For the response stimuli in the bottom row, the task was to click on hand icons under the words, not represented in this illustration, in the order of their presentation in the preceding list.

Materials. The material was identical to Experiment 3. The stimuli are presented in the online Supplementary Material C.

Design and procedure. The design and the procedure were as described in Experiment 3, except for the following changes (cf. Table 1). There were 36 trials (18 trials of immediate reconstruction of a six-word list, and 18 corresponding trials of delayed reconstruction of the same lists). In the immediate reconstruction task participants completed 6 trials, each involving a six-word list and 2 trials of each grouping condition (2 trials ungrouped, 2 trials temporal grouping, 2 trials spatiotemporal grouping).

Most critically, the six words were presented one-by-one in the immediate reconstruction task. More specifically, in the ungrouped condition the items were presented at a rate of one word per 2.4 s (2,000 ms on, 400 ms off). For both the temporal and spatiotemporal triad conditions, the words were presented one-by-one for 2 s (2,000 ms on, 0 ms off). After the presentation of the third word, there was a blank interval for 2 s demarking the two sets of triads. For the spatiotemporal condition, the items were presented one-by-one at the same spatial position on the screen the triads were presented in the previous experiments (Triad 1: left, middle, right; Triad 2; left, middle right).

Results

As shown in Figures 7 and 9, participants' performance was superior in the immediate (M=0.83, SD=0.27) compared to the delayed (M=0.46, SD=0.34) reconstruction condition. Overall participants' performance was similar at reconstructing the order of the words when they were presented in triads temporal condition (M=0.65, SD=0.36), in the triads spatiotemporal condition (M=0.65, SD=0.36) or in the ungrouped condition (M=0.64, SD=0.36). More importantly, unlike previous experiments, there was no beneficial effect of grouping for immediate reconstruction (ungrouped: M=0.82, SD=0.28; temporal: M=0.83, SD=0.26; spatiotemporal: M=0.83, SD=0.27), or for delayed reconstruction (ungrouped: M=0.46, SD=0.34; temporal: M=0.47, SD=0.36; spatiotemporal: M=0.46, SD=0.34).

Reflecting those trends there was a main effect of memory task, F (1, 59)=220.31, $\eta_p^2 = .79$, $BF_{10} > 1{,}000$, a main effect of position, F (5, 295)=39.92, $\eta_p^2 = .40$, $BF_{10} > 1{,}000$, but neither a main effect of grouping, $F < 1$, $\eta_p^2 = .00$, $BF_{01} = 779.08$, nor an interaction between memory task and grouping, $F < 1$, $\eta_p^2 = .01$, $BF_{01} = 234.66$. Thus, there was no grouping effect in either immediate or delayed reconstruction (Figure 7).

The interaction between memory task and position, F (5, 295)=4.26, $\eta_p^2 = .07$, did not contribute importantly to the full model, $BF_{01} = 105.77$. There was no interaction between grouping and position $F < 1$, $\eta_p^2 = .01$, $BF_{01} > 1{,}000$, and no three-way interaction, $F < 1$, $\eta_p^2 = .01$, $BF_{01} > 1{,}000$.

Pattern of errors in immediate reconstruction. In this experiment we wondered whether, in the absence of a benefit of grouping on immediate reconstruction, we could nevertheless find an effect of grouping on the error pattern. The ANOVAs of within-triad placement errors in immediate reconstruction, with the grouping condition as the repeated measure, showed no effect, F (2, 118)=2.53, $\eta_p^2 = .04$, $BF_{01} = 7.33$, with similar means (ungrouped, M=0.02, SD=0.07; temporal, M=0.01, SD=0.06; spatiotemporal, M=0.03, SD=0.08). There was also no difference, $F < 1$, $\eta_p^2 = .00$, $BF_{01} = 89.95$, in the between-half placement of items in immediate reconstruction (ungrouped, M=0.11, SD=0.18; temporal, M=0.11, SD=0.19; spatiotemporal, M=0.11, SD=0.19).

Discussion

Unlike previous experiments, participants were not better in the triad conditions (temporal, spatiotemporal) relative to the ungrouped condition in immediate reconstruction. This experiment can either suggest that the benefit of grouping required simultaneous presentation or that a more salient manipulation of temporal grouping is needed. We investigate the latter possibility in Experiment 6.

Experiment 6

Experiment 6 was designed to explore if the absence of grouping effects in Experiment 5 was due to the limited strength of the grouping manipulation, given that it yielded no grouping effect in immediate or delayed reconstruction. With visual items, it is possible that a too-slow presentation of sequential items does not result in items from a group occupying the focus of attention at the same time, preventing benefits of grouping even in immediate reconstruction. In this experiment, we modelled our timing after Experiment 1 of Hitch et al. (1996), who observed a sizable temporal grouping effect. In the temporal and spatiotemporal grouping (triad) conditions, items were presented one by one for 300 ms with a 150 ms blank interval between the presentation of each word, except that 900 ms blank interval was introduced after the presentation of the third item (see Table 1). In addition, again in the spatiotemporal condition, items were presented one-by-one in left, centre, and right spatial locations, as in Experiment 5, the same spatial positions used in all the previous experiments. Also, as in Experiment 5, both triad conditions were compared to an ungrouped condition with the same overall presentation length; the blank period following each word was 300 ms in this condition. If previous delayed-memory benefits were due to the simultaneous presentation of the items, the benefit of the temporal or spatiotemporal condition might be observed in the immediate but not in the delayed reconstruction condition.

