Available online at www.sciencedirect.com
SCIENCE DIRECT®
Acta Psychologica 120 (2005) 113–140
acta psychologica
www.elsevier.com/locate/actpsy

Verbal coding and the storage of form-position associations in visual–spatial short-term memory

Kevin Dent *, Mary M. Smyth

Department of Psychology, Lancaster University, Lancaster, UK

Received 18 October 2004; received in revised form 21 March 2005; accepted 22 March 2005
Available online 17 May 2005

Abstract

Short-term memory for form-position associations was assessed using an object relocation task. Participants attempted to remember the positions of either three or five Japanese Kanji characters, presented on a computer monitor. Following a short blank interval, participants were presented with 2 alternative Kanji, only 1 of which was present in the initial stimulus, and the set of locations occupied in the initial stimulus. They attempted to select the correct item and relocate it back to its original position. The proportion of correct item selections showed effects of both articulatory suppression and memory load. In contrast, the conditional probability of location given a correct item selection showed an effect of load but no effect of suppression. These results are consistent with the proposal that access to visual memory is aided by verbal recoding, but that there is no verbal contribution to memory for the association between form and position.

PsycINFO classification: 2300; 2323; 2340; 2343

Keywords: Short-term memory; Visual memory; Spatial memory; Verbal memory; Object

* Corresponding author. Address: Behavioural Brain Sciences Centre, School of Psychology, University of Birmingham, Edgbaston Birmingham B15 2TT, UK. Tel.: +44 121 4148509; fax: +44 121 4144897.
E-mail address: k.dent@bham.ac.uk (K. Dent).

0001-6918/$ - see front matter
doi:10.1016/j.actpsy.2005.03.004

1. Introduction

The visual world may be described along several distinct dimensions, for example, shape, colour, and location. Neurophysiological, neuropsychological, and behavioural data support the idea that the human visual system uses such a set of independent dimensions to code visual inputs (see Zeki, 1993 for a review). One of the most salient distinctions made both in everyday language and by the visual system is between "what" and "where" (Ungerleider & Miskin, 1982). When one considers the everyday tasks the human visual system must solve, the utility of the distinction between "what" and "where" becomes clear. Object recognition must achieve some degree of location invariance, so that objects may be recognised as the same despite appearing in distinct locations. Other tasks, involving overt movements directed towards objects, require precise location information. Treisman (Treisman, 1988; Treisman & Gelade, 1980) draws attention to the fact that a visual system employing independent descriptive dimensions, must solve a binding problem, so that multiple properties of particular objects are correctly perceived. In this paper, we are concerned with a particular limited form of this binding problem, that of maintaining the association between a complex shape and its location, and the contribution of verbal forms of representation to the solution of this problem.

Much of the research conducted on short-term memory through the 60s and 70s concentrated on memory for lists of linguistic material. Data from behavioural, neuropsychological, and brain imaging studies now supports a broad distinction between verbal and visuo-spatial short-term memory systems (see Baddeley & Logie, 1999 for a recent review). This distinction is embodied in Baddeley's (1986) working memory model which proposes independent verbal (phonological loop) and visuo-spatial (visuo-spatial sketchpad) subsystems. More recent evidence supports the idea that the distinctions between descriptive dimensions made in visual perception may be preserved in short-term memory. Studies requiring individuals to detect a change between two presentations of an image have shown that storing a number of values from a single dimension is more difficult than storing a number of values from distinct dimensions (Olson & Jiang, 2002; Wheeler & Treisman, 2002). Such findings support the idea of independent dimension specific memory stores. Other earlier studies support a broader distinction between memory for "what" and "where" type information. Studies employing interference paradigms have shown a double dissociation between memory for spatial and visual properties (Hecker & Mapperson, 1997; Kessels, Postma, & De Haan, 1999; Klauer & Zhao, 2004; Tresch, Sinnamon, & Seamon, 1993); all showing that certain irrelevant tasks interfere with memory for spatial but not visual material, but other irrelevant tasks interfere with memory for visual but not spatial information.

Another important dimension to consider in the context of visual–spatial short-term memory is that of "when". Many investigations of spatial-short term memory have employed dynamic tasks, in which locations are presented one at a time and recall must be in correct serial order (e.g. Jones, Farrand, Stuart, & Morris, 1995; Logie & Marchetti, 1991; Smyth & Scholey, 1994). Recent research (e.g. Zimmer, Speiser, & Seidler, 2003; see also Pickering, 2001, for a useful discussion of this issue) suggests that such dynamic tasks may recruit additional processes when compared with their static counterparts. Zimmer et al. (2003) compared participants' ability to memorise and subsequently reproduce the spatial arrangement of a set of objects, with their ability to reproduce a sequence of locations in correct serial order. Zimmer et al. (2003) showed that performance in the spatio-temporal task was severely impaired by irrelevant arm movements to spatial targets during a retention interval. In contrast static object relocation showed no effects of such arm movements when pictures of real objects were used and only minimal disruption when complex nonsense objects were used. These results are consistent with the idea that whereas the retention of spatio-temporal sequences may recruit a movement planning (Logie, 1995) or attention (Smyth & Scholey, 1994) based active rehearsal mechanism, the retention of static configurations makes use of such a mechanism to a much smaller extent. In the current paper, we confine our investigation to object-location memory in the context of a static task, without the requirement to retain temporal order information.

The key question addressed in this paper is, given the separation between visual and spatial memories, how does one maintain the association between the visual and spatial properties of objects, and in particular what is the role of verbal-linguistic processes and representations in this domain? Studies which have directly assessed the ability to store object-position associations have shown that such storage is indeed possible but highly capacity limited. Irwin (Irwin, 1992; Irwin & Andrews, 1996) examined performance in a trans-saccadic partial report bar probe task (Averbach & Coriell, 1961). Participants were briefly (<300 ms) shown an array of letters (Irwin, 1992) or coloured letters (Irwin & Andrews, 1996), after making an eye-movement, an empty location was cued, and participants reported the prior contents of this location. Successful performance requires that participants store the association between locations and the objects which occupied them. The results showed that people were able to store approximately 3–4 letter-position (Irwin, 1992) or letter-colour-position (Irwin & Andrews, 1996) units in memory. Given the brief timing parameters used in this study it is unlikely that verbal representations play a role in this task. Irwin (1992) suggests that associative memory in this context is supported by a specific kind of visual object representation; the object file. Object files (Kahneman & Treisman, 1984; Kahneman, Treisman, & Gibbs, 1992) preserve the integrity of multi-property object representations by setting up associations between spatio-temporal properties of objects and their other visual properties, however only a small number of such object files may be maintained at any one time.

