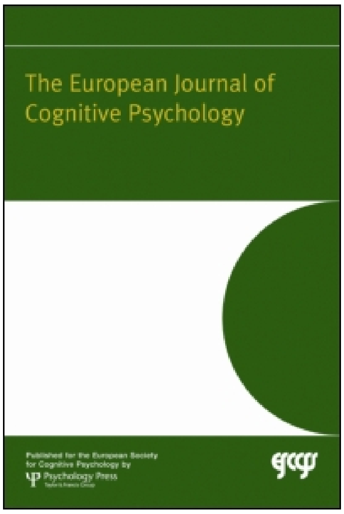


European Journal of Cognitive Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713734596

Spatial clustering and hierarchical coding in immediate serial recall

Carlo De Lillo [a]; Valerie E. Lesk [a]
[a] University of Leicester, Leicester, UK

First Published on: 27 July 2009

To cite this Article De Lillo, Carlo and Lesk, Valerie E.(2009)'Spatial clustering and hierarchical coding in immediate serial recall',European Journal of Cognitive Psychology,99999:1,

To link to this Article: DOI: 10.1080/09541440902757918

URL: http://dx.doi.org/10.1080/09541440902757918

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY
0000, 00 (00), 1–31

Psychology Press
Taylor & Francis Group

Spatial clustering and hierarchical coding in immediate serial recall

Carlo De Lillo and Valerie E. Lesk
University of Leicester, Leicester, UK

Four experiments are reported that demonstrate the benefits of clustering by spatial proximity in spatial serial recall and provide support for the notion that hierarchical coding underpins the retention of clustered sequences in spatial working memory. Sequences segregated by spatial clusters increased serial recall performance at different levels of sequence length in a variation of the Corsi test and produced a faster initial response time (RT), which indicates that they afforded data reducing processes. RT at cluster boundary increased in parallel with the number of items forming the clusters, suggesting that subroutines of different length were responsible for the ordering of items within clusters of different size. Evidence for hierarchical coding was also obtained in a serial recognition task, indicating this type of representation pertains to the retention of the sequences rather than exclusively to the organisation of the motor plan for the reproduction of the sequences.

Keywords: Chunking; Serial recall; Short-term memory; Spatial memory; Spatial organisation.

The ability to exploit the structure in the layout of the environment is considered a variable of extreme importance in several areas of cognitive research (De Lillo, Aversano, Tuci, & Visalberghi, 1998; De Lillo, Aversano, & Visalberghi, 1997; Kirsh, 1995). However, in studies of human spatial working memory relatively little attention has been dedicated to the understanding of how the use of spatial constraints can affect the organisation of serial recall. Human spatial working memory is typically assessed using the Corsi test (Corsi, 1972; Milner, 1971), which requires participants to

---

Correspondence should be addressed to Carlo De Lillo, School of Psychology, University of Leicester, Henry Wellcome Building, Lancaster Road, Leicester LE1 9HN, UK. E-mail: cdl2@le.ac.uk

This research was supported by Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/C007840/1 to CDL. We are grateful to Tony Andrews for developing the software used in this study. We would also like to thank Clelia Rossi-Arnaud and an anonymous reviewer of a previous version of the manuscript whose comments greatly helped to improve this article.

http://www.psypress.com/ecp DOI: 10.1080/09541440902757918

reproduce observed tapping sequences on trays of wooden blocks or computer monitors. It has been shown that intrinsic spatial coordinates are used for coding the position of each item in a Corsi display (Avons, 2007; Avons & Trew, 2006) and that the display is processed configurationally (Boduroglu & Shah, 2006). Nevertheless, the display of items in the Corsi test is typically irregular (see Berch, Krikorian, & Huba, 1998 for a review) and, with some notable exceptions discussed later, systematic assessments of how spatial structure affects recall are rare.

Kemps (1999) addressed the problem of the role of spatial structure within Corsi displays using the notions of quantitative and structural complexity. Quantitative complexity refers to the number of items in the display. Sequences of the same length are more difficult to reproduce if presented within larger arrays that contain some irrelevant items in addition to those in the sequence to be reported. Structural complexity refers to the particular arrangement of the blocks in the display irrespectively of the serial order in which they have to be reported and less complex arrays organised as square matrices generate a higher span than more complex arrays formed by randomly arranged locations (Kemps, 1999). Structural complexity is related to Gestalt notions of pattern "goodness", which can be quantified in terms of redundancy in information theory approaches to visual perception (e.g., Attneave, 1954; Garner, 1974). Simpler stimuli are more redundant and the higher span observed for sequences presented in simpler displays could be explained in terms of reduced information load.

The study by Kemp (1999) clearly indicates the importance of the spatial arrangement of the material used to test serial recall. However, it focused on the static structure of the Corsi board and explicitly avoided the issue of the relationship between the structure of the array and the serial organisation of the sequence to be reproduced. Yet, other studies show that even in relatively unstructured Corsi boards, different sequences are reproduced with a different level of accuracy (Smirni, Villardita, & Zappala, 1983). It is, therefore, important to evaluate the effect of the relationship between the *serial* organisation of the sequence to be reproduced and the *spatial* structure of the display which is used to present it. In fact, it has been shown (Bor, Duncan, Wiseman, & Owen, 2003) that in a Corsi display featuring a square matrix of locations, sequences that follow linear constraints (i.e., sequences where consecutive items appear within the same row, column, or diagonal) are reproduced at a higher level of accuracy. It has been suggested that this may be due to the fact that sequences conforming to the linear organisation of a square matrix generate meaningful visual structures which are easier to encode in memory. Moreover, the processing of these types of structures seems to have a specific neural substrate since their presentation produces a selective activation of the dorsolateral prefrontal cortex (Bor et al., 2003).

Other studies (De Lillo, 2004) have looked at the effect of segregation by spatial clusters of sequences presented within a display of items grouped on the basis of their spatial proximity and showed that recall accuracy is higher when consecutive items are in the same cluster rather than in different clusters. Moreover, RT analyses suggest that, in analogy with results reported in the literature on chunking in other domains (Klahr, Chase, & Lovelace, 1983; Povel & Collard, 1982) or on the organisation of motor sequences in rapid speech and typewriting (Sternberg, Knoll, Monsell, & Wright, 1988; Sternberg, Knoll, & Turock, 1990), participants form hierarchical representations of clustered Corsi sequences and that this type of organisation facilitates recall (De Lillo, 2004). The notion of "hierarchical representation" in this context refers to a representation whereby the order of report of higher order units (e.g., spatial clusters) is stored and can be accessed separately from that of individual items which is specified at a subordinate level.

Parmentier, Andres, Elford, and Jones (2006) replicated the findings of De Lillo (2004) by showing a superior serial recall for sequences segregated by spatial clusters in a similar task but also noted that the movement path followed during the presentation of sequences segregated by clusters is shorter than that of sequences which violate a clustering principle. It is, therefore, important to characterise better the role of hierarchical organisation (HO) in the immediate reproduction of spatial sequences in situations where the role of structure is not confounded by variables such as path length. We aimed to do so in the present study by focusing on the effect played by cluster size and number of clusters in which a sequence is segregated in relation to accuracy and timing during serial recall. The rationale for this being that if effects of sequence composition can be observed in conditions where path length does not vary systematically with one particular type of organisation, then the organisation of the sequence, and the type of representation that it affords, should be considered as an important determinant of serial spatial memory performance in their own right.

The analysis of RT in relation to the structure of the sequences to be reproduced is of particular interest. In fact, RT has been used for the characterisation of the hierarchical nature of motor plans in tasks requiring the reproduction of word sequences in speech and typewriting (Sternberg et al., 1988, 1990). In this domain, evidence for HO derives from the fact that sequence initiation time is a function of the number of words in the sequence, whereas RT at cluster boundary is a function of the number of syllables comprised in each word. In this model, the initial response time would be an expression of the time taken to access the motor plan for the entire sequence, whereas the length of the interval between the recall of consecutive words would be an expression of the time taken to recall a subroutine specifying the ordering of syllables within a single word.

Fischer (2001) demonstrated the usefulness of measuring RT in the standard Corsi test by showing that initial RT is proportional to the number of steps in the sequence for spatial sequences too. In Experiments 2 and 3 of the present study we aimed to use time measures to clarify the role of HO in the retention of clustered sequences by using sequences with the same number of sequential steps but segregated by clusters of different sizes. Evidence for RT at cluster boundary that is proportional to the number of items in each cluster would give support for the HO of sequences in clustered versions of the Corsi test.

Experiment 1 aimed, first of all, at corroborating the results of De Lillo (2004), by assessing whether sequences segregated by clusters are reproduced better than sequences not segregated by clusters. Moreover, it aimed to establish whether the incentive to deploy the organisational factors, which eventually lead to the benefits of clustering, only emerges when sequence length exceeds memory span. Finally, as the presence and numerosity of irrelevant display items has been shown to reduce the benefits of imposing structure in a Corsi-type task (Kemps, 1999) and produce changes in the serial position curve in other spatial recall tasks (Farrand, Parmentier, & Jones, 2001), in Experiment 1 we monitored the possible effects of this variable by manipulating the presence/absence of irrelevant additional items during the presentation of clustered and unclustered sequences.

EXPERIMENT 1

Methods

Participants

Twenty five participants (average age 23.21 years, $SD = 4.01$, reporting normal or corrected to normal vision) were initially recruited and paid £5 for taking part. However, two of them were discarded because of computer error during the presentation of the experiment.

Apparatus and stimuli

A PC equipped with a 19-inch ELO touch-sensitive monitor was used to display the stimulus sequences and to collect the participant's responses. The stimuli were grey square icons with a 4 cm side presented on a black background. Their possible location was defined within a virtual display of four spatial clusters of four locations each (see Figure 1a, c, e, g). The overall size of the array was 25 cm × 29 cm with a distance of 1 cm between items in the same cluster and a minimum distance of 4 cm between items in different clusters. The experiment was controlled using an E-prime (Psychology Software Tools Inc.) program, which enabled the presentation of a sequence of icons that "blinked" (disappeared for 0.5 s before reappearing again on the screen) and the recording of the timing and the order in which icons were touched during recall.