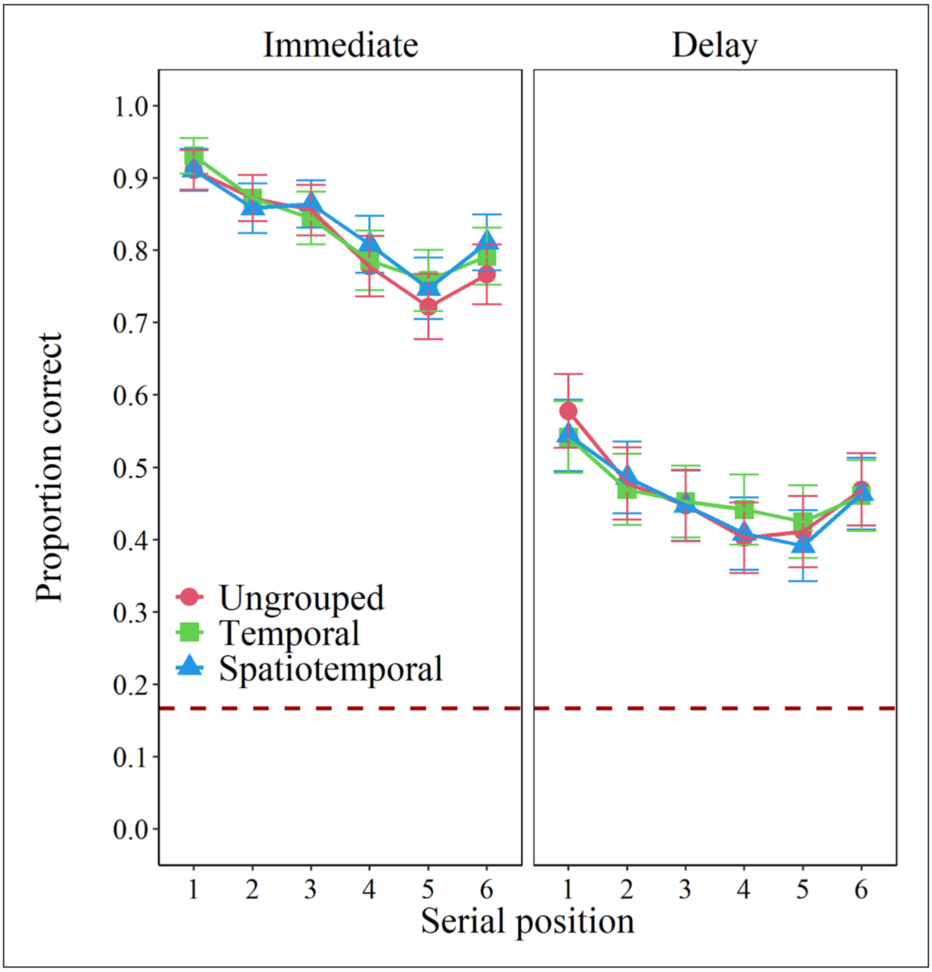

Figure 7. Proportion of correct reconstruction for Experiment 5 in which items were presented one by one as a function of serial position (1–6), grouping condition (ungrouped vs. temporal vs. spatiotemporal), and memory task (immediate reconstruction vs. delayed reconstruction).
Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure. The dashed line represents chance performance levels.

Method

Participants. Sixty different participants volunteered from Prolific (https://www.prolific.co/). The mean age was 22.49 (SD=3.76, range=18–30); 23 self-identified as female and 37 participants self-identified as male.

Materials, design, and procedure. The material, the design and the procedure were as described in Experiment 5, except for the following changes (cf. Table 1). The six words were presented one-by-one for 300 ms in the immediate reconstruction task. In the ungrouped condition, the items were presented at a rate of one word per 600 ms (300 ms on, 300 ms off). For both the temporal and spatiotemporal triad conditions, the words were presented one-by-one for 300 ms (300 ms on, 150 ms off). After the presentation of the third word, there was a blank interval for 900 ms demarking the two sets of triads.

Results

As can be seen in Figures 8 and 9, participants were better at reconstructing the order of the words when they were presented in temporal triad grouping condition (M=0.56, SD=0.37) and spatiotemporal triad grouping condition (M=0.56, SD=0.37) compared to the ungrouped condition (M=0.51, SD=0.36). As observed in previous experiments, participants' performance was superior in the immediate (M=0.77, SD=0.29) compared to the delayed (M=0.31, SD=0.28) reconstruction condition.

Reflecting the descriptive statistics, there was a main effect of memory task, F (1, 59)=480.59, $\eta_p^2 = .89$, $BF_{10} > 1{,}000$, a main effect of grouping, F (2, 118)=10.34, $\eta_p^2 = .15$, $BF_{10} > 1{,}000$, and a main effect of position, F (5, 295)=38.84, $\eta_p^2 = .40$, $BF_{10} > 1{,}000$. The interaction between memory task and grouping, F (2, 118)=4.36, $\eta_p^2 = .07$, was indeterminate, $BF_{01} = 1.02$. Below, we further investigate the latter two-way interaction, which might suggest that the benefit of grouping the items into triads was not preserved in delayed reconstruction.

There was an interaction between memory task and position, F (5, 295) = 13.00, $\eta_p^2 = .18$, $BF_{10} > 1{,}000$. However, there was no interaction between grouping and position, F (10, 590) = 1.87, $\eta_p^2 = .03$, $BF_{01} > 1{,}000$. Again, there was no three-way interaction, $F < 1$, $\eta_p^2 = .01$, $BF_{01} > 1{,}000$.

Effects of grouping examined separately in immediate and delayed reconstruction. In this section, we further explored the relation between grouping condition and memory task by running separate one-way BF ANOVAs and one-way ANOVAs for each memory task (immediate reconstruction, delayed reconstruction) with the grouping condition (ungrouped vs. temporal vs. spatiotemporal) as the only fixed factor. As observed in Experiments 1 and 2, there was a main effect of grouping condition for the immediate reconstruction condition, F (2, 118) = 13.08, $\eta_p^2 = .18$, $BF_{10} > 1{,}000$, but not for the delayed reconstruction condition, F (2, 118) = 1.48, $\eta_p^2 = .02$, $BF_{01} = 34.58$.

As in Experiments 1 and 2, post hoc comparison of the delayed reconstruction conditions confirmed that none of the grouping conditions differed importantly from one another. However, post hoc testing for the immediate reconstruction condition revealed that participants were better in the temporal compared to the ungrouped condition ($BF_{10} > 1{,}000$) and better in the spatiotemporal condition relative to the ungrouped condition ($BF_{10} = 60.81$). Finally, in immediate reconstruction the temporal and the spatiotemporal condition did not differ from another ($BF_{01} = 10.06$).

Pattern of errors in delayed reconstruction. As in Experiments 1 and 2 we wondered whether, in the absence of a benefit of grouping on delayed reconstruction, we could nevertheless find an effect of grouping on the error pattern in delayed reconstruction. ANOVAs of within-triad placement errors in delayed reconstruction like those carried out in Experiment 2, with the grouping condition as the repeated measure, showed no effect, F (2, 118) = 1.39, $\eta_p^2 = .02$, $BF_{01} = 30.05$, with similar means (ungrouped, M = 0.14, SD = 0.17; temporal, M = 0.12, SD = 0.16; spatiotemporal, M = 0.14, SD = 0.17). There was also no difference in the between-half placement of items in delayed reconstruction (ungrouped, M = 0.33, SD = 0.23; temporal, M = 0.34, SD = 0.24; spatiotemporal, M = 0.32, SD = 0.23), $F < 1$, $\eta_p^2 = .02$, $BF_{01} = 39.55$.

Like Experiments 1 and 2, we explored when participants made errors in the delayed reconstruction tasks if the participants might try to recall the order that they answered during the immediate reconstruction task, rather than the order of the presented items at encoding phase. Unlike Experiments 1 and 2, this measure showed that there was a similar number of errors based on immediate reconstruction across the conditions, F (2, 118) = 1.50, $\eta_p^2 = .02$, $BF_{01} = 20.31$, (ungrouped, M = 0.10, SD = 0.21; temporal, M = 0.09, SD = 0.20; spatiotemporal, M = 0.08, SD = 0.18). Note that in all three of these experiments, the result is not what one would expect if participants tried to recall the order in which they answered during the immediate reconstruction task even when it was erroneous.

Discussion

In Experiment 6, with a more salient manipulation of grouping based on Hitch et al. (1996), we observed grouping effects in immediate reconstruction. However, in line with the results of Experiments 1 and 2, the grouping effect was absent in the delayed reconstruction condition. This experiment provides further evidence that the benefit of grouping in immediate reconstruction indicates a temporary memory of the list that rarely becomes a prominent, part of LTM that could be observed through delayed reconstruction.