Other research supports the idea that the ability to store the associations between objects and their locations is highly capacity limited, and more limited than the ability to store either object or location information per se. Schumann-Hengsteller (1992) examined the ability of children to reconstruct from memory an array of visual objects. She showed that while children were very good at choosing the correct items, and at reconstructing the original locations, they were much poorer at associating objects and positions, often placing a correctly selected object at the location of a different object. Furthermore while the ability to select the correct items and to reconstruct the correct positions showed little development with age, older children were better at correctly pairing objects and locations.

Postma and De Haan (1996) studied an object relocation task in adults. They contrasted participants ability to remember the set of locations with their ability to remember which item occupied which location. In one condition all the items were identical and participants had to retain location information only. In another condition participants were presented with the set of locations used and the set of items used and they had to assign each item to its correct position. Successful performance in the object to position assignment condition requires memory for object-position associations.

Postma and De Haan (1996) showed that performance in these two conditions could be dissociated on two dimensions. The ability to assign objects to positions was substantially impaired by both increasing memory load and by articulatory suppression, in contrast memory for the positions showed only minimal impairment. In agreement with the work of Irwin (Irwin, 1992; Irwin & Andrews, 1996) and Schumann-Hengsteller these results demonstrate capacity limitations on the storage of form position associations in memory. Also in agreement with Schumann-Hengsteller they show that the storage of position information in memory is not tightly limited.

It is widely accepted that articulatory suppression interferes with the short-term retention of verbal information (e.g. Murray, 1968), and furthermore that it hinders the verbal recoding of visual information (Baddeley, Lewis, & Vallar, 1984), instead promoting the use of visual representations (Brandimonte, Hitch, & Bishop, 1992; Hitch, Woodin, & Baker, 1989). Therefore the finding of a detrimental effect of articulatory suppression on object-position assignment in the study of Postma and De Haan (1996), suggests a contribution of verbal coding to this ability. Importantly the effect of articulatory suppression was found, despite the use of nonsense characters which were not readily nameable by the participants.

In principle verbal-linguistic representations could provide a good way to preserve object location associations, in memory. Such representations do not respect the basic distinctions made between descriptive dimensions by visual representations. When recoded verbally spatial and object information is represented by a common linguistic representation. So for example, given a suitable verbal description of each of a set of objects, their positions could be specified relative to each other e.g., "the *X* is above the *Y* but below the *Z*", or with respect to a more global frame of reference, e.g. "the *X* is in the top left of the display". Walker and Cuthbert (1998) showed that at least in the case of shape-colour associations verbal recoding may be of great help. In particular when the shape and colour to be associated do not belong to the same object, and verbal recoding is prevented, shape-colour associations are lost.

Kessels and Postma (2002) further explored the impact of articulatory suppression in the object relocation task, comparing the effects of suppression during the encoding and maintenance phases of the task. They showed that the ability to relocate everyday namable objects back to their original locations was substantially impaired by articulatory suppression at encoding, but that suppression during maintenance only had no effect. In addition, they showed that memory for the positions only was unimpaired by suppression during encoding and maintenance. These findings place important constraints on understanding the role of verbal resources in this task. In particular, it seems that verbal resources play an important role in the derivation of an appropriate representation of either the objects or object-position associations, but that such resources play no subsequent role in maintaining these representations. One specific possibility is that elaborative verbal processing makes explicit certain properties of the stimulus, which are not explicitly coded in visual perception. As Kessels and Postma suggest, it may be that once derived these representations persist passively in long-term memory, rather than relying on active rehearsal in a verbal short-term store.

However one may question whether Postma and De Haan's (1996) articulatory suppression finding really does implicate verbal representations in the storage of object-position associations. Postma and De Haan (1996) assume that their object-position assignment condition targets object-position memory, arguing that since participants are given both all the objects and all the positions they only need to remember the associations. However this is not the case, in order to correctly reassign an object to its prior location, both object and position information must be accessed in memory. A perceived object must access a representation of its prior occurrence in memory, and from this an associated position must be retrieved. It is clear that if I failed to store object or location information in memory my object-position reassignment performance would be impaired. Therefore Postma and De Haan's object-position reassignment condition is not a pure test of the storage of object-position associations in memory. As such the effect of articulatory suppression on object-position relocation may not necessarily implicate verbal representations in the storage of object-position associations. The effect could equally arise from a disruption of memory for objects per se.

Simons (1996) also investigated the contribution of verbal-linguistic representations to memory for different properties of arrays of objects. Participants judged if two arrays separated by a blank interval were the same or different. Changes were of three types: (i) identity change, in which a new object appeared at an old location, (ii) location change, where an old object appeared at a new location, (iii) switch change in which two old objects exchanged locations. In order to detect identity or location changes participants must store identity or location information, respectively. To successfully detect switch changes participants must store identity, location, and their association. Simons manipulated the contribution of verbal coding by (i) comparing memory for real and novel objects, and (ii) examining the effects of verbal shadowing on memory for arrays of everyday objects.

Location changes were easiest to detect and performance here was unaffected by manipulating the availability of verbal coding. In contrast both identity and switch changes were more difficult to detect, and performance with both kinds of changes was impaired by reducing the contribution of verbal coding. However in none of the experiments reported was switch detection significantly poorer than identity change detection, and neither manipulation of verbal coding selectively impaired switch performance whilst leaving identity performance unimpaired. These results are consistent with the idea that the primary contribution of verbal representations in this context is to memory for object identity, but inconsistent with the idea that verbal representations primarily support object-position associations. However, unfortunately there are several problems with the Simons (1996) study. Firstly, performance is at or near chance in the identity and switch conditions, making it difficult to detect a specific contribution of verbal coding to switch performance. Secondly, one may question whether the manipulations of verbal coding employed are specifically verbal in nature. Novel objects might be more visually similar than everyday objects, and verbal shadowing, in comparison with, for example, articulatory suppression, may involve more general attentional and semantic interference, in addition to preventing verbalisation.

The primary aim of this paper is to determine if the suppression effect reported by Postma and De Haan (1996) arises from a disruption of object or object-position knowledge. To this end, we used a modified version of the object relocation task. This modified version includes a separate test of object knowledge allowing us to determine if suppression affects item or item position knowledge.