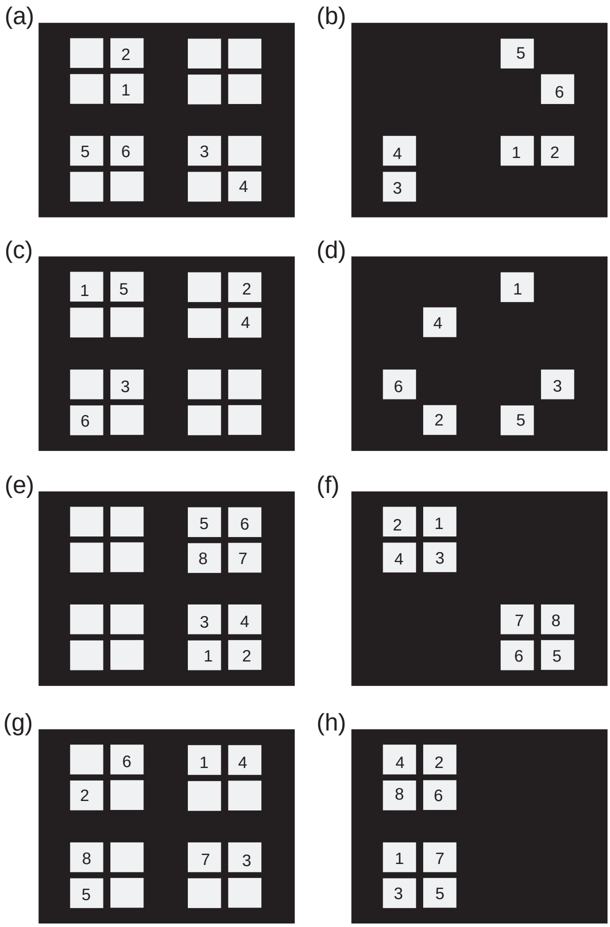


Figure 1. Examples of trials for some of the conditions used in Experiment 1: (a) AIP 3C $\times$2I; (b) AIA 3C $\times$2I; (c) AIP 0C $\times$6I; (d) AIA 0C $\times$6I; (e) AIP 2C $\times$4I; (f) AIA 2C $\times$4I; (g) AIP 0C $\times$8I; (h) AIA 0C $\times$8I. The numbers on the icons are used in the figure to refer to the order in which the items blinked and had to be touched during serial recall and were not part of the presentation during the experiment. For reasons of space examples for the conditions AIP 2C $\times$3I, AIA 2C $\times$3I, AIP 4C $\times$8I, and AIA 4C $\times$2I are not shown in the figure. See text for further explanations.

Design

Each participant received all the experimental conditions (described later). The variables manipulated in the experiment were: (1) the length of the sequence; (2) its structure; and (3) the presence or absence of irrelevant additional icons (AI), which were presented on the screen but did not pertain to the sequence to be reproduced. Sequence length was manipulated by using sequences of six and eight ordinal steps. These values correspond roughly to the average standard Corsi span and to a supraspan value, respectively (Kessels, Zandvoort, Postma, Kappelle, & de Haan, 2000; Vandierendonck, Kemps, Fastane, & Szmalec, 2004). The structure of the sequence was manipulated by presenting, for each level of sequence length, sequences segregated by clusters (clustered sequences), which in turn could be segregated in clusters of different size and numbers, and sequences that were not segregated by clusters (nonclustered sequences). In particular, sequences of length 6, were segregated into two clusters of three icons each (2C × 3I condition), three clusters of two icons (3C × 2I condition), or could be nonclustered (0C × 6i condition). Sequences of length 8 were divided into two clusters of four icons (2C × 4I condition), four clusters of two icons (4C × 2I condition), or were nonclustered (0C × 8I condition). The effect of the presence of AI was tested by presenting, for each of the conditions mentioned above, trials with "additional icons present" (AIP) and of trials with "additional icons absent" (AIA). In the AIP trials, 16 icons were displayed during both the presentation of the sequence and during serial recall, irrespectively of the number of items featured in the sequence to be reported. Thus, in this condition participants had to identify which items appeared in the original sequence as well as serial order. In the AIA condition only the icons used in the sequence were displayed. Therefore, only serial order had to be encoded in this condition. Examples of trials of different conditions are presented in Figure 1a–h.

One 120-trial testing session comprised AIP trials only and another session AIA trials only. The use of separate sessions for the presentation of AIP and AIA trials was preferred over the alternative of interspersing them in the same testing session in order to avoid potential additional effects induced by having to switch continuously and without warning between two types of task. The order of presentation of the two sessions was counter-balanced across participants. Within one session, 60 trials featured sequences of length six and comprised 20 trials for each of the three sequence type (2C × 3I, 3C × 2I, and 0C × 6I). The other 60 trials comprised 20 trials for each of the sequence type conditions (2C × 4I, 4C × 3I, and 0C × 8I) of length 8. Different trials were presented in a pseudorandom order in which trials of the same type could not appear more than twice consecutively.

Procedure

Participants were tested individually in a quiet room. They were asked to sit comfortably and at a distance which enabled them to easily reach for any corner of the screen. After instructions, the experiment began. Each trial started with the presentation of the display of icons for 1 s. The icons then started to "blink" at a rate of one item per 1.2 s until the entire sequence was presented. At this point, a blank screen was presented for a 1 s interval, during which the participant had to retain the sequence in memory. A static icon display was then presented and the participant had to report the sequence by touching the screen. Following recall, either correct or incorrect, a blank screen was presented for 1 s, before the presentation of the subsequent trial. To ensure that the finger movement required to perform the initial response always started from a fixed location and at a constant distance from the screen, during this 1 s interval participants were asked to keep their index finger on the central key of a response box.

Results

Accuracy

The percentage of items correctly recalled is reported in Figure 2.

Sequence length 6. A repeated measures ANOVA 3 (type of sequence: $2C \times 3I$ vs. $3C \times 3I$ vs. $0C \times 6I$) $\times$ 2 (presence of additional icons: AIP vs. AIA) $\times$ 6 (serial position of item) carried out on the percentage of correct items revealed a significant main effect for type of sequence, $F(2, 44) = 91.35$, $p < .001$. There was a significant difference between each type of sequence with the lowest score registered for the $0C \times 6I$ condition and the highest for the $3C \times 2I$ condition, $2.15 < t(22) < 11.38$, all $p < .05$. This trend can be observed in Figure 2c. The presence of AI also significantly affected recall, $F(1, 22) = 28.00$, $p < .001$, with higher scores recorded for the AIA conditions (Figure 2c). The interactions of type of sequence by serial position, $F(10, 220) = 22.72$, $p < .001$, and of presence of AI by serial position, $F(5, 110) = 5.00$, $p < .001$, were both significant but the interaction presence of AI by type of sequence was not (Figure 2a).

Sequence length 8. Unsurprisingly, the percentage of correct items was lower, $t(22) = 19.91$, $p < .001$, for sequences of length 8 than for sequences of length 6. Despite this overall difference in accuracy as a function of sequence length, the analysis of the differences between the different conditions within the eight item sequences, revealed a pattern of results similar to that observed with six item sequences. In fact, a repeated measures ANOVA 3 (type of sequence) $\times$ 2 (AIP/AIA) $\times$ 8 (serial position) revealed

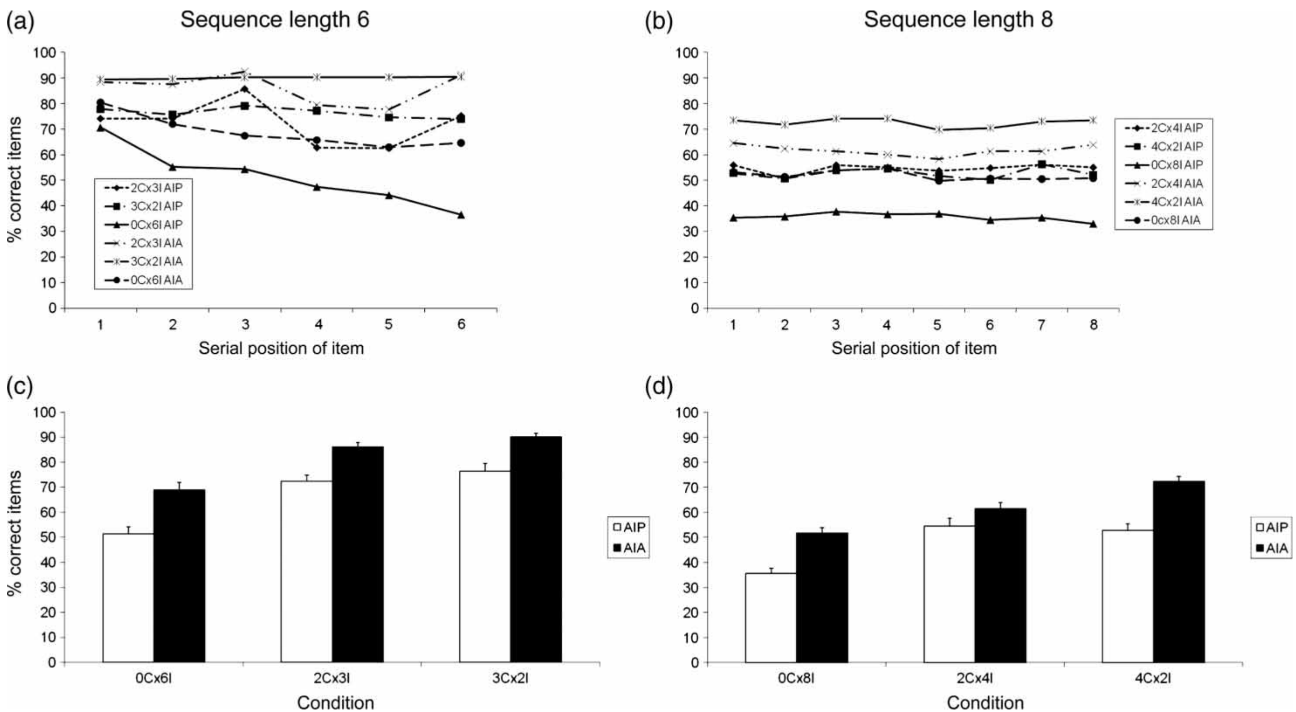


Figure 2. Accuracy values observed in Experiment 1: (a) Percentage of items correctly recalled at each serial position in the different sequences of length 6 with additional icons present (2C × 3I AIP, 3C × 2I AIP, 0C × 6I AIP) and additional icons absent (3C × 2I AIA, 3C × 2I AIA, 0C × 6IAIA); (b) percentage of items correctly recalled at each serial position in the different sequences of length 8 with additional icons present (2C × 4I AIP, 4C × 2I AIP, 0C × 8I AIP) and absent (2C × 4I AIA, 4C × 2I AIA, 0C × 8I AIA); (c) overall percentage of correct items recorded for the different sequences of length 6 with additional icons present (AIP) and absent (AIA); (d) overall percentage of correct items recorded for the different sequences of length 8 with AIP and with AIA.