General discussion

The current study was aimed at investigating whether grouping of items would be helpful in LTM using a reconstruction of order task. As was mentioned in the introduction, we know of very little evidence addressing this topic and it was unclear what results would be obtained. The question of whether grouping assists long-term recall is an important one, as it is useful in constraining theories of STM or working memory and their contribution to LTM. As we will discuss, many major theories can predict an effect on LTM, but we believe that all need to be adjusted to account for the conditions that do or do not yield effects of grouping on LTM.

This question was investigated in the current study with a series of six experiments using a reconstruction of order task. In our first three experiments, we presented a novel sequence of six words on each trial for immediate reconstruction of order. Every block of immediate-reconstruction trials was followed by a 60-s distracting period filled with arithmetic, and then delayed reconstruction of order trials for the same lists (see Table 1 for more details). In Experiment 4, we tested reconstruction immediately or after an unfilled or filled (arithmetic) delay of 20 s. For the first three experiments, the lists were presented in three different spatiotemporal groupings: presentation of words one-by-one (six ungrouped words), two-by-two (three pairs), and three-by-three (two triads). For the fourth experiment, the words were presented ungrouped or as triads. Finally, in Experiments 5 and 6, each item was always presented alone but the time for processing each item was about the same as in the previous two experiments (in Experiment 5) or was

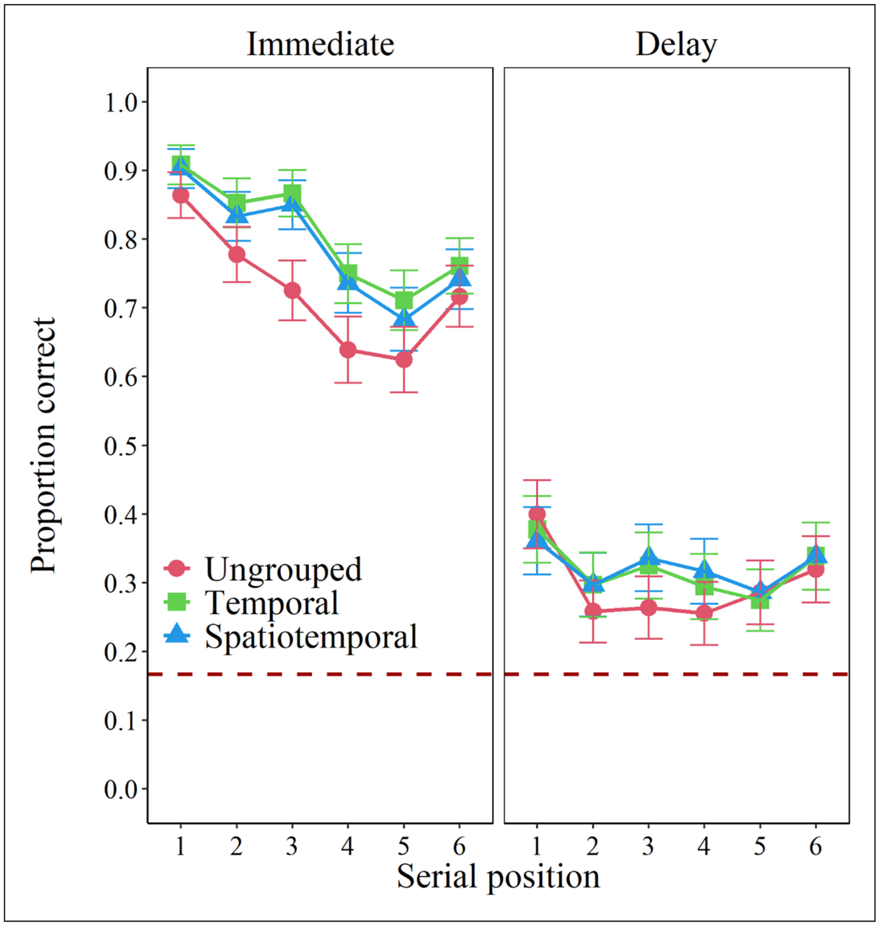

Figure 8. Proportion of correct reconstruction for Experiment 6 in which items were presented one by one as a function of serial position (1–6), grouping condition (ungrouped vs. temporal vs. spatiotemporal), and memory task (immediate reconstruction vs. delayed reconstruction).
Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure. The dashed line represents chance performance levels.

much less (in Experiment 6). The separation of stimuli into groups was carried out with all items at the centre of the screen (temporal grouping) or with items at the left, centre, and right spatial positions used in the previous experiments (spatiotemporal grouping). These conditions were compared to an ungrouped condition with the same overall presentation time.

The results are shown for proportion correct in Figure 9 and can be summarised as follows. With concurrent presentation of items in a group, in immediate reconstruction we observed a robust benefit of triad grouping, whereas there was only a small or absent benefit of the paired condition when it was included (Experiments 1–3). These findings of grouping in immediate reconstruction conform to those of Hitch et al. (1996) using temporal grouping. In delayed reconstruction, the performance was lower but well above chance in all experiments. Yet, there was little or no benefit of spatiotemporal grouping with concurrent presentation of groups in delayed reconstruction when the initial encoding time was 1 s per item (Experiments 1 and 2). It did emerge for triads when the encoding time was doubled to 2 s per item (Experiments 3 and 4).

With items presented one at a time but including either temporal or spatiotemporal cues to grouping, compared to an ungrouped condition, the pattern differed. In Experiment 5, there was no longer an effect of grouping in immediate reconstruction, suggesting that this effect in our previous experiments may have depended on the concurrent presentation of items in a triad. Neither was there an effect of grouping in delayed reconstruction in that experiment. Finally, in Experiment 6, by speeding up the rate of presentation, we found that a robust effect of grouping in immediate reconstruction reappeared despite the one-at-a-time presentation, consistent with Hitch et al. (1996); yet, no effect of grouping in delayed reconstruction emerged.

Taken together (Figure 9), the results provide important insight into what it takes for grouping to assist STM and LTM. Grouping into triads with a slow presentation rate was of great help for the immediate reconstruction of order, but only with concurrent presentation of items (Experiments