The paper also addresses the issue of storage capacity for form and form-position information. Firstly is it the case, as suggested by Schumann-Hengsteller (1992), that memory for form information outstrips memory for form-position associations? Secondly is it the case that only a small fixed number of form-position units, may be stored in memory?

1.1. The object relocation task

The task adopted was modeled after the object-position assignment variant of the object relocation task reported by Postma and De Haan (1996). The basic task was to view and memorise the positions of either three or five Japanese Kanji characters, displayed simultaneously each at a unique position on a computer screen. In the task of Postma and De Haan the original set of items were redisplayed along with the original set of locations and the task was to relocate all the items back to their original locations. The strategy of redisplaying the original set of stimulus objects makes it impossible to tease apart item and item-position knowledge, since there was no independent test of item accessibility. Therefore in the current implementation we introduced an independent test of item accessibility. In common with Postma and De Haan (1996) the full set of stimulus locations were marked at test. However in contrast with this earlier study participants relocated only a single item, which they selected from two alternatives, only one of which was present in the original stimulus. The outcome of this decision can be used to determine the extent to which access to representations of the items in and of themselves succeeds or fails. The analytic strategy is then to restrict ones attention to only those trials where item identity has been successfully retrieved. One may then proceed to determine the proportion of these "valid access" trials which indicate the storage of position information as well as form information; the conditional probability of position given identity ($P(pos|id)$), referred to simply as p.

Postma and De Haan's (1996) strategy of requiring relocation of the entire set of stimuli also induces certain mutual dependencies in relocation, which may inflate error. If an item is incorrectly relocated to the location of another item, this item must also be mislocated. In this way an error early in relocation may have knock-on consequences for subsequent relocations. This feature of their task may result in the inflation of the effects of experimental variables like articulatory suppression. Our strategy of requiring relocation of only a single item overcomes this problem, and will allow a better assessment of the impact of suppression in this context.

1.2. A simple "competitive matching" model of performance in the object relocation task

During viewing of the stimulus set it is assumed that a representation of each of the items is registered in memory. Support for this assumption comes from the work of Musen and Treisman (1990) who showed that participants were able to correctly select one of 17 previously viewed novel items from three alternatives 77% of the time, corresponding to a capacity of 12 items, and far exceeding the maximum of five items used here. The stored representation is assumed to be a description of the visual form of a stimulus item. These representations are assumed not to be perfect but to vary in terms of quality or fidelity. When a participant is presented with the selection alternatives a process of competitive matching occurs the outcome of which determines the decision made by the participant. The idea that retrieval from memory takes the form of a competitive matching process is not novel. Such competitive matching models have been proposed for a range of tasks in which a single output must be selected from a range of alternatives (e.g. Humphreys, Riddoch, & Quinlan, 1988, for visual object recognition, Levelt, Roelofs, & Meyer, 1999, for speech production).

For each possible pairing of perceived alternative and memory representation the degree of match is determined and a "match value" is returned for each possibility. The pairing with the highest match value wins the competition and the perceived alternative involved is selected. This competitive matching process is illustrated, for a five item case, in Fig. 1. The match value is driven by two factors, firstly the similarity of the perceived alternative–memory representation pair, the greater the similarity the higher the match value, secondly the fidelity of the memory representation. Reducing the fidelity of the memory representation has two consequences, (i) for pairings involving representations of the same item it will reduce the match value,

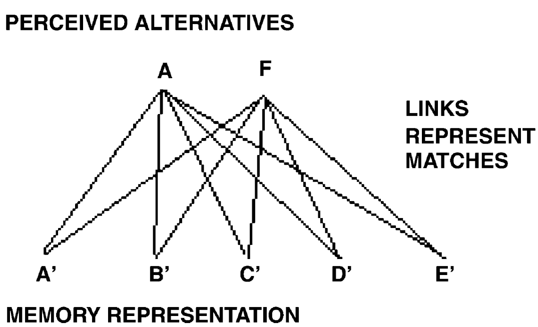

Fig. 1. Illustration of the competitive matching process for a five item case.

in essence making it more difficult to say that two things are the same, (ii) for pairings involving representations of different items it will increase the match value, making it more difficult to say that two things are different. In conditions where fidelity is less than perfect and there is some degree of similarity between items there will be occasions where the perceived target–memory target pair does not win the competition. This kind of event may occur because no pairing reached a sufficiently high value or because a non-target pairing had the highest match value. This kind of event is regarded as a retrieval failure, it does not indicate that a representation of the target was not stored, merely that this representation has not been accessed in this case. Conversely trials where the target–target match wins the competition are regarded as instances of successful retrieval and provide a valid probe into memory.

It is only for these "valid access" trials that it is sensible to ask on what proportion of them position information was available. These valid access trials represent a series of one item samples taken from the set of items stored in memory. This sampling procedure is random since a randomly selected item serves as the target on each occasion, therefore the proportion of these trials which result in a position access should reflect the proportion of item representations in memory which also have a position representation.

1.2.1. Isolating "valid access" trials

Including a foil item in the relocation task allows occasions when the representation of the target has been accessed in memory and occasions when it has not to be distinguished. Therefore it is possible to remove item failures from consideration and restrict ones attention only to trials when access to item memory was successful. This allows one to distinguish memory for items from memory for item-position associations.

The data of an individual participant may be represented in a two by two contingency table the cells of which represent all possible combinations of item correct or incorrect and position correct or incorrect, each cell contains the number of trials which yielded that particular outcome. This situation is illustrated below in Table 1. Cells b and d represent the outcomes of trials in which memory access failed. Cells a and c represent a combination of valid access trials and invalid trials. It should be possible to use the observed values b and d to remove invalid trials from the values a and c.

Table 1
Contingency table for response types in the modified object relocation task

		Item	
		✓	×
Position	✓	a	b
	×	c	d

Where $a + b + c + d$ = total number of trials.
✓ = correct; × = incorrect.

The values b and d are composed of responses resulting from three basic kinds of error. First there is the situation when no match pair yields a sufficient match value to win the competition. Under these circumstances simple guessing occurs, a response alternative is selected at random and located in a random position. On half of the trials the target is selected and on half of the trials the foil is selected. Each of these halves is further divided into $1/n$th correct position and $(n - 1)/n$ incorrect position trials, where n is the memory set size. Since the target and the foil are selected equally often the response matrix is symmetrical and we know that when people guess, for every incorrect guess there is a correct guess which needs to be removed, both for position correct and incorrect trials.