an effect of condition, $F(2, 44) = 110.68$, $p < .001$, which, as for the six item sequences, was due to a difference between each of the conditions, $3.51 < t(22) > 16.29$, all $p < .05$. There was also an effect of AI with more items recalled in the AIA conditions (Figure 2d). An overall main effect for serial position, $F(7, 154) = 2.58$, $p < .05$, also emerged but only the interaction type of sequence by presence of AI was significant, $F(2, 44) = 9.52$, $p < .001$ (Figure 2b).

RT

RT analyses aimed at providing information about how the structure of the sequence affects its encoding in working memory. If sequences are encoded differently as a function of their structure, then a different pattern of RT should emerge for each sequence type and the following specific predictions were made. First, as the time taken to initiate the sequence is typically dependent on the number of items in the sequence (see Fischer, 2001; Sternberg et al., 1988, 1990), we should observe that the initial RT should be shorter for clustered sequences, if the better retention of clustered sequences is mainly due to data reduction. Second, if the encoding of clustered sequences is based on a hierarchical representation whereby the order of report of items within a cluster is encoded at a subordinate level, then we should expect that RT at cluster boundaries should be affected by the size of the cluster to be reported.

Time measures were based on RT for items correctly recalled. Data that were two standard deviations or more either above or below the mean (4.2% of the data from the $3C \times 4I$ condition and 4.6% of the data from the $4C \times 3I$ condition overall) were removed as outliers. The data set before outliers were removed showed the same pattern of results as the data set with outliers removed. Figure 3 shows RT registered for each serial position in the different conditions and Table 1a summarises mean RT at critical points of serial recall. A visual inspection of the graphs shows the longest RT at the beginning of the sequences, as predicted if subjects plan the execution of the sequence at that stage.

RT for sequences of length 6 was analysed using an ANOVA 3 (type of sequence) $\times$ 2 (presence of additional icons) $\times$ 6 (serial position) and revealed a main effect of condition, $F(2, 44) = 42.34$, $p < .001$, with the $2C \times 3I$ condition producing about the same RT as the $3C \times 2I$ condition and the $0C \times 6I$ condition producing the slowest RT. There was also an effect of serial position, $F(5, 110) = 260.61$, $p < .001$, and a significant Condition $\times$ Serial position interaction, $F(10, 220) = 28.70$, $p < .001$. No other main effects or interactions were significant. The interaction condition by serial position is easily explained by the peaks at cluster boundaries which occur at different serial positions in the two clustered conditions.

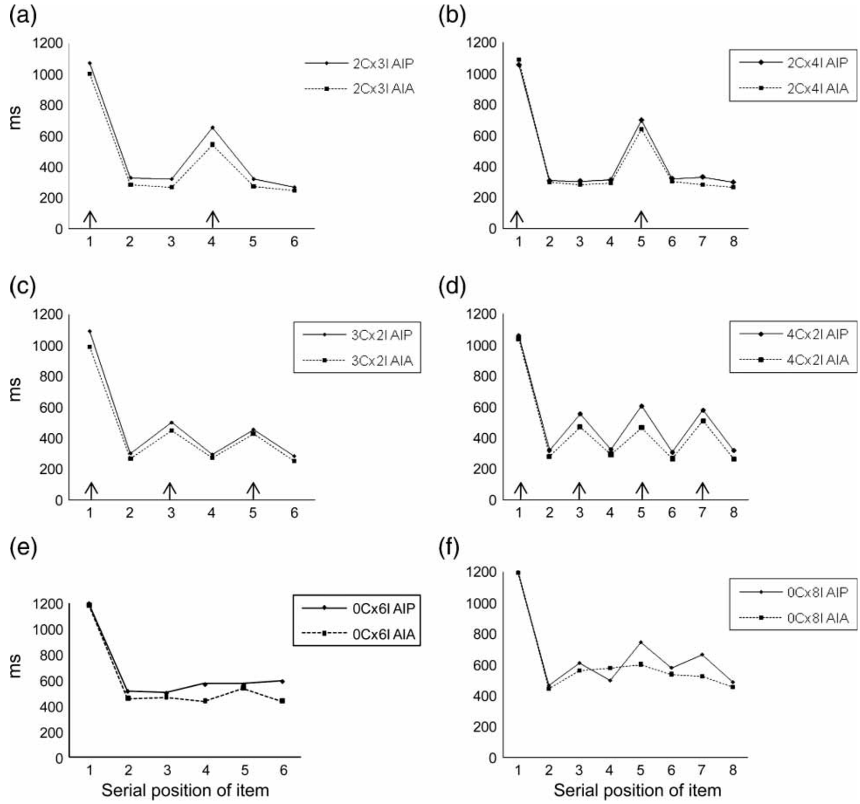


Figure 3. Response time (RT) recorded for correct touches at each ordinal position during serial recall of sequences of the different conditions featured in Experiment 1: (a) 2C × 3I AIP and 2C × 3I AIA; (b) 2C × 4I AIP and 2C × 4I AIA; (c) 3C × 2I AIP and 3C × 2I AIA; (d) 4C × 2I AIP and 4C × 2I AIA; (e) 0C × 6I AIP and 0C × 6I AIA; (f) 0C × 8I AIP and 0C × 8I AIA. In conditions with sequences segregated by clusters, arrows indicate the serial position corresponding to the first item within each cluster.

In clustered conditions, a longer RT was recorded for serial positions at cluster boundaries compared to serial positions within cluster boundaries, $4.42 < t < 14.58$, $p < .01$ with Bonferroni correction. By contrast, as the nonclustered conditions do not explicitly afford a segmentation of the sequence on the basis of spatiotemporal patterns, conspicuous peaks cannot be detected in the curves for the 0C × 6I (Figure 3e) and for the 0C × 8I (Figure 3f) conditions. In fact, in the nonclustered conditions none of the peaks at serial positions 2–6 compared using *t*-tests yielded statistical significance.

A similar pattern of results was obtained for the RT observed in sequences of length 8. In fact, a repeated measures ANOVA 3 (type of sequence) × 2

TABLE 1a
Response time (RT, in ms) observed in Experiment 1

Nonclustered sequences	*Mean RT*	*Initial RT*	*RT within sequence*	
0C × 6I	623.9 (29.1)	1193.4 (79.8)	510.1 (50.4)	
0C × 8I	628.5 (25.5)	1179.14 (78.5)	549.81 (56.5)	
Clustered sequences	*Mean RT*	*Initial RT*	*RT CB*	*RT within CB*
2C × 3I	466.3(18.3)	1036.4 (75.5)	600.3 (39.0)	294.8 (21.2)
3C × 2I	468.3 (13.5)	1043.0 (65.5)	461.2 (14.6)	285.3 (11.1)
2C × 4I	435.1 (14.4)	1006.8 (70.1)	669.6 (37.6)	301.8 (10.6)
4C × 2I	476.0 (14.4)	1040.7 (58.8)	530.6 (22.2)	314.0 (12.1)

Mean RT = mean of RT for all serial positions; Initial RT = average RT registered for the first item in the sequence; RT within sequence = average of all RT excluding RT for serial position 1 in the nonclustered sequences; RT CB = average RT at cluster boundary (i.e., RT at serial position 4 for the 2C × 3I condition; mean RT at serial positions 3, and 5 for the 3C × 2I; RT at serial position 5 for the 2C × 4I condition and mean RT at positions 3, 5, and 7 for the 4C × 2I condition); RT within CB = average RT within cluster boundary (i.e., average of RT at serial position 2, 3, 5, and 6 for the 2C × 3I condition; average RT at serial position 2, 4, and 6 for the 3C × 2I condition; average RT at serial position 2, 3, 4, 6, 7, and 8 for the 2C × 4I condition and average RT at serial position 2, 4, 6, and 8 for the 4C × 2I condition. Reported in parentheses are standard errors.

(presence of additional icons) × 8 (serial position) revealed a main effect of type of sequence, $F(2, 40) = 101.08$, $p < .001$, with the 2C × 4I registering faster RT than the 4C × 2I and the 0C × 8I condition registering the slowest RT (mean = 628.48 ms, $SE = 25.54$ ms). There was also an effect of serial position, $F(7, 140) = 164.46$, $p < .001$. The presence of a significant Condition × Serial position interaction, $F(14, 280) = 12.631$, $p < .001$, could also be easily explained by the peaks at different serial positions in the various conditions. In the clustered conditions peaks at cluster boundary were always higher than values within cluster boundaries, $3.39 < t < 9.50$, $p < .01$, with Bonferroni correction. In the nonclustered conditions none of the peaks at serial positions 2–8 compared using t-tests yielded statistical significance, with the only exception of a significant difference between RT at serial position 5 and 2, $t(22) = 4.17$, $p < .01$. None of the other main effects or interactions was significant.