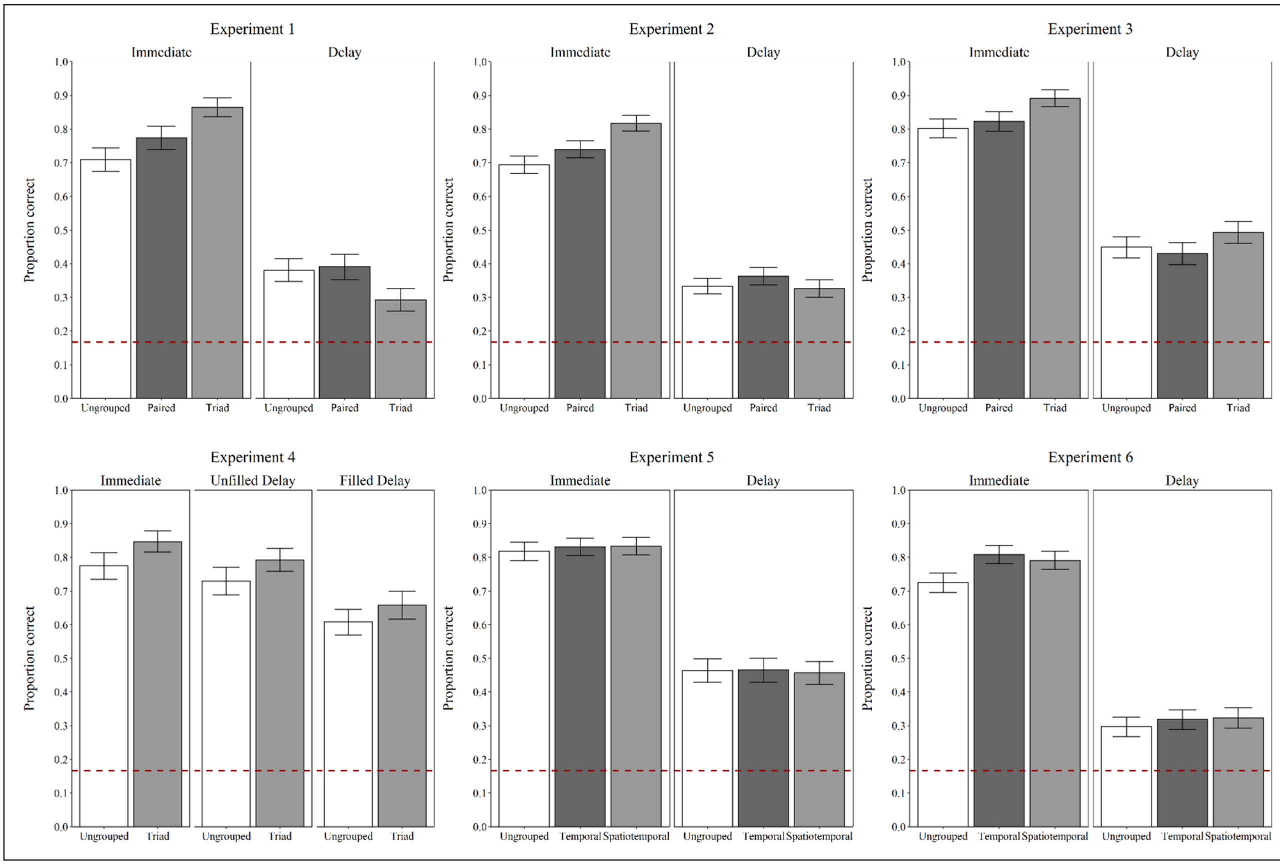

Figure 9. Proportion of correct reconstruction for each experiment (panels) as a function of grouping condition (*X* axis: ungrouped vs. paired vs. triad in Experiments 1–3; ungrouped vs. triad in Experiments 4; ungrouped vs. temporal vs. spatiotemporal in Experiments 5 and 6), and memory task (sub-panels: immediate reconstruction vs. delayed reconstruction, except for Experiment 4, which includes unfilled as well as filled delay tasks).

Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure. The dashed line represents chance performance levels.

1–4) and not with a one-by-one presentation schedule (Experiment 5). In Experiment 6, with a much faster presentation rate for one-by-one presentation, immediate reconstruction was helped by temporal grouping and the spatial placement of items did not matter. The magnitude of effects of grouping in immediate reconstruction, however, did not directly predict the magnitude of effects of grouping in delayed reconstruction. For the latter, concurrent presentation of items in a triad with a slow rate was the only situation that worked (Experiments 3 and 4). Faster rates (Experiments 1, 2, and 6) did not allow long-term learning that would aid delayed reconstruction of order.

To obtain grouping effects in delayed reconstruction, the presence of grouping effects in immediate reconstruction seems necessary; they were present whenever we obtained grouping effects in delayed reconstruction, that is, in Experiments 3 and 4. However, grouping effects in immediate reconstruction were not sufficient to obtain them in delayed reconstruction; in Experiments 1, 2, and 6, we obtained grouping effects in immediate, but not delayed, reconstruction. When grouping effects in delayed reconstruction were obtained, they were obtained on lists that had been tested previously in immediate reconstruction, in Experiment 3, and on lists that had not been tested previously in immediate reconstruction, in Experiment 4. In most of our experiments, there were intervening lists before the delayed test, but Experiment 4 was an exception to that method (Table 1). A limitation of our finding, therefore, is that we have not looked for an effect of grouping on delayed reconstruction following multiple intervening trials, but without immediate reconstruction tests preceding delayed reconstruction. Therefore, it remains possible that immediate reconstruction testing provides an important component of learning that reinforces grouping benefits helping the effect to withstand intervening, interfering information from other trials.

Relation to previous empirical findings

In this section, we briefly highlight the relation between the key results of the present study and the previous ones in the literature. One notable finding in our study, observed in Experiment 3, is the presence of a reliable effect of grouping on the accuracy of delayed reconstruction of triads, but not pairs. This pattern of results is consistent with previous demonstrations that have shown a robust benefit for items grouped into triads in the verbal domain, which has been less strongly the case for items grouped into pairs (Ryan, 1969a; Severin & Rigby, 1963; Wickelgren, 1964).

Interestingly, more long-lasting effects for triplets than for pairs have been shown also in the case of chunking, in which the grouped items correspond to known units. In particular, Norris et al. (2020) showed a redintegration process for smaller, two-word chunks like *brief case* but a data compression process only for larger, three-word chunks like *leather brief case*. In the redintegration process, knowing part of a chunk facilitates retrieval of the rest of the chunk from memory, but in a data compression process, the chunk also allows more space for other items in STM compared to a non-chunked control stimulus. This difference suggests that learning a triad may be more worthwhile than learning a pair in terms of the benefits for later retention. In Experiment 3 in our study, it is possible that a group of three items presented for a sufficient period concurrently are associated with one another to form a new long-term chunk (Cowan, 2019) that allows compression as in Norris' study, whereas pairs are not learned as well because such learning would not benefit retention nearly as much. Further examination of this possibility would have to control for factors like list length and number of groups.

The absence of a benefit of grouping items into triads for delayed reconstruction when the processing time per item is short, even when there are benefits on immediate reconstruction, is consistent with a number of previous empirical findings. For instance, these results are in line with Sukegawa et al. (2019), who found a temporal grouping benefit in immediate recall, but no benefit of grouping for the Hebb repetition effect.

An exception to this absence of an effect of grouping on long-term learning is an implicit learning procedure by Stadler (1993). Participants saw asterisks and were to press a key as quickly as possible to the location (out of 4) that corresponded with the location of the asterisk. Different fingers were used for the different stimulus/response locations. The finding was that implicit learning of a repeating set of 12 trials with a fast pace was facilitated when a 12-trial set was divided by gaps of 2 s of into consistent subsets of 4, 4, and 4 trials, compared to no division or inconsistent division.

The implicit learning procedure of Stadler (1993) and the list-recall or list-reconstruction procedures differ on many key aspects (see, for example, Guérard et al., 2011). One key difference is that, in implicit learning procedures, the correct response is practised many times. In list recall or reconstruction, errors are made. With repetition in the Hebb effect, errors are made until the correct response is approximated or learned. Perhaps this imperfect retrieval process introduces variance in the representations that subvert the long-term learning process.

All of these empirical findings combined with our results seem to support the notion that a temporary organisation of representations derived from grouping do not leave highly usable permanent impressions. However, when items are presented concurrently with sufficient processing time, we have shown that this presentation method does allow associations that assist delayed reconstruction (Experiments 3 and 4).