The second source of errors are occasions when either of the two perceived alternatives match a non-target memory representation and this pairing wins the competition, these trials also yield a symmetrical response matrix. In this situation, it is equally probable for either the target or the foil to match a non-target in memory. In addition, both when the target is selected and when the foil is selected the distribution of responses between position correct and incorrect trials is driven by the probability of knowing the position of the non-target and the probability of correctly guessing when it is not known. Thus we have the same distribution for item correct and item incorrect trials. Again we have a symmetrical matrix and we can safely assume that for every item incorrect trial there is a corresponding item correct trial.

The situation is not quite so simple for the third source of errors, instances where the foil matches the representation of the target in memory and wins the competition. This is logically speaking quite an odd situation. What we have is a situation in which the match value obtained for the pairing of the visible target and a representation of the visible target is less than the value obtained for the match between the representation of the target and an unrelated foil. However when this situation occurs the distribution of responses between position correct and incorrect will be driven by the probability that the position of the target is known and the probability of guessing the position correctly. Therefore for errors resulting from these circumstances the ratio of position correct to position incorrect trials will be exactly the same as the ratio resulting from valid access trials, when the target–target match wins. In other words the position correct:position incorrect ratio resulting from this situation should be the same as the uncontaminated valid access ratio we wish to derive.

We can use these facts about the response matrices produced by invalid access trials to remove contamination in the item correct trials. Both when participants guess and when a match involving a non-target memory item occurs the response matrix is symmetrical, for every incorrect item there is a corresponding correct item trial, therefore one may simply subtract these invalid item incorrect trials from the corresponding item correct trials in order to remove them from consideration. However the observed item incorrect trials also receive a contribution from foil-memory target matches and these trials are not symmetrical. Foil-memory target matches have the same ratio of position correct to position incorrect responses as valid trials. Simple subtraction of two identical ratios will yield a result with the same ratio so long as there is a sufficient difference in the size of the numbers making up each ratio.

Therefore if one subtracts the invalid foil-memory target matches from the valid trials the distribution of responses between the two categories position correct and position incorrect will remain unchanged. In order to remove invalid trials from the item correct trials one may simply subtract item incorrect trials from item correct trials. This procedure will have the effect of removing all guesses and non-target memory matches, in addition some valid trials will also be removed but this will leave the ratio of position correct to incorrect trials unchanged.

1.2.2. Correcting for guessing position on valid access trials

The subtraction procedure will leave a set of valid access trials uncontaminated by failures of item access. However on a proportion of trials where position is not known participants will guess correctly, and so the observed proportion of position correct trials will be an overestimate of the conditional probability, position correct given item correct (p).

In order to correct for guessing we simply make the assumption that when participants do not know the position of a memory item, they simply make a random placement to one of the available locations. Thus with n possibilities participants will be correct on $1/n$ trials and incorrect on $(n - 1)/n$ trials. Making this assumption one knows that the observed conditional probability of position incorrect given item correct $(1-p)$ is $(n - 1)/n$th of its true value where n is known, and we can use this fact to derive the true guess free value of p. The resulting equation for converting observed p into true p is given below:

$$\text{true } p = 1 + ((n(p \text{ observed} - 1))/(n - 1))$$

where $p = P(\text{pos}|\text{id})$.

Corrected p is the conditional probability of knowing position given a successful item retrieval and is taken to reflect the proportion of items in memory stored as item-position units.

2. Experiment 1

Experiment 1 assessed the contribution of verbal recoding to a visual object relocation task, by comparing conditions with and without articulatory suppression. The experiment employed Japanese Kanji characters as stimuli. These items were unfamiliar to the participants, they were difficult to code verbally and are unlikely to have any stored representations in long-term memory. Any verbal contribution to performance is therefore non-trivial, and cannot be accounted for by the retrieval of name codes for the items.

Participants viewed three or five characters, which they attempted to memorise. Following a short interval they saw two items, one which had been present just before and one which had not. They attempted to select the correct item and relocate it back to its original position.

In the analysis two distinct measures of performance are assessed: (i) the proportion of trials resulting in a correct item selection, (ii) the probability that the correct position is selected given a correct item selection. If Postma and De Haan are correct and articulatory suppression selectively interferes with item-position associations, leaving both item and position memory unaffected, then articulatory suppression should reduce the probability of a correct relocation given a correct item selection, while leaving unaffected the proportion of correct item selections.

2.1. Method

2.1.1. Participants and design

A total of 32 undergraduate and postgraduate students from Lancaster University, participated in exchange for £3. The participants were divided into two groups, half of the participants took trials with articulatory suppression (eight males and eight females, mean age 21.9, range 19–26, three left handed) and half without (eight males and eight females, mean age 20.3, range 19–26, two left handed). Within each of these groups participants completed two experimental blocks of trials, one block with three characters and one block with five characters. The order of the two blocks was counterbalanced over participants.

2.1.2. Materials and equipment

The experiment was conducted using a Macintosh computer running HyperCard. A 29 cm monitor was used to display all stimuli. The stimuli used were Japanese Kanji characters. A set of 102 Japanese Kanji characters were selected from a larger set of 150 characters. The initial set of 150 characters were the first 150 presented in Henshall (1988), these characters are of the lowest possible complexity and are not composites of other Kanji. The experimenter discarded 48 characters which were either too simple i.e. composed of a single line, or to complex, i.e. composed of too many component parts. The complete set of 102 Kanji used in this study may be seen in Appendix.

Three sets of location arrays were used, each consisting of 48 arrays, one set with each of two, three and five locations. These sets of locations were pre-generated by the experimenter so as to be non-regular and evenly spaced.

2.1.3. Procedure

At the beginning of the experiment participants were seated in front of the computer and instructions were given verbally by the experimenter. Following this participants in the articulatory suppression group were introduced to the secondary task. The articulatory suppression task was simply to repeat the syllable "la" at a rate of three per second throughout the presentation of the stimulus, the retention interval, and the response phase. The required rate was illustrated using a digital metronome, the experimenter then demonstrated the suppression task, participants then articulated in time to the metronome and then without support until they performed satisfactorily. Although performance in the articulatory suppression task was not measured formally, the Experimenter was present throughout to ensure compliance. No participant showed any difficulty in performing the suppression task at the required rate. All participants first completed five practice trials with two characters in order to familiarise them with the task. They then completed two blocks of 20 trials one block with three items and one block with five items. Each block comprised an initial five practice trials followed by 15 experimental trials.