Peaks at cluster boundaries in sequences segregated by clusters are consistent with the notion of a hierarchical representation for these sequences but their presence can be partially explained by the longer movements required to touch icons located in different clusters rather than within the same cluster. Further information derives from a comparison of the different conditions in terms of their initial RT and RT cluster boundaries. In fact, since the participants always started the sequence from a fixed hand position and since the average length of the movement required to move between clusters in the different conditions was not systematically affected by the type of condition, these comparisons could not be affected by movement length. We used paired sample *t*-tests with Bonferroni correction for these comparisons. As for the time analyses presented earlier, the presence of AI did not produce a significant effect and given the lack of interaction between serial position and the presence of AI, RT for the AIA and AIP conditions was combined.

The results of the tests are summarised in Table 1b. For both sequences of length 6 and of length 8, the initial RT in the two clustered conditions was significantly shorter than in the nonclustered conditions but no difference in this value emerged between the two clustered conditions. RT at cluster boundaries was longer for the conditions featuring larger clusters in both six item and eight item sequences.

Discussion

Sequences segregated by clusters always produced a higher level of recall. Within clustered sequences, an increase in the size of the clusters in the sequence negatively affected recall. The presence of irrelevant items negatively affected recall in both clustered and nonclustered sequences.

Whereas accuracy data indicate the benefits of imposing particular forms of serial spatial organisation on the sequences to be reported, RT analyses provide information concerning how sequences of different composition are represented in memory and why some sequences are easier to reproduce than others. Consistently with the presence of a hierarchical representation for clustered sequences, initial RT was faster for the clustered than for nonclustered sequences. This effect would be explained by the fact that fewer

TABLE 1b
Planned comparisons between different conditions of Experiment 1 for RT recorded at critical points of serial recall

Comparison	*t*	*p*
Initial RT		
2C × 3I vs. 0C × 6I	4.02 (22)	< .01
3C × 2I vs. 0C × 6I	3.95 (22)	< .01
2C × 3I vs. 3C × 2I	.318 (22)	*ns*
2C × 4I vs. 0C × 8I	4.79 (22)	< .01
4C × 2I vs. 0C × 8I	6.32 (22)	< .01
2C × 4I vs. 4C × 2I	1.00 (22)	*ns*
RT at cluster boundary		
3C × 2I vs. 2C × 3I	4.54 (22)	< .01
4C × 2I vs. 2C × 4I	4.53 (22)	< .01

In parentheses are reported the degrees of freedom.

decision nodes need to be traversed before serial recall of these sequences can be initiated. RT at cluster boundary was always longer for sequences segregated in larger clusters, consistently with a model of serial representation where subroutines specifying the order of selection of items within a particular superordinate unit are called upon during recall when the appropriate point of the serial representation is accessed (Sternberg et al., 1988, 1990). However, one potentially confusing factor in this experiment is the presence of conditions featuring clusters of two items only. These conditions may be a special case of clustered sequences affording other data reduction strategies such as retaining in memory only the first item of each clusters and deriving the second by default during sequence reproduction. Therefore, a second experiment focused on clustered sequences of twelve icons segregated into three clusters of four items each or into four clusters of three items each.

EXPERIMENT 2

Method

Participants

Thirty undergraduates (average age of 20.6 years, $SD = 4.1$), reporting normal or corrected to normal vision, were given course credits for taking part in Experiment 2.

Apparatus and material

Apparatus and individual icons were the same as in Experiment 1. However, 12 icons were used for Experiment 2 and were arranged either as three clusters with four icons each ($3C \times 4I$ condition) or four spatial clusters of three icons each ($4C \times 3I$ condition). The two types of display are presented in Figure 4 and measurements provided in the caption.

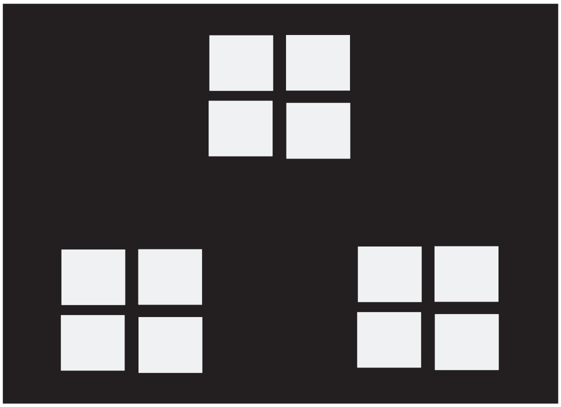
Design

Each participant received 30 $3C \times 4I$ trials and 30 $4C \times 3I$ trials randomly interspersed within a 60-trial testing session.

Procedure

The same general procedure used in Experiment 1. Distances between consecutive items at cluster boundary for all the sequences in the $3C \times 4I$ (mean = 17.90 cm, $SD = 3.07$ cm) and the $4C \times 3I$ (mean = 18.90 cm, $SD =$ 3.47 cm) condition were measured and compared using an independent-samples t-test which did not prove significant, $t\ (58) = -0.95$, *ns.* Thus,

(a)



(b)

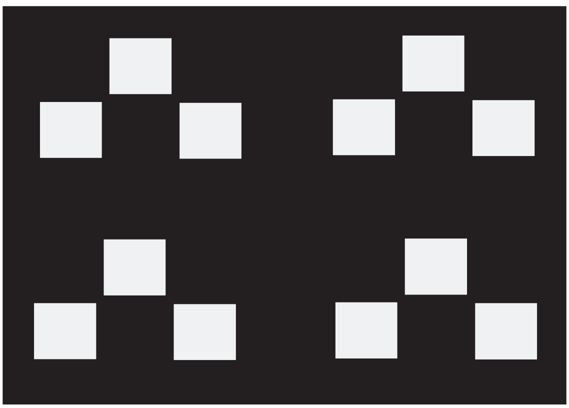


Figure 4. The two displays used in Experiment 2: (a) three clusters of four icons each (3C ×4I); (b) four clusters of three icons each (4C ×3I). In the 3C ×4I display, the shortest distance within the clusters was 5 cm and the longest was 7 cm. The shortest distance between icons of different clusters was 11 cm and the longest was 27 cm. For the 4C ×3I sequence the shortest distance within the clusters was 7 cm and the longest was 9 cm. The shortest distance between icons of different clusters was 10 cm and the longest was 30 cm.

systematic differences in the length of the finger movement required in the two conditions which could have spuriously affected RT at cluster boundaries were unlikely to be present.

Results

Accuracy

The percentage of correct items for each serial position in two conditions is reported in Figure 5a. A repeated measures ANOVA, 2 (type of

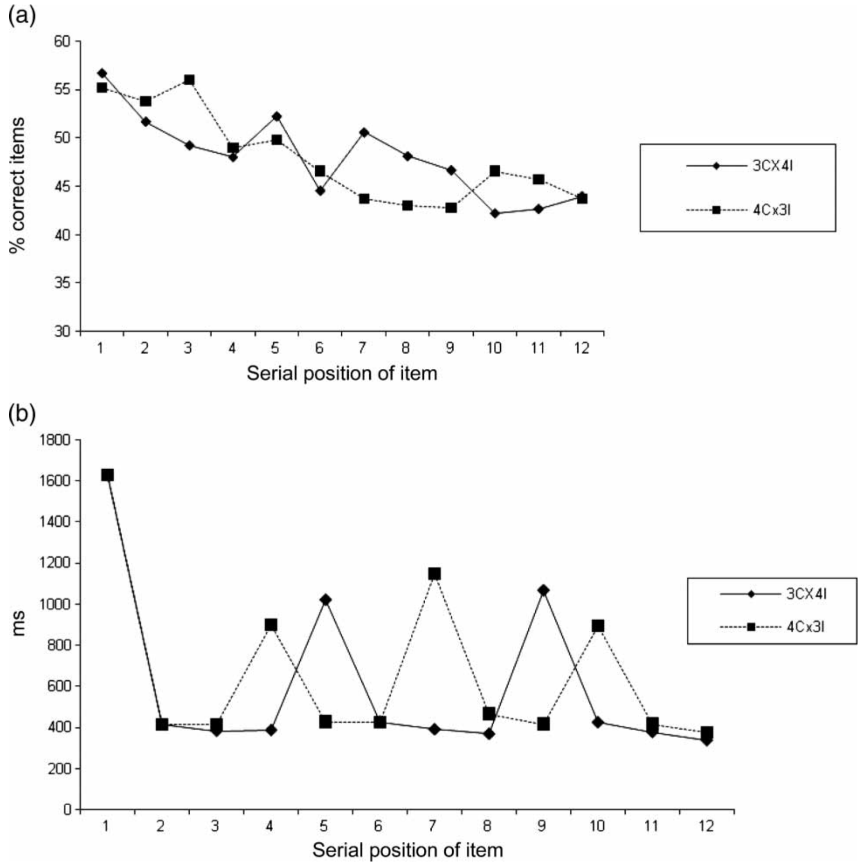


Figure 5. Results of Experiment 2: (a) percentage of correct responses recorded for each serial position in the 3C × 4I and the 4C × 3I conditions; (b) RT corresponding to correct responses recorded for each serial position in the 3C × 4I and the 4C × 3I conditions; (c) mean RT recorded in the two conditions.

sequence) × 12 (serial position) indicated there was no difference, $F(1, 29) = 0.002$, *ns* between the 3C × 4I and the 4C × 3I condition. The main effect for serial position, $F(11, 319) = 9.36$, $p < .001$, was significant but no clear primacy or recency effect are evident in the serial position curve (see Figure 5a). The interaction condition by serial position, $F(11, 319) = 3.16$, $p < .00$, was also significant.