Relation to extant theories

In this section, we briefly discuss the important theoretical implications of our results for traditional theories of information processing, unitary memory theories, and the embedded processes view. The aim of this section is to offer possible guidance for theorists on how to account the current results and improve their current theories and/or expand their scope.

Traditional theories of information processing. According to a core assumption of Atkinson and Shiffrin's (1968) theory of information processing, during the time that information is present in a short-term store, transfer of information from the short-term store to a long-term store will occur. Based on this traditional theory, we were expecting that grouping would lead to a temporary memory of the list that would become a prominent part of LTM. This account nicely fits with the results of Experiments 3 and 4 in which the grouping of items into triads led to usable LTM representation in the delayed reconstruction. However, our initial interpretation of the theory did not predict the absence of grouping observed in Experiments 1, 2, 5, and 6. One possible way to reconcile the theory with these results is to argue that, in these experiments, the information about different items was not present concurrently in the short-term store long enough to allow associations to be formed in the long-term store. When encoding time was increased in Experiments 3 and 4 and items within a group were presented concurrently, these associations could be formed.

The expected role of sequential grouping was to parse the list into smaller subsets that each fit into the short-term store. An assumption was that the concurrent presence of items in the short-term store would facilitate a representation that included the order of items. In a six-item list divided into two sets of three items, grouping might facilitate learning of the serial order of the first three items relative to one another, and the last three items relative to one another. The serial position functions might help in the assessment of that suggestion. For immediate recall, in some experiments (e.g., the last experiment: see Figure 8) the STM functions in the triad condition show an especially large drop in accuracy between Serial Positions 3 and 4, that is, between triads. This serial position function might indicate that the first triad, encoded without a load, was preserved better than the second triad, encoded following imposition of a three-item load. The type of encoding that was so helpful for the first triad, however, did not transfer to LTM and there was no grouping effect on the pattern of errors except for an isolated effect in Experiment 1. Therefore, STM encoding seemed insufficient for LTM encoding.

What seems different about the triad condition in experiments in which an LTM effect of grouping was obtained, Experiments 3 and 4, is the superior encoding of the second triad, resulting in relatively little difference between the performance levels for the first and second triads. This finding suggests that the additional encoding time available for triads in these experiments was used to memorise each triad, perhaps by producing deeper, more semantic associations between items in a triad that are more likely to be retrievable from LTM compared to phonological representations (Craik & Tulving, 1975).

Unitary memory theories. Unitary memory theories, which do not distinguish between STM and LTM except as a result of the passage of filled time and consequent interference (e.g., Bjork & Whitten, 1974; Brown et al., 2007; Crowder, 1993; Gardiner & Gregg, 1979; Glenberg & Swanson, 1986; Nairne, 1990, 2002) might also account for the overall results in this study, but it will require at least some fine tuning. Broadly speaking, in this view, it is proposed that there is no difference between the mechanisms in play within STM and LTM tasks; differences in results are thought to occur because the context of the presentation needed for retrieval (any information about events taking place at the time of presentation of particular items to be remembered that can be used as retrieval cues) is still mostly present in the short term, but dissipates with the passage of time and/or interference from materials presented earlier, later, or in similar contexts to the targeted material. The derived prediction according to the view was the effects of stimulus grouping that are strong in the short term should show up in weakened form in the long term. If there are gaps between groups of items, these gaps distinguish the temporal context of the items on either side of the gap, but the distinction is relative to the time between presentation and test, so it is greatly diminished in delayed reconstruction.

The unitary memory view can account for our results that the effects of the stimulus grouping in Experiments 1, 2, 5, and 6 were not strong enough and were weakened either by interference and the resulting change in retrieval context below a usable level for the delayed reconstruction. However, when the effects of the stimulus grouping were stronger, in Experiments 3 and 4, the effect persisted in delayed reconstruction. On the contrary, unitary memory theorists should expect effects of grouping on delayed reconstruction even when items were presented one-by-one in temporal and spatiotemporal grouping, which did not occur in our Experiments 5 and 6. The duration of presentation of a triad were equally long in Experiments 3 and 5 and in addition there was a longer inter-triad times for Experiment 5 (see Table 1), yet grouping effects did not occur in the latter case. From the point of view of temporal distinctiveness this may be especially puzzling because stronger temporal cues to order within a triad were present in Experiment 5. On the contrary, non-temporal cues to triad group membership and order might have been stronger for Experiment 3. So, although the unitary memory theorists could explain our results, they would be relying heavily on non-temporal cues to grouping and order to do so, and these may be aspects of grouping that have not previously been addressed in unitary memory theories (see also Liu & Caplan, 2020).

Multimodal views of memory. A few theorists have explicitly addressed the issue of grouping in STM and LTM. In multicomponent models of working memory such as that of Baddeley (2000) and Vandierendonck (2016), there are different modules for different storage functions (separate ones for visuospatial and phonological storage) and for related processes to manage information in the stores. On one hand, there is strong evidence that the activity of the phonological buffer is critical for creation of long-term phonological memories (Baddeley et al., 1998). Nevertheless, some multicomponent views specify separate copies of information in STM and LTM (Norris, 2017, 2019) and the route from STM to LTM is not fully specified. A network modelling implementation of a multicomponent model of working memory has included explicit mechanisms for long-term learning (Burgess & Hitch, 1999, 2006). It showed that using the same temporal grouping of items in repetitions of the same list in immediate recall was important for learning the list across trials (the Hebb effect; Hebb, 1961). The model and evidence were later further reinforced (Hitch et al., 2009 see also Bower & Winzenz, 1969). This finding and model, however, still does not indicate whether a particular grouping is better than some other grouping, and it is not clear that the model can predict it without additional, proper simulations. A suggestion that it does is the statement that "the Hebb Effect will be sensitive to experimental manipulations affecting the context-timing signal" (Hitch et al., 2009, p. 100) that is used to encode the serial order of items.

Farrell (2012) provided a theoretical framework for thinking of STM and LTM as coming from the same mechanisms. Within that framework he explained grouping effects as follows (p. 246): "In common with unitary models, the model assumes that short- and LTM work in the same fashion and with the same machinery. Nonetheless, there is something akin to an STM mechanism in the model: If the group context does not have to be reinstated prior to retrieving the items in that group, recall of those items will be lent an advantage" (cf. Nairne, 1992). Instead of thinking of STM items as in an active state or in a special buffer, Farrell stated (p. 246) that "The alternative perspective offered by the current model is that the distinguishing feature of working memory is not in the storage or activation of items but rather the association of that information with the current episodic (i.e., group) context." This theoretical perspective leads to the notion that effects of grouping will be reduced in delayed performance rather than immediate performance, but the details of how this diminished effect of grouping will be manifested was not specified in the theory.