All breaks between the various parts of the experiment were signaled to participants by on-screen text. Participants clicked on an on-screen ''ok'' button to indicate readiness to move on. Progress through the experiment was self paced, in between trials a screen appeared to ask participants if they were ready to continue. After participants clicked the mouse a screen containing only a square outline frame 10.5 cm × 10.5 cm appeared for 1.5 s, at this point participants in the articulatory suppression group began the secondary task. Following this the Japanese characters were displayed at unique positions within the frame. The set of characters used on each trial were selected at random from the complete set of 102 characters. No more than a single instance of a particular character could appear within a single display. In order to avoid repetition of characters between trials, successive displays of the same character were always separated by at least 3 intervening trials. The positions occupied by the characters were selected at random from a set of 48 pre-generated arrays of three or five positions as appropriate, every trial used a unique array. The selected characters were mapped to the selected positions at random. Fig. 2 shows a sample stimulus, at approximately 75% of the original size.

The character stimuli were displayed for a period of 7 s, there then followed a retention interval of 2 s during which participants viewed the outline square frame only. Participants were then presented with the response array, the response array consisted of the positions displayed in the original stimulus clearly marked by outline circles measuring 6 mm in diameter. The set of positions were surrounded by an outline square frame, above this frame two Japanese characters were displayed, one of these characters formed part of the original stimulus and one did not. Fig. 2 shows a sample response array, at approximately 75% of the original size. The task of the participant was to choose the character which had been shown in the original stimulus and relocate it to the position it had occupied, using the mouse. When participants completed this task a button appeared at the bottom of the screen labelled ''done''. At this point participants in the articulatory suppression group stopped repeating the nonsense syllable. Participants clicked this button to move on and when they did so they were again shown the original stimulus for feedback purposes. Feedback was given in order to make the task more engaging and to increase the participants' motivation to perform well. The feedback display was terminated by a mouse click, after which the sequence began afresh.

2.2. Results

2.2.1. Valid memory retrievals

In all conditions participants selected the correct item more often than expected by chance alone (three items without suppression $t(15) = 31.275$, SE = 0.0149, three items with suppression $t(15) = 17.828$, SE = 0.0215, five items without suppression $t(15) = 11.942$, SE = 0.0318, five items with suppression $t(15) = 10.584$, SE =

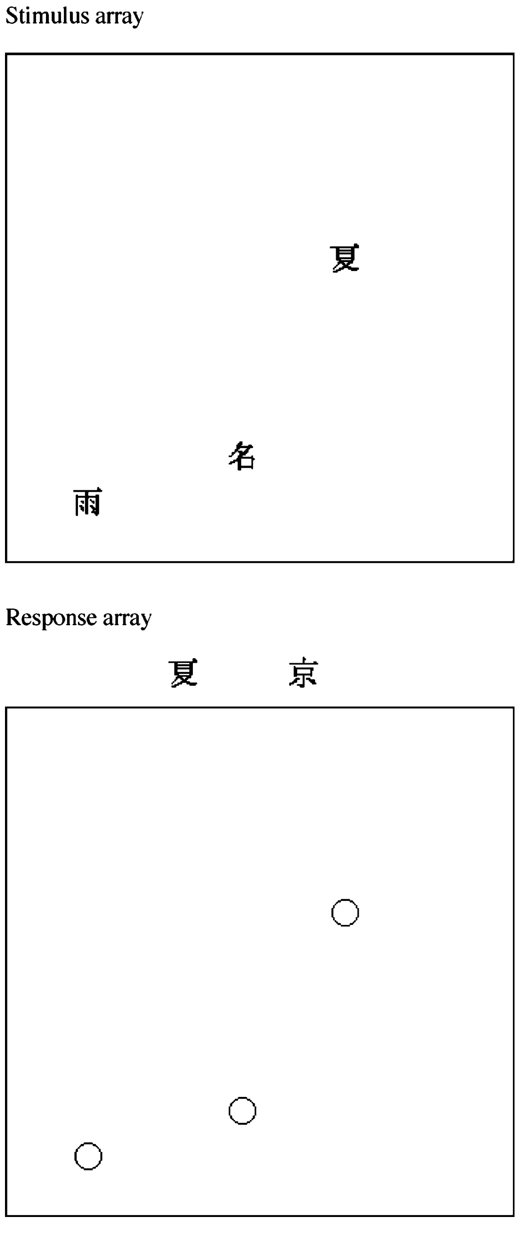

Fig. 2. Sample stimulus and response arrays, three item trial. Shown at approximately 75% of the original size.

0.0582, all *p*s < 0.0001). The proportion of trials in which the target was selected was analysed using a mixed ANOVA with one between groups factor, (articulatory

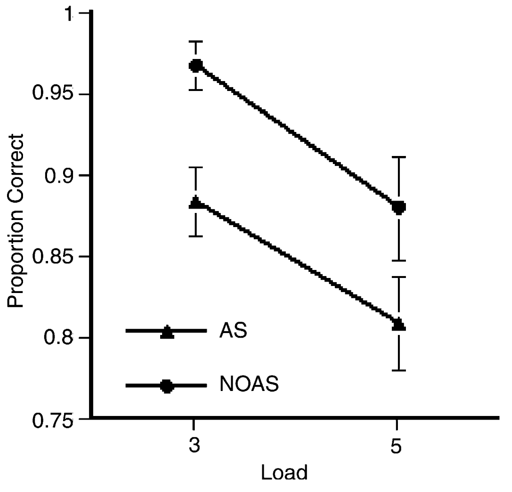

Fig. 3. Valid access trials: Mean proportion item correct as a function of memory load (three or five items), with articulatory suppression (AS) and without articulatory suppression. Means for the groups with and without articulatory suppression plotted with separate lines. Error bars show standard errors.

suppression with two levels suppression or no suppression) and one within subjects factor (memory load with two levels, three or five characters). The means may be seen in Fig. 3. There was a significant main effect of suppression $F(1,30) = 6.713$, MSE = 0.01416, $p < 0.05$, performance was significantly worse with than without suppression. There was also a significant main effect of memory load $F(1,30) = 17.14$, MSE = 0.00616, $p < 0.0005$, participants performed significantly better with three characters than with five. The interaction was not significant, $F < 1$.