RT

Outliers (5.6% of values for the 3C × 4I and 5.3% for the 4C × 3I condition) were removed from the analysis as in Experiment 1. However, the pattern of results was the same with and without outliers. Only correct responses were included in the analysis. An ANOVA 2 (type of sequence) × 12 (serial position) revealed a significant effect of type of sequence, $F(1, 29) = 39.50$, $p < .001$, with faster RT for the 3C × 4I condition than for the 4C × 3I condition. Thus, an increase in the number of clusters slowed recall (see values in Table 2a).

The ANOVA also revealed a significant main effect of serial position, $F(11, 319) = 1190.80$, $p < .001$. The values of RT for each serial position are presented in Figure 5b. *T*-tests with Bonferroni correction used for planned comparisons revealed that the main effect of serial position was due to longer RT at serial position 1 compared to serial positions at cluster boundary and that RT at cluster boundary were in turn longer than within cluster boundary (all $ps < .05$) in the two conditions (see Figure 5b).

The presence of a longer RT in the condition featuring four clusters (see Table 2a) could not be explained merely by the fact that it had one additional peak at cluster boundary compared to the condition featuring three clusters. In fact, planned comparisons carried out between the two conditions using the average RT of only touches within cluster boundaries revealed that the difference was maintained even at this level and thus that the condition with the higher number of clusters produced the slowest RT for individual items (see Table 2a and 2b for means and *t*-test results).

A significant interaction was also observed between sequence type and serial position, $F(11, 319) = 422.32$, $p < .001$. As for Experiment 1, this significant interaction can be easily explained by the presence of RT peaks at different serial positions in the two conditions (see Figure 5b).

RT at cluster boundary was longer for the condition with clusters of four icons (see Table 2a and 2b). According to a hierarchical model of sequence reproduction, the initial RT can be considered as the sum of two separate components: one related to the time taken to decide the serial order of the clusters in the sequence (superordinate) and one related to the time taken

TABLE 2a
Response time (RT, in ms) observed in Experiment 2

Condition	*Mean RT*	*Initial RT*	*RT CB*	*RT within CB*	*RT1–RTCB*
3C × 4I	600.6 (58.8)	1625.8 (171.4)	1043.0 (117.9)	388.5 (42.9)	646.1 (116.9)
4C × 3I	659.9 (67.0)	1626.0 (159.2)	979.9 (125.9)	419.2 (44.7)	582.81 (141.1)

Mean RT = mean of RT for all serial positions; RTCB = average RT at cluster boundary (i.e., average of RT at serial positions 5 and 9 for the 3C × 4I condition and at serial positions 4, 7, and 10 for the 4C × 3I condition); RT within CB = average RT within cluster boundary (i.e., average of RT at serial position 2, 3, 4, 6, 7, 8, 10, 11, and 12 for the 3C × 4I condition and of serial position 2, 3, 5, 6, 8, 9, 11, and 12 for the 4C × 3I condition; RT1 – RTCB = difference of RT at serial position 1 minus RT at cluster boundary. Reported in parentheses are standard deviations.

TABLE 2b
Experiments 2 and 3: Planned comparisons between critical RT values of different conditions

Comparison	*t*	*p*
RT within cluster boundary		
3C × 4I vs. 4C × 3I (Experiment 2)	5.44 (29)	<.01
3C × 4I vs. 4C × 3I (Experiment 3)	3.41 (7)	< 01
RT at cluster boundary		
3C × 4I vs. 4C × 3I (Experiment 2)	2.75 (29)	<.01
3C × 4I vs. 4C × 3I (Experiment 3)	2.40 (7)	<.05
Initial RT minus RT at cluster boundary		
3C × 4I vs. 4C × 3I (Experiment 2)	2.18 (29)	<.05
3C × 4I vs. 4C × 3I (Experiment 3)	2.56 (7)	<.05

The degrees of freedom are reported in parentheses.

to decide the order of report of the icons within the first cluster (subordinate). This latter component should be similar in length to the RT observed at cluster boundary (as RT at cluster boundary should only be an expression of the time taken to decide the order in which the icons within a cluster need to be reported). Therefore, we should be able to estimate the time taken to decide the order of report at the superordinate level by subtracting, separately for each condition, the mean RT at cluster boundary from the initial RT. Once this measure has been obtained, we can evaluate a prediction of the hierarchical model, namely, that the difference between the initial RT and the RT at cluster boundary should be longer for sequences containing more clusters. The value of the initial RT minus the value of the RT at cluster boundary for the two conditions is reported in Table 2a. A paired-samples *t*-test confirmed that this value was higher for the 3C × 4I than for the 4C × 3I (see Table 2b for *t*-test results).

Discussion

The different pattern of RT observed for the two sequence types are consistent with hierarchical models of the representation of these types of sequences. A first interesting finding is that sequences segregated in four clusters produced a longer average RT than sequences segregated in three clusters. In analogy with what has been suggested for other types of serially organised motor productions from memory (Sternberg et al., 1988, 1990), this would indicate that individual responses within a sequence are affected by the global properties of the sequence. In the present context this result suggests that spatial working memory is sensitive to the serial-spatial segregation of the material to be reported. RT at cluster boundary was longer for the condition with clusters of more items than for the condition with clusters with fewer items. Thus, in accordance with our predictions, there is a positive relationship between the time taken by the participants to decide in which order the items within a cluster had to be reported and the size of the cluster. Also consistent with a hierarchical model is our second and related finding concerning the time taken by the participants in selecting the order of report of the different clusters (rather than the icons within a cluster). When we isolated the time taken to process this level of the hierarchy, by subtracting the RT at cluster boundary from the initial RT, we found that this value was significantly longer for sequences containing more clusters.

It was important to use 12 item sequences in this experiment as this is the minimum number of icons where the size and numerosity of clusters can be manipulated so that there are always more than two items within a cluster and more than two clusters in the sequence. In fact the absence of a significant difference in accuracy here would indicate that the superior performance observed in the first experiment in conditions featuring clusters of two icons only may indeed have been due to the fact that these conditions require the encoding of only one item per cluster, with the second item derived by default. However, with twelve items to recall only a low number of sequences were correctly reproduced in their entirety. As only correct responses allow a meaningful analysis of RT, our analyses were based on items reported in the correct order within otherwise incorrectly reproduced sequences. Although this may not have been a problem in itself, we considered it appropriate to verify our results in situations where a more affluent set of correct responses were recorded and where it was possible to conduct a time analysis on touches recorded within correctly reproduced sequences. We did so in our third experiment.

EXPERIMENT 3

In Experiment 3 we used the same paradigm as in Experiment 2 but we selected proficient participants and gave them extensive task practice in order to generate a better performance overall which would have allowed an analysis of RT in sequences reported correctly in their entirety. Moreover, the protracted testing given in Experiment 3 allowed us to analyse possible changes in the pattern of results due to tasks practice.

Method

Participants

Eight undergraduates (average age of 19.6 years, $SD = 1.5$) with normal or corrected to normal vision took part. They were selected on the basis of their high performance in Experiment 2 or in a similar experiment (De Lillo & Lesk, 2006, 2009). They were given either course credits or were paid £5.

Apparatus and materials

These were the same as used for Experiment 2.

Design and procedure

Apart from the presentation of additional trials, the design and procedure of Experiment 3 were the same used in Experiment 2. In Experiment 3, each testing session comprised 100 trials: 50 of the 3C × 4I condition and 50 of the 4C × 3I condition, presented in random order without replacement. Participants received 400 trials in four consecutive daily testing sessions.

Results

Accuracy

Participants reproduced entire sequences correctly on 38% of the 3C × 4I trials and 43% of the 4C × 3I trials. An ANOVA 2 (type of sequence) × 2 (session) carried out on these values showed that recall in the two conditions was not significantly different, $F(1, 7) = 1.70$, *ns*. There was a main effect of session, $F(3, 21) = 175.66$, $p < .001$, with significant linear, $F(1, 7) = 53.45$, $p < .001$, and quadratic, $F(1, 7) = 7.54$, $p < .05$, trends indicating that performance improved with task practice (see Figure 6a).

RT

Only RT for responses within sequences reproduced correctly in their entirety were analysed. Outliers were removed (3.5% of the values for the 3C × 4I and 3.4% for the 4C × 3I condition) as for previous experiments. A repeated measures ANOVA 2 (type of sequence) × 12 (serial position) × 4 (session) revealed a main effect for type of sequence, $F(1, 6) = 15.29$, $p < .01$, with shorter RT for the 3C × 4I than for the 4C × 3I condition. Mean RT values are reported in Table 3.