Embedded processes view. In the embedded processes view (Cowan, 1988, 1999, 2001, 2005/2016, 2010, 2019), a subset of information is said to be in a temporarily activated state embedded in LTM. In the subset of active information, a smaller subset is in the focus of attention, which is supposed to be quantitatively limited to approximately 3–4 separate meaningful units of information, or chunks, at once. A key process in this view is rapid long-term learning responsible for the formation of new associations that are needed on a trial, but were not previously present in LTM, such as the associations between list items and their serial positions (Cowan, 1999, 2019) A basic prediction derived from this view is that representations held concurrently in the focus of attention should be associated to become new units or events in LTM, observable if the initial encoding is sufficiently strong to be used despite the interference and change in context from immediate to delayed reconstruction.

This view is fairly closely aligned with the Atkinson and Shiffrin (1968) view. The results are consistent with embedded processes view and provide evidence that the long-term representation would sometimes be difficult to retrieve, as was observed in Experiments 1 and 2 but, with a sufficiently robust encoding, that LTM representation can be retrieved as observed in Experiments 3 and 4. However, the theory did not specify the basis of the additional processing that would occur given additional processing time. It is compatible with the notion that as compared to rapidly formed phonological representations, more slowly formed semantic representations and associations between items on that basis could be more important for LTM (Craik & Tulving, 1975).

Given that we espouse the embedded processes view, it is of special importance to point out a challenge to this view from the present findings. The challenge is to account for how the concurrent presence of items within a group could have produced a grouping effect in immediate reconstruction but not in delayed reconstruction, as we found in Experiments 1, 2, and 6. In these experiments, the presentation rate was rapid (1.8–6 s per list) whereas, when grouping effects in delayed reconstruction were obtained, in Experiments 3 and 4, the presentation was much slower (12 s per list). The associations between item may begin with arbitrary links that are prone to interference at the time of retrieval. With more time, the associations could become deeper. For example, if a triad consisted of the words *drab, wrap, beer*, it might take a few seconds to think of an integrated concept such as beer served in bottles with nondescript labels. This example suggests that it may be important to acknowledge a depth-of-processing principle (e.g., Craik, 2020; Craik & Tulving, 1975) within the embedded-processes approach, and this is grounds for further research.

Other theoretical views. Other theories could also be consistent with our results but, again, the results constrain those theories. Some theories, for instance, may not be clear on whether or not grouping in STM influences an LTM representation; our data specify conditions under which it does. Thus, there could be versions of the separate-copy theory (cf. Norris, 2017), temporary binding theory (Oberauer & Lin, 2017), and time-based resource sharing theory (Barrouillet & Camos, 2014) in which the STM representation used in immediate reconstruction does nothing to enhance LTM for the list, but such versions would have to be replaced with versions in which at least concurrent spatial presentation with a sufficiently long presentation period does produce an LTM representation, whereas certain other conditions might not.

One possibility within the time-based resource sharing theory is that we should expect that grouping should enhance LTM insofar as it influences refreshing, without regard to its potential influence on rehearsal. However, mixed support has been found for the influence of refreshing in LTM (see Camos et al., 2018 for a review). As an example, Camos and Portrat (2015) found initially that the refreshing mechanism resulted in improvement in LTM representations whereas verbal rehearsal did not, when participants had to recall the items. However, more recent evidence with young and older participants have highlighted that refreshing had no benefit on LTM representations (Bartsch et al., 2019; Loaiza & Camos, 2018) with a recognition procedure. Another possibility that was ruled out is that refreshing might involve elaborative processes that could benefit from concurrent spatial presentation. For instance, some early evidence from complex span tasks shows that refreshing produces LTM representations (see, for example, Loaiza & McCabe, 2012). However, Souza and Oberauer (2017) have shown that an alternative factor, processing time, was the determining factor in this LTM benefits. It is therefore unclear how the time-based resource sharing theory can account for the current set of findings. Overall, our results provide important constraints to time-based resource sharing theory.

Our results might also be consistent with other theoretical views. For instance, Popov and Reder (2020) proposed a resource-limited theory in which information in LTM is constrained by the limited resource available in working memory that is involved in encoding and later recuperation of information in LTM. Importantly, that limited resource recovers over time. Therefore, in Experiments 3 and 4, when presentation time was increased, it is possible that the resource had time to recover sufficiently after the presentation of the first triad to store the information of the second triad in LTM compared to when items were presented at a faster rate. However, it remained unclear within this view, why this benefit was not observed in Experiment 5 with a sequential presentation of items and a slower presentation time. Overall, our results provide valuable information to constrain and improve theories and models of memory.

Future directions

The current study highlights the theoretical value of studying the conditions that affect STM representations and how they will translate into LTM representations. This line of research is of high importance to the development of memory models. It is therefore unsurprising that there is a recent interest in this kind of research. For instance, working memory theorists are examining the consequences of working memory tasks for long-term learning using a variety of manipulations of the processing at the time of encoding and maintenance in working memory (e.g., Bartsch et al., 2018; Forsberg et al., 2021; Loaiza et al., 2011; Loaiza & McCabe, 2012; McCabe, 2008; Rose & Craik, 2012; Souza & Oberauer, 2017, 2018 for a recent review see Oberauer, 2019b).

Nevertheless, despite this recent interest, considerable further empirical testing will be needed before we can identify the best account of the present data. According to one theoretical account (Farrell, 2012), immediate recall involves participants' spontaneous, mental grouping of stimuli to facilitate recall but Spurgeon et al. (2015) showed that the model sometimes overpredicts the use of spontaneous grouping, and showed that stimulus grouping can help in immediate free and serial recall. Nevertheless, it is possible that one reason our effects of grouping on LTM were not larger is that some effects of mental grouping play a role even in the ungrouped control condition. Cowan et al. (2002) found that mental grouping was apparent using printed, but not spoken lists, so further work with spoken lists could be theoretically important and might show further effects of stimulus grouping on STM because that modality more thoroughly avoids mental grouping in the ungrouped control condition.

More work is also needed to understand our finding that spatiotemporal grouping with sufficient inter-group intervals seems to produce retrievable LTM memory traces, whereas simple temporal grouping does not. It could be that the spatiotemporal grouping initially is represented through item-to-context associations that distinguish one group from another and, within groups, a spatially represented sequential ordering (cf. Lee & Estes, 1981). Much of this representation might be rapidly lost but, along the way, it could give rise to item-to-item associations within a group that can be learned in a manner that no longer depends on context, either complete enough to be considered a newly learned chunk (Cowan, 2019), or only partially, as sets of weak or imperfect associative links (cf. Cowan et al., 2004). It is not yet possible to tell the difference between complete and partial item-to-item associations because even incomplete knowledge about groups in LTM may be enough to allow perfect ordering of a list on some delayed recall trials, and there may be stored knowledge that is not retrieved at the right moment.

There is a possibility that the disappearance of the grouping effect in delayed reconstruction in Experiments 1, 2, and 6 might be due to the loss of the item-to-context associations that distinguish one group from another. More precisely, immediate reconstruction testing before the delayed reconstruction test may have disturbed the essential retrieval context needed for the weak manipulation of grouping to benefit LTM. The slower presentation rate of Experiment 3, compared to the faster rate in Experiments 1 and 2, led to transferable grouping benefits in LTM despite the changing retrieval context. If the retrieval context remained intact between the list presentation and delayed test, as it did in Experiment 4 with spatiotemporal presentation, it might even allow sequential presentation (as in Experiment 6) to produce a grouping benefit in delayed reconstruction. Further studies should investigate this important issue of the emergence of delayed benefits of grouping depending on the retrieval context.