2.2.2. Conditional probability of position given identity (p)

The frequency distributions for three and five item trials may be seen in Fig. 4. The distribution for three item trials showed a clear ceiling in the data, the majority of the participants had a value of p around 1 with 72% of participants having a value of p between 0.8 and 1. The distribution for five item trials was quite different, there was no pronounced ceiling in the data, the modal category for p was 0.5–0.6, seven participants had values in this range. Mean p was significantly below 1 for the five item trials, $t(31) = 10.107$, $p < 0.0001$. In all conditions values of p were significantly higher than expected on the basis of random relocation (three items without suppression $t(15) = 38.122$, SE = 0.0237, three items with suppression $t(15) = 18.223$, SE = 0.0458, five items without suppression $t(15) = 12.893$, SE = 0.0414, five items with suppression $t(15) = 5.345$, SE = 0.0898, all ps < 0.0001).

Values of p were analysed using a two factor ANOVA, memory load was a repeated measures factor with two levels (three or five items), and articulatory suppression was a between groups factor with two levels (suppression or no suppression). The means can be seen in Fig. 5. There was a significant effect of memory load,

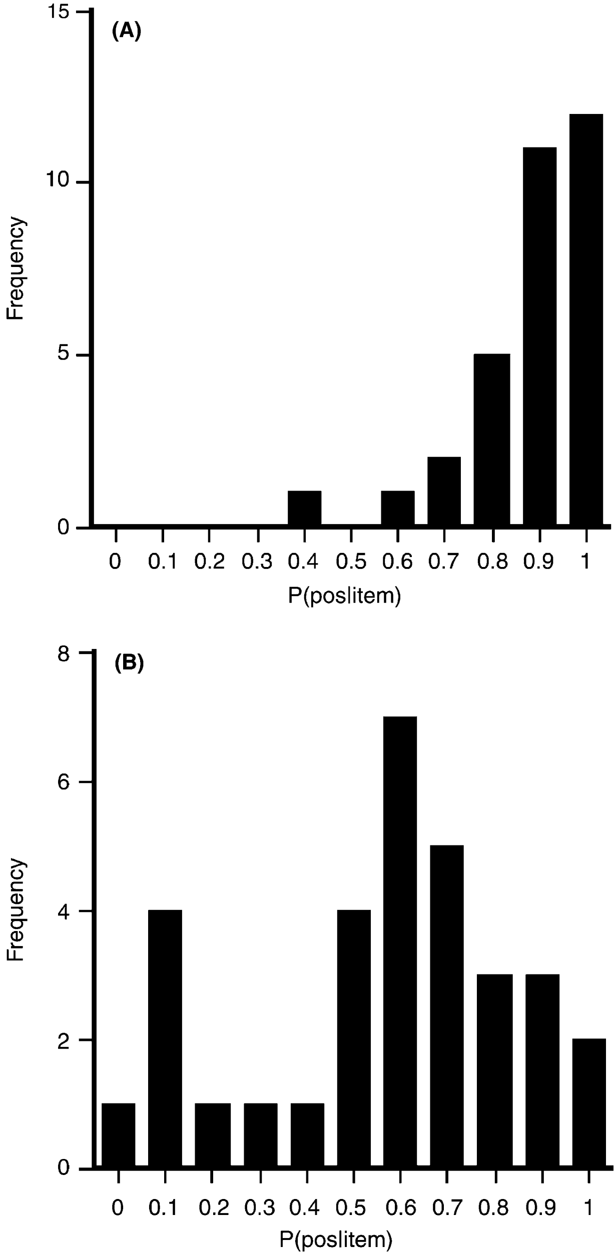

Fig. 4. Panel A, frequency histogram of P(pos|item) for three item trials. Panel B, frequency histogram of P(pos|item) for five item trials. Each bar represents the number of participants with a value of P(pos|item) equal to or less than the value shown below that bar but greater than the next value to the left.

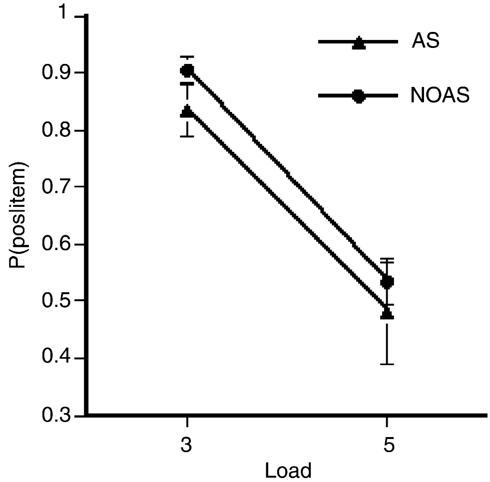

Fig. 5. Mean values of P(pos|item) as a function of memory load (three or five items), with articulatory suppression (AS) and without articulatory suppression. Means for the groups with and without articulatory suppression plotted with separate lines. Error bars show standard errors.

$F(1,30) = 63.493$, $MSE = 0.331$, $p < 0.0001$, the conditional probability was lower with five items than with three items. There was no significant main effect of suppression and no significant interaction between memory load and suppression (both Fs < 1). In order to rule out the possibility that the absence of an effect of suppression was due to performance being at ceiling with three items, we conducted a separate ANOVA on just the five item trials, here too we found no significant effect of suppression, $F < 1$.

2.3. Discussion

2.3.1. Valid memory retrievals

The results of the above analysis clearly show that carrying out articulatory suppression during this task reduces the availability of item information. When participants carry out suppression they are less able to select the correct item. Within the competitive matching framework presented above this means that with suppression the target item is less likely to win the competition. When interpreting findings from dual task situations one must inevitably face the issue of whether any interference observed is general or specific. However a range of empirical data now exists from dual task studies and there seems to be general agreement that articulatory suppression imposes minimal general demands and can be considered largely verbal specific (see Baddeley, 1986; Logie, 1995).