RT recorded for each serial position during each testing session in the two conditions are reported in Figure 6b and 6c. A significant main effect emerged for the factors serial position, $F(11, 66) = 49.64$, $p < .001$, and session, $F(3, 18) = 10.61$, $p < .001$. The first order interactions Condition × Serial position, $F(11, 66) = 26.69$, $p < .001$, and Session × Serial position, $F(33, 198) = 2.80$, $p < .001$, were both significant, as it was the third order interaction Condition × Serial position × Session, $F(33, 198) = 1.80$, $p < .01$. Planned comparisons (see Table 3 for mean values), carried out using *t*-tests with Bonferroni correction, revealed that the RT at cluster boundary was

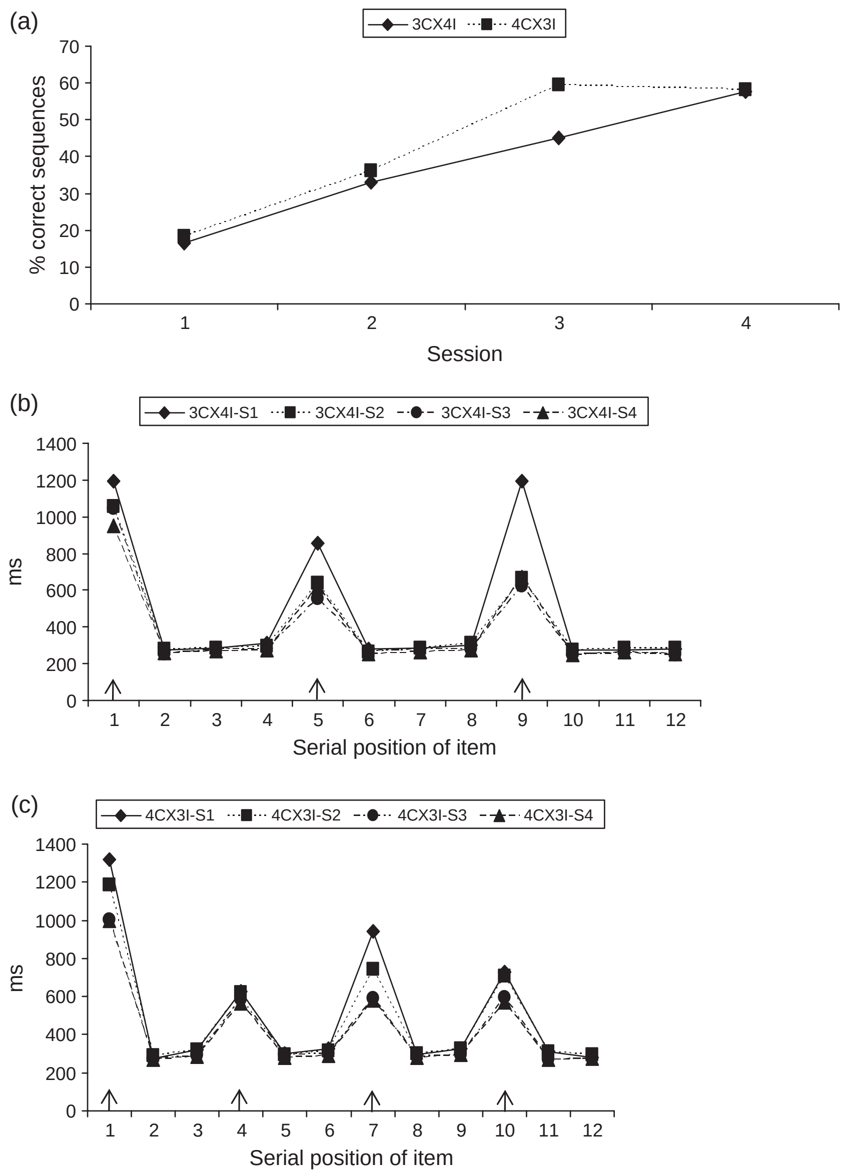


Figure 6. Results of Experiment 3: (a) percentage of correct sequences correctly reproduced in each of the four testing sessions presented for the 3C $\times$4I and the 4C $\times$3I conditions; (b) 3C $\times$4I condition: mean RT recorded for each serial position in sequences correctly reproduced during each of the four testing sessions, arrows indicate the serial position corresponding to the first item within each cluster; (c) 4C $\times$3I condition: mean RT recorded for each serial position in sequences correctly reproduced during each of the four testing sessions, arrows indicate the serial position corresponding to the first item within each cluster.

TABLE 3
Response time (RT, in ms) observed in Experiment 3

Condition	*Mean RT*	*Initial RT*	*RT CB*	*RT within CB*	*RT1 – RTCB*
3C × 4I	408.7 (76.2)	1061.0 (295.8)	690.8 (208.3)	275.00 (53.7)	348.5 (198.6)
4C × 3I	455.5 (76.1)	1125.9 (328.1)	629.3 (165.2)	294.6 (16.7)	447.24 (26.4)

Mean RT = mean of RT for all serial positions; RTCB = average RT at cluster boundary (i.e., average of RT at serial positions 5 and 9 for the 3C × 4I condition and at serial positions 4, 7, and 10 for the 4C × 3I condition); RT within CB = average RT within cluster boundary (i.e., average of RT at serial position 2, 3, 4, 6, 7, 8, 10, 11, and 12 for the 3C × 4I condition and of serial position 2, 3, 5, 6, 8, 9, 11, and 12 for the 4C × 3I condition; RT1 – RTCB = difference of RT at serial position 1 minus RT at cluster boundary. Reported in parentheses are standard deviations.

longer than the RT within cluster boundary and the RT recorded for serial position 1 was in turn longer than the RT at cluster boundary for both the 3C × 4I and the 4C × 3I condition (all *p*s $< .01$). As for previous experiments, the different serial position of the RT peaks in the two conditions easily explains the significant Condition × Serial position interaction. A trend analysis carried out for the factor session revealed a significant linear component, which indicates that participants were getting faster with task practice, $F(1, 6) = 19.86$, $p < .01$. As can be seen from Figure 6b and 6c, and consistently with the results of Experiment 2, for both conditions the initial RT was the slowest and peaks in RT can also be observed at cluster boundary. The mean RT at cluster boundary and within cluster boundary for the 4C × 3I condition and the 3C × 4I are reported in Table 3. RT at cluster boundary was slower for larger clusters. By contrast, as for Experiment 2, RT within cluster boundary was longer for condition 4C × 3I, which contains more clusters. The value obtained by subtracting RT at cluster boundary from the initial RT in each of the two conditions is reported in Table 3. It was longer for the 3C × 4I condition than for the 4C × 3I condition (see Table 2b for *t*-test results).

Discussion

Experiment 3 allowed us to evaluate the robustness of the results of Experiment 2, in a situation where data based on entire sequences correctly reproduced could be analysed. This was important as it was possible that when correct touches were derived from sequences containing some mistakes, RTs may not have been an accurate measure of the processes underpinning the errorless retention of entire sequences. Overall, the results of Experiment 3 confirmed the major findings of Experiment 2. Thus, further support was provided for the notion that the memory representation of spatially clustered sequences is affected by the size and the number of clusters and it is likely to be hierarchically organised.

The results of the first three experiments, however, need clarification in at least two important respects. The first is the fact that the two measures that we used for the assessment of the role of HO are not independent from each other. The initial RT minus the RT at cluster boundary is a measure that is obviously affected by the value of RT at cluster boundary. The significant increase of RT for larger clusters that we found at cluster boundary could therefore have determined the effect that we found when comparing initial RT in the two conditions. Although the observed effect at cluster boundary in itself would be difficult to explain without considering that spatial working memory is sensitive to the serial-spatial organisation of the sequences, an additional independent test of this assertion would provide further support to the notion of HO.

Furthermore, the observation of patterns of RT consistent with a hierarchical model in a serial task such as that featured in Experiments 1, 2, and 3 do not allow us to assess whether the inferred HO pertains to the memory representation of the sequence, independently from the planning of the serial motor output necessary during the reproduction of the sequences, or rather to the motor planning of the sequence to be executed, or both. In Experiment 4 we sought further support from the notion of HO in serial spatial memory in a recognition task not involving serial recall and the motor planning processes associated with it.

EXPERIMENT 4

As HO often pertains to the domain of motor control (see Sternberg et al., 1988, 1990), it is important to assess whether the evidence for this type of representation observed in the experiments described previously is related to the retention of serial spatial information and not only to processes related to the organisation of motor output.

Thus, in Experiment 4 we used a paradigm based on recognition that does not require the production of long sequences of motor responses. Following the presentation of sequences segregated by clusters, Experiment 4 featured the presentation of different testing conditions which, if participants had coded the sequences hierarchically, would have required access to either the superordinate or subordinate level of the hierarchical representation, in order to provide a “yes” or “no” answer. As explained in more detail later, different patterns of RT would have been expected in the two conditions according to hierarchical and nonhierarchical models of serial order representation.

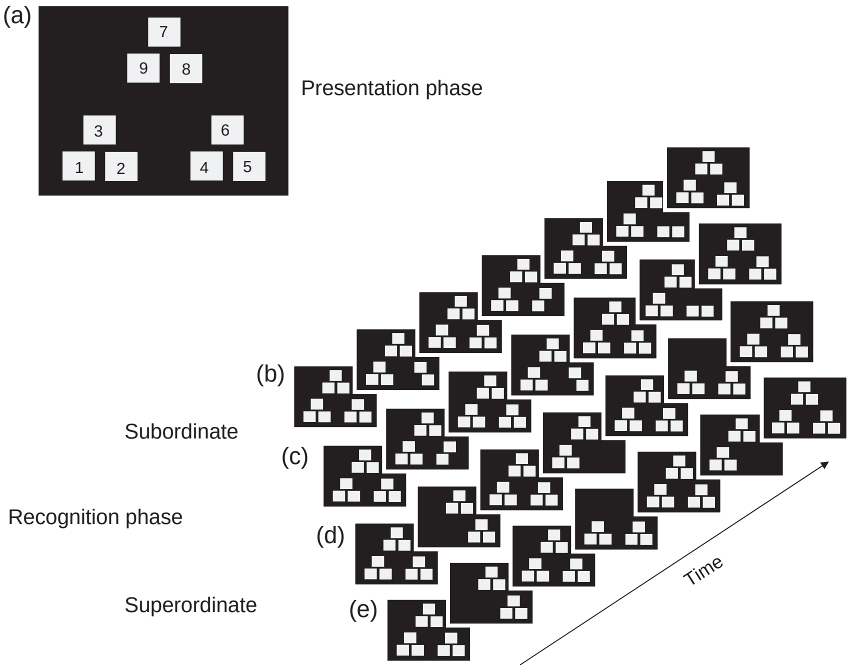


Figure 7. Examples of displays and sequences used in Experiment 4: (a) An example of the presentation of a sequence: numbers indicate the order in which the icons blink; (b) example of a "subordinate" trial requiring the response "Y". for the sequence presented in (a); (c) example of a "subordinate" trial requiring the response "N" for the sequence presented in (a); (d) example of a "superordinate" trial requiring the response "Y" for the sequence presented in (a); (e) example of a "superordinate" trial requiring the response "N" for the sequence presented in (a).