Conclusion

The current study does not prove any one view of information processing but it provides important new empirical constraints on all views, by providing an evaluation of when the grouping of items for an STM test is still helpful in a subsequent LTM test. Overall, our results provide evidence that at least concurrent presentation of items in a group with spatiotemporal grouping can be helpful for LTM representation. Furthermore, our manipulation uncovers an important factor in the usability of the LTM representation with concurrent presentation, namely presentation time. It is remarkable that even though participants engaged in immediate reconstruction of the stimuli, in some circumstances the grouping effects that assisted that immediate task were ineffective in subsequent delayed reconstruction. Although more evidence is needed to determine why the effects of grouping are not always the same in STM and LTM, our findings have important implications to constrain models of memory.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a Natural Sciences and Engineering Research Council of Canada graduate scholarship to D.G., a Discovery grant from the Natural Sciences and Engineering Research Council of Canada to J.S.-A., and NIH Grant R01-21338 to N.C.

ORCID iDs

Dominic Guitard https://orcid.org/0000-0002-4658-3585
Jean Saint-Aubin https://orcid.org/0000-0002-4799-6912
Nelson Cowan https://orcid.org/0000-0003-3711-4338

Data accessibility statement

The data and materials from the present experiment are publicly available at the Open Science Framework website: https://doi.org/10.17605/OSF.IO/B7UGC.

Supplementary material

The supplementary material is available at qjep.sagepub.com.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173. https://doi.org/10.1037/0033-295X.105.1.158

Barrouillet, P., & Camos, V. (2014). *Working memory: Loss and reconstruction*. Psychology Press.

Bartsch, L. M., Loaiza, V. M., Jäncke, L., Oberauer, K., & Lewis-Peacock, J. A. (2019). Dissociating refreshing and elaboration and their impacts on memory. *NeuroImage*, *199*, 585–597. https://doi.org/10.1016/j.neuroimage.2019.06.028

Bartsch, L. M., Singmann, H., & Oberauer, K. (2018). The effects of refreshing and elaboration on working memory performance, and their contributions to long-term memory formation. *Memory & Cognition*, *46*, 796–808. https://doi.org/10.3758/s13421-0180805-9

Bjork, R. A., & Whitten, W. B. (1974). Recency sensitive retrieval processes in long term free recall. *Cognitive Psychology*, *6*, 173–189. https://doi.org/10.1016/0010-0285(74)90009-7

Bower, G. H., & Winzenz, D. (1969). Group structure, coding, and memory for digit series. *Journal of Experimental Psychology*, *80*(2 Pt 2), 1–17. https://doi.org/10.1037/h0027249

Broadbent, D. E., & Broadbent, M. H. (1981). Articulatory suppression and the grouping of successive stimuli. *Psychological Research*, *431*(1), 57–67. https://doi.org/10.1007/BF00309638

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539–576. https://doi.org/10.1037/0033295X.114.3.539

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*, 551–581. https://doi.org/10.1037/0033-295X.106.3.551

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, *55*, 627–652. https://doi.org/10.1016/j.jml.2006.08.005

Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, *1424*, 19–32. https://doi.org/10.1111/nyas.13616

Camos, V., & Portrat, S. (2015). The impact of cognitive load on delayed recall. *Psychonomic Bulletin & Review*, *22*, 1029–1034. https://doi.org/10.3758/s13423-014-0772-5

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, A*, *33*, 497–505. https://doi.org/10.1080/14640748108400805

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, *104*, 163–191. https://doi.org/10.1037/0033-2909.104.2.163

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.006

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2016). *Working memory capacity*. Psychology Press. (Original work published 2005)

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*, 51–57. https://doi.org/10.1177/0963721409359277

Cowan, N. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin*, *145*, 822–847. https://doi.org/10.1037/bul0000108

Cowan, N., Chen, Z., & Rouder, J. N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science*, *15*, 634–640. https://doi.org/10.1111/j.0956-7976.2004.00732.x

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*, 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

Cowan, N., Saults, J. S., Elliott, E. M., & Moreno, M. (2002). Deconfounding serial recall. *Journal of Memory and Language*, *46*, 153–177. https://doi.org/10.1006/jmla.2001.2805

Craik, F. I. M. (2020). Remembering: An activity of mind and brain. *Annual Review of Psychology*, *71*, 1–24. https://doi.org/10.1146/annurev-psych-010419-051027

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition*, *21*, 142–145. https://doi.org/10.3758/BF03202725

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, *119*, 223–271. https://doi.org/10.1037/a0027371

Forsberg, A., Guitard, D., & Cowan, N. (2021). Working memory limits severely constrain long-term retention. *Psychonomic Bulletin & Review*, *28*, 537–547. https://doi.org/10.3758/s13423-020-01847-z

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language*, *24*(2), 200–209. https://doi.org/10.1016/0749-596X(85)90024-5

Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 469–479. https://doi.org/10.1037/0278-7393.15.3.469

Gardiner, J. M., & Gregg, V. H. (1979). When auditory memory is not overwritten. *Journal of Verbal Learning & Verbal Behavior*, *18*, 705–719. https://doi.org/10.1016/S0022-5371(79)90411-0

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning & Verbal Behavior*, *5*(4), 351–360. https://doi.org/10.1016/S0022-5371(66)80044-0

Glenberg, A. M., & Swanson, N. C. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *12*, 3–15. https://doi.org/10.1037/0278-7393.12.1.3

Guérard, K., Saint-Aubin, J., Boucher, P., & Tremblay, S. (2011). The role of awareness in anticipation and recall performance in the Hebb repetition paradigm: Implications for sequence learning. *Memory & Cognition*, *396*(6), 1012–1022. https://doi.org/10.3758/s13421-011-0084-1

Hartley, T., Hurlstone, M. J., & Hitch, G. J. (2016). Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism. *Cognitive Psychology*, *87*, 135–178. https://doi.org/10.1016/j.cogpsych.2016.05.001

Healy, A. F., Havas, D. A., & Parker, J. T. (2000). Comparing serial position effects in semantic and episodic memory using reconstruction of order tasks. *Journal of Memory and Language*, *42*, 147–167. https://doi.org/10.1006/jmla.1999.2671

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). Blackwell.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 116–139. https://doi.org/10.1080/713755609

Hitch, G. J., Flude, B., & Burgess, N. (2009). Slave to the rhythm: Experimental tests of a model for verbal short-term memory and long-term sequence learning. *Journal of Memory and Language*, *61*, 97–111. https://doi.org/10.1016/j.jml.2009.02.004

Jiang, Q., & Cowan, N. (2020). Incidental learning of list membership is affected by serial position in the list. *Memory*, *28*, 669–676. https://doi.org/10.1080/09658211.2020.1761398

Kuhn, J. R., Lohnas, L. J., & Kahana, M. J. (2018). A spacing account of negative recency in final free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1180–1185. https://doi.org/10.1037/xlm0000491

Lawrence, M. A. (2016). *ez: Easy analysis and visualization of factorial experiments* (R package version 4.4-0). https://CRAN.R-project.org/package=ez

Lee, C. L., & Estes, W. K. (1981). Item and order information in short-term memory: Evidence for multilevel perturbation processes. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 149–169. https://doi.org/10.1037/0278-7393.7.3.149