Assuming verbal coding may contribute to performance in this task, can we answer the question of how it contributes? One simple common sense suggestion might be that when free to verbalise participants are able to generate arbitrary names for the items which are then stored in verbal memory. This possibility may be ruled out on purely logical grounds. An essential feature of a successful verbal strategy must be that the verbal memory item can be accessed directly from perceiving the response alternatives bypassing visual short-term storage. In the case of a completely novel item–arbitrary name pairing this is not possible. In order for the name to be accessed at all the correct item must be accessed from visual memory first, then any associated name may be retrieved. If participants adopted the arbitrary naming strategy the competitive matching process would have to proceed in exactly the same way as if no verbal coding occurred, any verbal information will be retrieved only after this process and will therefore be redundant. Given this situation arbitrary naming cannot in principle improve performance.

If verbal coding is to be of any help to participants in this task then the relevant verbal information must be available before the completion of the competitive matching process. The verbal information must be accessible from a perceptual representation of an item before access to VSSTM, for this to be possible the verbal code must bear a non-arbitrary relationship to its referent character. There are two ways to satisfy these requirements. Firstly one could take advantage of existing recognition processes to determine the similarity between characters and common everyday items and then code a character as "like an *X*" where *X* may be any everyday item, a tree, a person etc. Given the visual properties of the character this description will be available following a perceptual analysis and before VSSTM access. The second possibility does not involve accessing knowledge of common items, participants take advantage of the possibility of verbalisation by building propositional descriptions of visual structure, for example "long vertical, short horizontals", these descriptions are then stored in verbal memory. Since there is clearly a non-arbitrary relationship between the structural description and the visual form of the character, these descriptions are accessible given the perceptual availability of an item without VSSTM access.

Both of these forms of verbal coding make extra information available about the stimuli in this task; this additional information may be used to more effectively discriminate between the target and the foil thereby increasing correct item selections. However it is important to note that the forms of verbal coding used in this task do not amount to simple naming but are best described as forms of elaborative verbal coding.

Turning to the memory load manipulation, the results showed that participants were less likely to validly access the memory store with more items. At first glance this finding appears to show the existence of a capacity limitation for storing item form information. However it is possible to distinguish two plausible explanations for this finding each of which has quite different implications for the issue of capacity limitations. Firstly it could be that memory for visual form is subject to a capacity limitation such that as one increases the number of items in store the quality or fidelity of each representation is reduced. This would mean that in the matching competition lower values would arise for target–target pairs and higher values would arise for other pairs involving two quite different items. In other words with reduced fidelity in memory it is more difficult to tell that two identical items are the same and that two different items are different. This has the consequence of changing the relative match values for correct and incorrect pairs, increasing the likelihood of selecting the foil over the target. However a much simpler explanation which assumes no capacity limit cannot be ruled out. It could simply be that as one increases memory load one also increases the probability that the foil is visually similar to a stimulus item. This has no implication for capacity limits but predicts that as memory load is increased, the discrimination of the target from the foil becomes more difficult. Even though at first glance the finding of a memory load effect seems to imply a capacity limitation, it is impossible to rule out the alternative explanation that the difference is a consequence of the structure of the task and not the nature of VSSTM.

In summary, memory load is likely to influence performance by changing the match value of the target-target pairing relative to other competing pairings. Articulatory suppression may act to remove or reduce the availability of additional propositional constraints, which ordinarily aid in disambiguating target from foil. The idea that the effects of suppression and load stem from independent sources is supported by the finding that these two factors do not interact statistically. Following additive factors logic (Sternberg, 1969; see also Schweickert, 1985, for an extension to accuracy data) two factors which have additive effects are likely to operate on distinct components of processing.[1]

2.3.2. *Conditional probability of position given identity (p)*

Interestingly and in contrast to the analysis of item access, the analysis of p did not reveal an effect of suppression. Given that item information has been successfully retrieved participants are no more likely to know that item's position when they are free to verbalise compared to when they engage in articulatory suppression. It is clear that contrary to predictions derived from Postma and De Haan (1996) the action of suppression is not specific to the association between item and position knowledge. It is not the case that articulatory suppression disrupts memory for the association between item and position information, while leaving memory for item and position itself unimpaired.

On the contrary item-position memory was impaired by increasing memory load. The probability that a correct item selection led to a correct placement, after correcting for guessing, was greater for three compared to five items. This memory load effect cannot be the result of an underlying limitation on memory for position, since Postma and De Haan (1996) convincingly showed that large numbers of locations (up to 10) can be retained with remarkable degrees of accuracy in this context. Nor can the load effect stem from limitations on item memory, since performance is conditionalised on correct item selection. Therefore in agreement with Postma and De Haan (1996), it seems that memory for the association between form and position is rather tightly capacity limited. Memory capacity for form information

[1] Thanks are due to Albert Postma for pointing out the application of additive factors logic in this context.

exceeds capacity for the association between form and position, as shown by: (i) extremely high levels of performance for item selection, and (ii) values of p substantially and significantly below 1 with five items. These results are consistent with the storage of a small fixed number of form-position units together with a larger number of form representations. We defer further discussion of the exact nature and mechanism of this capacity limitation until the general discussion.

It is important to be absolutely clear about what can and what cannot be concluded on the basis of the presented data. The experiment set out to test whether suppression and load have an independent effect on the association between form and position over and above any effect on access to form memory. However one weakness with the experiment is that it cannot, and is not designed to, detect any influence of the experimental variables on form-position associations that is not independent of any effect on access to form memory. If it were the case that suppression had a very general effect on processing the stimuli, or if suppression tended to cause the loss of both form and form-position information from memory, p would show no effect of suppression. Therefore we cannot rule out the possibility that articulatory suppression has some effect on form-position associative memory that is not independent of the effect on form memory.

Nonetheless the results still provide major constraints on any theory of VSSTM in the context of object relocation. Firstly any impact of suppression in this task is at best minimal. Access to form memory is remarkably resistant to the effects of suppression showing only about an 8% decrement in overall. Any additional effect of suppression on form-location associations can be no larger than this. Such a small effect of suppression contrasts sharply with the magnitude of the effect reported by Postma and De Haan (1996), which was 20%. What the results do clearly demonstrate is that item-position associations can be disrupted independently of item representations, since memory load selectively disrupts these associations, as shown by smaller values of p with five compared to three items. Despite the fact that it is clearly in principle possible to selectively disrupt item-position associations, it is equally clear that suppression does not cause such selective disruption. It is therefore safe to conclude that articulatory suppression does not have an independent effect on form-position memory after successful retrieval of form.