Methods

Participants

Thirty-one undergraduates with an average age of 19.2 ($SD = 1.3$) and normal or corrected to normal vision took part in exchange for course credits.

Apparatus and materials

The same PC and icons used for the previous experiments. In Experiment 4, a display of nine icons arranged in three clusters of three icons each was used (Figure 7). According to the procedure described next, sequences of blinking icons had to be recognised by the participants in a subsequent recognition phase. A response box was used to collect responses. Participants had to press a button labelled Y, for "Yes", when they recognised the

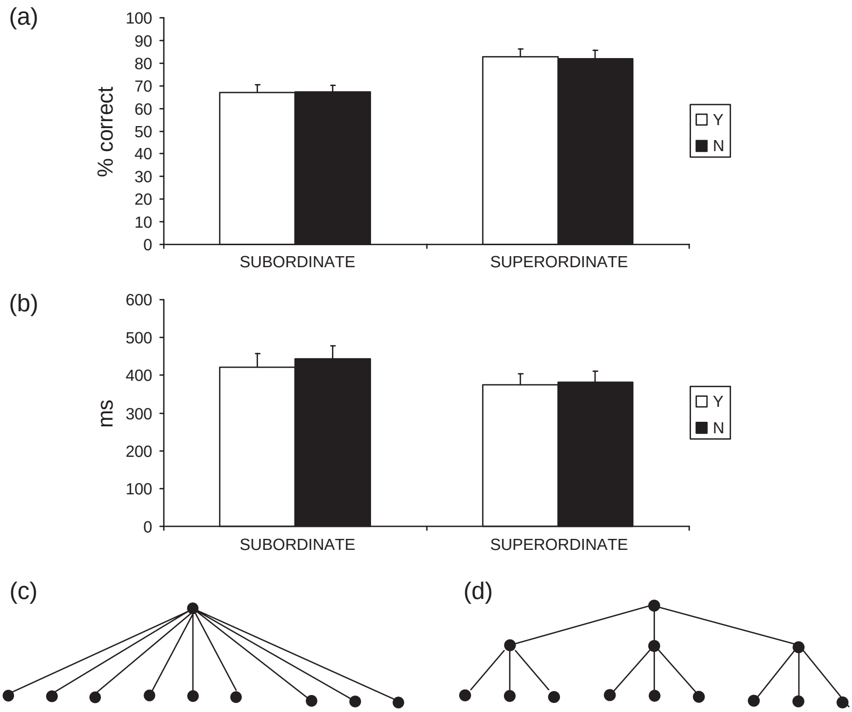


Figure 8. Experiment 4: (a) percentage of correct responses recorded for Y and N trials in the subordinate and superordinate conditions; (b) average RT registered for correct responses for Y and N trials in the two conditions; (c) tree structure of a nonhierarchical representation of the sequences; (d) tree structure of a hierarchical representation of the sequences.

previously presented sequence and a button labelled N, for "No", when they did not.

Design and procedure

Nine-step sequences of blinking icons were presented which were all segregated by spatial clusters (see Figure 7a for an example). Following a 2 s delay, a recognition test ensued, which could conform to either a subordinate or a superordinate condition.

Subordinate condition. The test involved presenting a sequence of three blinking icons within one and only one particular cluster. Participants were required to decide whether the order of blinking icons in that particular cluster matched the order of blinking icons in the same cluster in the presentation phase, and press one of the keys labelled Y for "yes" or N for "no" on the response box, accordingly. This type of test is illustrated in Figure 7b–c.

Superordinate condition. The test involved showing a sequence where the three clusters blinked in a particular order (all the icons within a cluster blinking simultaneously) and the participants were required to judge whether this order matched the order in which each cluster (irrespectively of the order in which each icon blinked within a particular cluster) was traversed in the presentation phase (see Figure 7d–e), by pressing the "Y" or "N" key on the response box. This type of test is illustrated in Figure 7d–e.

A total of 144 trials were administered: 72 trials (50%) featured a recognition test of the subordinate level and the other 72 of the superordinate level. For each of the subordinate and superordinate conditions, 36 trials (50%) presented the same sequence as shown in the presentation phase and therefore required a "Y" response and 36 trials featured a different sequence as that previously shown and required an "N" response. The hierarchical model predicted shorter response times in the superordinate condition compared to the subordinate condition.

Results

Accuracy

Figure 8a shows the percentage of correct responses observed in the superordinate and in the subordinate conditions. An ANOVA 2 (condition: subordinate/superordinate) $\times$ 2 (response type: Y/N) carried out on these values revealed a significant main effect of condition, $F(1, 28) = 17.53$, $p < .001$. Participants were more accurate in recognising superordinate than subordinate sequences. Y and N responses produced the same level of accuracy, $F(1, 28) = 0.02$, *ns*.

RT

Only RT for correct responses was analysed. Figure 8b shows RT observed in the superordinate and the subordinate conditions. An ANOVA 2 (condition) $\times$ 2 (response type) revealed a significant main effect of condition, $F(1, 28) = 5.43$, $p < .05$. Superordinate sequences were recognised faster than subordinate sequences. "Y" and "N" responses were produced at the same speed, $F(1, 28) = 2.27$, *ns*.

Discussion

Using clustered sequences, different predictions for a nonhierarchical model (see Figure 8c) and a hierarchical model (see Figure 8d) of memory representation can be made in terms of the pattern of RT observed in test conditions requiring the recognition of the order of presentation of clusters of items irrespectively of the order of presentation of individual items within clusters (superordinate condition) or the order of presentation of individual items within a cluster, independently of the order of presentation of the clusters (subordinate condition). A hierarchical model predicts shorter response times in the superordinate condition as the participants will only need to travel through the three upper nodes of the representation to access the necessary information. By contrast, longer response times will be expected for the subordinate condition as at least four nodes will need to be accessed (one superordinate node to access the relevant cluster and three subordinate nodes to access information about item order within that cluster).

A nonhierarchical model would, by contrast, predict that the longest response time should be observed in the superordinate condition. In fact, a minimum of seven nodes needs to be traversed in the nonhierarchical model to access information concerning the order in which the three clusters were interrogated (three nodes for the first cluster, three for the second, and at least the first node of the third). By contrast, the time required to produce correct responses in the subordinate condition should be shorter as an average of six nodes (depending on the position of the relevant icons, sometimes three, sometimes six, and sometimes nine nodes will need to be interrogated, thus an average of three when the position of the test items is balanced) would need to be interrogated to retrieve the relevant information. It is important to note that, although different nonhierarchical models have been proposed for the representation of serial order in short term memory, such as associative chaining, positional coding, and ordinal theory (see Henson, 2001), all of them would predict a shorter response time for the subordinate test in the present paradigm.

Thus, our results showing faster RT in the superordinate then in the subordinate condition are consistent with hierarchical models of serial spatial recall, which predict that fewer nodes of a hierarchical representation need to be traversed to retrieve the relevant information at the superordinate level rather than at the subordinate level.

Moreover, as the procedure used in this experiment did not require the participants to perform a sequence of movements, the results seem to indicate that the hierarchical nature of the representation pertains to the memory representation of the sequence and not just the motor plan of the sequence.

GENERAL DISCUSSION

Although the role of structure on spatial working memory has started to be assessed only recently, evidence for its positive effects on recall is mounting

(Bor et al., 2003; De Lillo, 2004; Kemps, 1999; Parmentier et al., 2006; Ridgeway, 2006). One beneficial form of organisation of spatial working memory is provided by segregating sequences on the basis of groupings of items defined by their relative spatial proximity (De Lillo, 2004) and this type of segregation may induce a hierarchical coding of the sequence (see De Lillo, 2004, and Parmentier et al., 2007, for a discussion).

The aims of the present study were to clarify: (1) the extent to which results indicating a benefit of clustering can be generalised to sequences of different length and composition and with additional irrelevant items not pertaining to the sequence to be reported (Experiment 1); (2) whether evidence for HO can be obtained by manipulating the segregation in terms of cluster size and number of clustered sequences (Experiments 2 and 3); and (3) if so, whether the planning of a serial response is required for evidence of hierarchical coding to emerge (Experiment 4).

The results of Experiment 1 confirmed that the beneficial effects of clustering are robust and hold for both sequences of length six and eight. It has been shown that the Corsi test taps executive resources differentially as a function of sequence length (Vandierendonck et al., 2004) and that executive and frontal functions may support the use of organisational strategies (Bor et al., 2003; Stuss et al., 1994). With our manipulation of the length of the sequence we tested the possibility that strategic factors are recruited only for longer sequences where the need to minimise the memory demand of the task is particularly acute. Our finding that recall is more accurate for structured sequences, even in sequences of six ordinal steps, indicates that the processes responsible for taking advantage of serial spatial structure are deployed also for a sequence length around the typical spatial span of about six items observed in the Corsi test (see Kessels et al., 2000; Vandierendonck et al., 2004).

In studies of the effects of the presence of structure in the array of items, irrespective of whether the serial component of the sequence was compatible with the structure of the array or not, it has been observed that additional irrelevant items reduced the benefits of structure (Kemps, 1999). In our case, where the structure of the sequence was defined by its *serial-spatial* structure, this was not the case. Thus, the beneficial effects of serial-spatial structure as assessed in this study are solid and are not reduced by the presence of additional items.

The differential pattern of RT for sequences of 12 items, depending of the composition of the sequence, observed in Experiments 2 and 3 provided a strong indication that participants formed hierarchical representations. With rather different connotations, the notion of HO in cognition has been proposed in the context of list learning (Estes, 1972), spatial learning of static locations of objects (McNamara, 1986), action planning, and motor control (Lashley, 1951; Sternberg et al., 1988, 1990), but it has been rarely assessed in the context of serial spatial working memory. In this study we provide supporting evidence for the fact that HO emerges in this latter task domain. Hierarchical representational structures in working memory could be formed at one or more of the following stages: encoding, rehearsal, or when organising the motor output required for reporting the sequence (analogously to what has been proposed for speech and typewriting; Sternberg et al., 1988, 1990).