Liu, Y. S., & Caplan, J. B. (2020). Temporal grouping and direction of serial recall. *Memory & Cognition*, *48*, 1295–1315. https://doi.org/10.3758/s13421-020-01049-x

Loaiza, V. M., & Camos, V. (2018). The role of semantic representations in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 863–881. https://doi.org/10.1037/xlm0000475

Loaiza, V. M., & McCabe, D. P. (2012). Temporal–contextual processing in working memory: Evidence from delayed cued recall and delayed free recall tests. *Memory & Cognition*, *40*, 191–203. https://doi.org/10.3758/s13421-011-0148-2

Loaiza, V. M., McCabe, D. P., Youngblood, J. L., Rose, N. S., & Myerson, J. (2011). The influence of levels of processing on recall from working memory and delayed recall tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1258–1263. https://doi.org/10.1037/a0023923

McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language*, *58*, 480–494. https://doi.org/10.1016/j.jml.2007.04.004

McElree, B. (1998). Attended and non attended states in working memory: Accessing categorized structures. *Journal of Memory and Language*, *38*, 225–252. https://doi.org/10.1006/jmla.1997.2545

Medler, D. A., & Binder, J. R. (2005). MCWord: An on-line orthographic database of the English language. http://www.neuro.mcw.edu/mcword/

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*, 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes Factors for common designs* (R package version 0.9.12-4.2). https://CRAN.R-project.org/package=BayesFactor

Nairne, J. S. (1990). Similarity and long-term memory for order. *Journal of Memory and Language*, *29*, 733–746. https://doi.org/10.1016/0749-596X(90)90046-3

Nairne, J. S. (1992). The loss of positional certainty in long term memory. *Psychological Science*, *3*, 199–202. https://doi.org/10.1111/j.1467-9280.1992.tb00027.x

Nairne, J. S. (2002). Remembering over the short term: The case against the standard model. *Annual Review of Psychology*, *53*, 53–81. https://doi.org/10.1146/annurev.psych.53.100901.135131

Nairne, J. S., & Neath, I. (2001). Long term memory span. *Behavioral and Brain Sciences*, *24*, 134–135. https://doi.org/10.1017/S0140525X01003922

Norris, D. (2017). Short-term memory and long-term memory are still different. *Psychological Bulletin*, *143*, 992–1009. https://doi.org/10.1037/bul0000108

Norris, D. (2019). Even an activated long-term memory system still needs a separate short-term store: A reply to Cowan (2019). *Psychological Bulletin*, *145*, 848–853. https://doi.org/10.1037/bul0000204

Norris, D., Kalm, K., & Hall, J. (2020). Chunking and redintegration in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*, 872–893. https://doi.org/10.1037/xlm0000762

Oberauer, K. (2019a). Is rehearsal an effective maintenance strategy for working memory? *Trends in Cognitive Sciences*, *23*, 798–809. https://doi.org/10.1016/j.tics.2019.06.002

Oberauer, K. (2019b). Working memory capacity limits memory for bindings. *Journal of Cognition*, *2*, 1–13. https://doi.org/10.5334/joc.86

Oberauer, K., & Bialkova, S. (2009). Accessing information in working memory: Can the focus of attention grasp two elements at the same time? *Journal of Experimental Psychology: General*, *138*(1), 64–87. https://doi.org/10.1037/a0014738

Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review*, *124*, 21–59. https://doi.org/10.1037/rev0000044

Öztekin, I., Davachi, L., & McElree, B. (2010). Are representations in working memory distinct from representations in long-term memory? Neural evidence in support of a single store. *Psychological Science*, *21*(8), 1123–1133. https://doi.org/10.1177/0956797610376651

Popov, V., & Reder, L. M. (2020). Frequency effects on memory: A resource-limited theory. *Psychological Review*, *127*(1), 1–46. https://doi.org/10.1037/rev0000161

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rose, N. S., & Craik, F. I. M. (2012). A processing approach to the working memory/long-term memory distinction: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1019–1029. https://doi.org/10.1037/a0026976

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*, 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237. https://doi.org/10.3758/PBR.16.2.225

Ryan, J. (1969a). Grouping and short-term memory: Different means and patterns of grouping. *The Quarterly Journal of Experimental Psychology*, *21*, 137–147. https://doi.org/10.1080/14640746908400206

Ryan, J. (1969b). Temporal grouping, rehearsal and short-term memory. *Quarterly Journal of Experimental Psychology*, *21*, 148–155. https://doi.org/10.1080/14640746908400207

Severin, F. T., & Rigby, M. K. (1963). Influence of digit grouping on memory for telephone numbers. *Journal of Applied Psychology*, *47*, 117–119. https://doi.org/10.1037/h0045301

Soemer, A., & Saito, S. (2016). Domain-specific processing in short-term serial order memory. *Journal of Memory and Language*, *88*, 1–17. https://doi.org/10.1016/j.jml.2015.12.003

Souza, A. S., & Oberauer, K. (2017). Time to process information in working memory improves episodic memory. *Journal of Memory and Language*, *96*, 155–167. https://doi.org/10.1016/j.jml.2017.07.002

Souza, A. S., & Oberauer, K. (2018). Does articulatory rehearsal help immediate serial recall? *Cognitive Psychology*, *107*, 1–21. https://doi.org/10.1016/j.cogpsych.2018.09.002

Spurgeon, J., Ward, G., Matthews, W. J., & Farrell, S. (2015). Can the effects of temporal grouping explain the similarities and differences between free recall and serial recall? *Memory & Cognition*, *43*, 469–488. https://doi.org/10.3758/s13421-014-0471-5

Stadler, M. A. (1993). Implicit serial learning: Questions inspired by Hebb (1961). *Memorv & Cognition*, *21*, 819–827. https://doi.org/10.3758/bf03202749

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*, 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolKit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*, 24–31. https://doi.org/10.1177/0098628316677643

Sutterer, D. W., Foster, J. J., Adam, K. C., Vogel, E. K., & Awh, E. (2019). Item-specific delay activity demonstrates concurrent storage of multiple active neural representations in working memory. *PLoS biology*, *17*(4), e3000239. https://doi.org/10.1371/journal.pbio.3000239

Sukegawa, M., Ueda, Y., & Saito, S. (2019). The effects of Hebb repetition learning and temporal grouping in immediate serial recall of spatial location. *Memory & Cognition*, *474*(4), 643–657. https://doi.org/10.3758/s13421-019-00921-9

Thompson, V. A., & Campbell, J. I. D. (2004). A power struggle: Between- vs within-subjects designs in deductive reasoning research. *Psychologia: An International Journal of Psychology in the Orient*, *47*(4), 277–296. https://doi.org/10.2117/psysoc.2004.277

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104132. https://doi.org/10.1037/0033-295X.114.1.104

Vandierendonck, A. (2016). A working memory system with distributed executive control. *Perspectives on Psychological Science*, *11*, 74–100. https://doi.org/10.1177/1745691615596790

Wickelgren, W. A. (1964). Size of rehearsal group and short-term memory. *Journal of Experimental Psychology*, *68*, 413–419. https://doi.org/10.1037/h0043584