The results of Experiment 4 indicate that motor planning may not be necessary for evidence of HO to emerge. This does not exclude the likely possibility that HO pertains also to the motor planning of sequences in the Corsi-type tasks but our results indicate that it must occur at other stages of processing too, such as encoding and/or rehearsal. f-MRI studies indicating a differential activation of brain areas for structured and unstructured Corsi-type sequences during the presentation of the sequence rather than in the retention interval (Bor et al., 2003) would suggest that specific forms of processing for structured material are likely to take place at encoding.

Clusters of icons create visual configurations. It has been shown that Corsi sequences are processed configurationally by encoding the location of each item in the sequence in relation to the other icons in the display rather than on the basis of egocentric coordinates centred on the body of the participant (Avons, 2007; Avons & Trew, 2006). Some facilitating effects for clustered sequences could stem from the particular visual shapes obtained by mentally drawing the connections between items in clustered sequences and using visual memory to help guiding their recall. Should this be the case, memory aid could derive from: (1) recognising familiar shapes, (2) segregating particular static shapes from the background on the basis of Gestalt laws of organisation, or (3) both. These possibilities raise interesting interpretative hypotheses with important implications for working memory models. The possibility that sequences segregated by clusters generate static visual configurations which are easier to code because they match LTM templates for familiar shapes would imply the presence of an interface between working memory systems and LTM. A similar hypothesis has been proposed in verbal working memory to explain the benefits of prose recall and the recall of semantically clustered list over serial recall for unconnected words and has led to the revision of Baddeley and Hitch's (1974) working memory model in order to include an additional component responsible for such interface (see Baddeley, 2000). The hypothesis of the involvement of LTM representations in the processing of spatial structure would be supported by the fact that the structured material presented in matrices and conforming to linear principles in a Corsi-type task also activates the fusiform gyrus which is a brain area often associated with object recognition (Bor et al., 2003). However, it is interesting to note that, in contrast with the verbal domain, where the role of LTM in explaining the difference between semantically organised material and unconnected words cannot be disputed, structured visual patterns could be easier to encode by virtue of their simplicity and redundancy in terms of Gestalt principles and informational content apart from their familiarity or other factors dependent on LTM. The fact that recall of supraspan sequences in Experiment 3 improved with extensive task practice suggests that some learning can take place even when only temporary retention is required by the task. However, since we did not observe an interaction between session and condition in Experiment 3, we conclude that practice effects are likely to be due to the development of general strategies which are independent from the amount of structure or redundancy present in the sequences to be reported. This issue, however, would deserve further study especially considering its vast potential implications for current theories of memory.

Temporal grouping and phrasing improves recall of verbal and auditory material (Mayberry, Parmentier, & Jones, 2002). Imposing pauses in at the specific points of the presentation of the sequence has also been shown to affect recall in structured versions of the Corsi task (Bor et al., 2003). We did not directly manipulate the temporal structure of the sequence during its presentation. Nevertheless, the imposition of temporal structure could in principle have played a role as part of rehearsal processes. Some rehearsal processes in the Corsi test are assumed to be based on mental imagery or be mediated by overt eye movements. Thus, it is possible that when scanning the representation by means of internal processes or eye movements (see. Postle, Idzikowsky, Logie, & Baddeley, 2006, for a recent discussion), pauses would naturally emerge at cluster boundaries given the increased spatial separation of items there. This may have contributed to the emergence of HO. It is, however, unlikely that temporal grouping alone could be responsible for HO and improved recall as also each unstructured sequence has transitions of different lengths which should lead to pauses in the scanning process. As recall accuracy is always lower in unclustered sequences, it seems that temporal phrasing has to be consistent with spatial structure in order to have a beneficial effect on recall.

In studies based on variations of the Corsi test, some the benefits of serial spatial structure have been attributed to variables related to the characteristics of the path through the display of items, such as the path length and the number of path crossing required by different sequences (see Parmentier et al., 2006). To those studies the present findings add the notion that different spatial sequences can be represented in different ways depending on characteristics related to the segmentation of the sequences in a different number of groupings of different size and that the nature of the representation does not pertain exclusively to the organisation of motor plans.

In light of this evidence, current models of the representation of serial order will need to be refined in order to take into account the mounting evidence for a positive role of structuring in spatial working memory (see also Bor, Duncan, & Owen, 2001; Bor et al., 2003; Kemps, 1999). This could be beneficial not only for advances in the theoretical understanding of the factors affecting the representation of serial order but also for the refinement of diagnostic tools for the characterisation of spatial memory deficits that go beyond the measurement of a general memory span in the Corsi test.

Original manuscript received May 2008
Revised manuscript received January 2009
First published online manth/year

REFERENCES

Attneave, F. (1954). Some information aspects of visual perception. *Psychological Review*, *61*, 183–193.

Avons, S. E. (2007). Spatial span under translation: A study of reference frames. *Memory and Cognition*, *35*, 402–417.

Avons, S. E., & Trew, L. (2006). What kind of space is remembered in spatial span? *Cognitive Processing*, *7*(Suppl. 1), S157–S159.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). *London: Academic Press.*

Berch, D. B., Krikorian, R., & Huha, E. M. (1998). The Corsi block-tapping task: Methodological and theoretical considerations. *Brain and Cognition*, *38*, 317–338.

Boduroglu, A., & Shah, P. (2006). Configural representations in spatial working memory. *Cognitive Processing*, *7*(Suppl. 1), S160–S162.

Bor, D., Duncan, J., & Owen, A. M. (2001). The role of spatial configuration in tests of working memory explored with functional neuroimaging. *Scandinavian Journal of Psychology*, *42*, 217–224.

Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, *37*, 361–367.

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International*, *34*(02), 891B (University microfilms No. AA105-77717).

De Lillo, C. (2004). Imposing structure on a Corsi-type task: Evidence for hierarchical organisation based on spatial proximity in serial-spatial memory. *Brain and Cognition*, *55*, 415–426.

De Lillo, C., Aversano, M., Tuci, E., & Visalberghi, E. (1998). Spatial constraints and regulatory functions in monkeys' (Cebus apella) search. *Journal of Comparative Psychology, 1998*, *112*(4), 353–362.

De Lillo, C., Aversano, M., & Visalberghi, E. (1997). The organisation of exhaustive searches in a patchy space by capuchin monkeys (Cebus apella). *Journal of Comparative Psychology*, *111*(1), 82–90.

De Lillo, C., & Lesk, V. (2006, 10–12 April). *Executive functions and hierarchical organisation in the spatial working memory for spatial locations.* Paper presented at the meeting of the Experimental Psychology Society; Birmingham, UK.

De Lillo, C., & Lesk, V. (2009). *Executive functions and hierarchical organisation in the spatial working memory for spatial locations.* Manuscript in preparation.

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–190). Washington, DC: Winston.

Farrand, P., Parmentier, F. B. R., & Jones, D. M. (2001). Temporal-spatial memory: Retrieval of spatial information does not reduce recency. *Acta Psychologica*, *106*, 285–301.

Fischer, M. H. (2001). Probing spatial working memory with the Corsi blocks task. *Brain and Cognition*, *45*(2), 143–154.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Lawrence Erlbaum Associates, Inc.

Henson, R. N. A. (2001). Short-term memory for serial order. *The Psychologist*, *14*, 70–73.

Kemps, E. (1999). Effects of complexity on visual-spatial working memory. *European Journal of Cognitive Psychology*, *11*, 335–356.

Kessels, R. P. C., Zandvoort, M. J. E., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The Corsi-block tapping task: Standardization and normative data. *Applied Neuropsychology*, *7*(4), 252–258.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence*, *73*, 31–68.

Klahr, D., Chase, W. G., & Lovelace, E. A. (1983). Structure and process in alphabetic retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *9*, 462–477.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms and behavior* (pp. 112–136). New York: Wiley.

Mayberry, M. T., Parmentier, F. B. R., & Jones, D. M. (2002). Grouping of list items reflected in the timing of recall: Implications for models of serial verbal memory. *Journal of Memory and Language*, *47*, 360–385.

McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, *18*, 87–121.

Milner, B. (1971). Interhemispheric differences in the localisation of psychological processes in man. *Cortex*, *27*, 272–277.

Parmentier, F. B. R., Andres, P., Elford, G., & Jones, D. M. (2006). Hierarchical organisation in visuo-spatial serial memory: Interaction of temporal order with spatial and temporal grouping. *Psychological Research*, *70*, 200–217.

Postle, B. R., Idzikowsky, C., Della Sala, S., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology*, *59*, 100–120.

Povel, D. J., & Collard, R. (1982). Structural factors in patterned finger tapping. *Acta Psychologica*, *52*, 107–123.

Ridgeway, D. (2006). Strategic grouping in the spatial span memory task. *Memory*, *14*, 990–1000.

Smirni, P., Villardita, C., & Zappala, G. (1983). Influence of different paths on spatial memory performance in the Block-Tapping Test. *Journal of Clinical Neuropsychology*, *5*, 355–359.

Sternberg, S., Knoll, R. L., Monsell, S., & Wright, C. E. (1988). Motor programs and hierarchical organisation in the control of rapid speech. *Phonetica*, *45*, 175–197.

Sternberg, S., Knoll, R. L., & Turock, D. L. (1990). Hierarchical control in the execution of action sequences: Tests of two invariance properties. In M. Jeannerod (Ed.), *Attention and performance XIII: Motor representation and control* (pp. 3–55). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Stuss, D. T., Alexander, M. P., Palumbo, C. L., Buckle, L., Sayer, L., & Pogue, J. (1994). Organisational strategies of patients with unilateral and bilateral frontal lobe injury in word list learning tasks. *Neuropsychology*, *8*, 355–373.

Vandierendonck, A., Kemps, E., Fastane, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology*, *95*, 57–79.