In summary, while Postma and De Haan (1996) are correct in their suggestion that memory for item-position associations is tightly capacity limited, it is not the case that articulatory suppression causes a similar selective disruption of these associations. Rather the data are consistent with the idea that elaborative verbal coding makes a small but significant contribution to item memory. Articulatory suppression may thus cause substantial difficulties in a task where item access is a necessary prerequisite for relocation, and there exist mutual contingencies in item placement.

3. Experiment 2: Similarity ratings

This sub-experiment attempted to determine the role of visual similarity in Experiment 1. In Section 1 a model was presented for understanding the relocation task used in Experiment 1. This model shows how information is selected from visual memory. Visual codes representing the appearance of the memory items are stored in VSSTM. When the participant is presented with the two response alternatives a competitive matching process ensues, for each logically possible pairing of visible and memory item a match value is computed, the magnitude of the match value is determined by two primary factors, (i) the degree of articulation or fidelity of the memory representation, (ii) the visual similarity of a potential pairing. There are two main sources of error responses according to the model, (i) no match pair yields a sufficiently high match value and both item and position are guessed, (ii) the target–memory target pair is out-competed by an incorrect pair. On those occasions where the target memory target pair is outmatched by a competing pair there should be a relation of similarity between the items involved in the winning pair. Furthermore the similarity of the winning pair should be higher than that of the other competing pairings which did not win the competition. This feature of the model provides the motivation for making the assumption that a selection failure does not imply a storage failure.

The prediction of the model, that winning pairs should be more similar than non-winning pairs may be tested empirically. In Experiment 2, we took all the trials from Experiment 1 which resulted in an error, that is all trials except correct item correct position trials. The stimuli involved in these trials were formed into match pairs, excluding target–target identical pairs. A group of participants then provided similarity ratings for these match pairs. The winning match pair was assumed to be the pair consisting of the item the participant selected and the item occupying the position to which the selected item was relocated. The average ratings for these winning or actual pairs were then compared with the average ratings for the other non-winning but possible matchings. One should find higher ratings for the actual compared to the possible pairs. If one can show that this is the case, support will be provided for the competitive matching framework and the assumptions it embodies. In addition a visual similarity effect of this sort will clearly implicate the involvement of a specifically visual memory in this task and argue against complete mediation by verbal memory.

3.1. Method

3.1.1. Participants

Twelve Postgraduate students from Lancaster University volunteered to take part in this study. Six were male and six female. Mean age 27.4, range 23–38.

3.1.2. Materials

The primary materials for this study were provided by the set of 102 Japanese Kanji used previously. The characters involved in all the error trials of the articulatory suppression condition of Experiment 1 were used. Pairs of items were constructed from the character chosen by the participant in a trial and all other characters present in that trial, with the exception of the identical target–target pair.

Table 2
Composition of the pools of pairs used in Experiment 1 (B)

	3			5			Total pairs
Item	×	×	✓	×	×	✓	
Position	×	✓	×	×	✓	×	
Set 1	5	2	9	8	3	24	190
Set 2	5	3	10	8	3	23	191
Set 3	4	2	9	9	3	24	192
Set 4	5	2	9	8	4	24	195

The table shows the number of each type of error trial in each pool. Total number of pairs in each pool shown in the rightmost column.
✓ = correct; × = incorrect.

Thus for correct item incorrect position trials there were two or four pairings for three or five item trials, respectively, for trials in which the incorrect item was selected there were three or five pairings for three or five item trials, respectively. The total set of 206 trials were divided up into four pools with each pool containing an approximately equal number of different error trials. Table 2 shows the composition of each of the pools, along with the resulting total number of pairs in each pool. Using this method all pairs belonging to the same trial are contained in the same pool, but each pool is composed of a random set of trials from different participants.

The pairs of items were printed onto A4 sheets of white paper, each sheet contained up to a maximum of 48 pairs arranged in a regular eight × six matrix, with each character measuring approximately 0.8 × 0.8 cm. Items within a pair were separated horizontally by 1 cm and pairs were separated from other pairs horizontally by 2 cm, each row of the matrix was separated vertically by 1.5 cm. Each pool of pairs were printed in a random order without respect to the particular trials they came from. Each pool consisted of five sheets of paper containing the set of pairs.

3.1.3. Design

Each pool of character pairs were rated by a separate group of three participants. A single independent variable was manipulated within subjects, that of pair type. Some pairs in the pools were designated "actual" pairs, these were the pairings of the selected item and the item which had occupied the position to which the selected item was relocated. The remainder of the pairs were designated "possible" pairs, these were pairings of the selected item and items which had occupied positions to which the selected item was not relocated.

3.1.4. Procedure

Participants were given the set of sheets for the appropriate pool and the task was explained verbally by the experimenter, full instructions were also printed on a separate sheet and attached to the ratings sheets. The task of the participants was to work through each pair of items and to rate how visually similar each pair looked. Ratings were made using a 5-point scale where a rating of 1 indicated the items were not at all similar and a rating of 5 indicated that the items were extremely similar. Participants worked through each pair of items and recorded their rating on the sheet at the side of each pair. Participants completed the task in their own time and then returned the completed sheets to the experimenter.

3.2. Results

For each participant the mean value of the similarity ratings was calculated for the actual and for the possible pairs. These means were analysed in a single factor repeated measures ANOVA with pair type as the experimental factor. The "actual" pairs ($M = 2.605$, SE = 0.097) were rated as being more similar than the "possible" pairs ($M = 2.933$, SE = 0.089), $F(1, 11) = 52.34$, MSE = 0.0123, $p < 0.00005$.

3.3. Discussion

The results of the above analysis show that visual similarity plays a role in the relocation task. When participants make an error the similarity of the selected item and the item which occupied the position to which the selected item is relocated is higher than that of other possible pairings. This outcome is exactly as predicted by the competitive matching model of performance in which item selection is controlled by a competitive matching process based on visual similarity. According to this model members of a selected pairing should be more similar on average than members of non-selected but possible pairings; this is exactly what we found. This finding therefore validates the overall conceptual framework presented here. In particular the finding provides support for the assumption that an incorrect response does not always imply a failure of storage, and the resulting absence of a representation, an incorrect response may result from a failure of access due to out competition.

4. General discussion

The purpose of the study presented here was to investigate memory for the association between visual form and position. In particular we wished to evaluate the contribution of verbal processes and representations to this ability. An earlier study by Postma and De Haan (1996) appeared to show that verbal recoding does make a contribution to form-position associative memory. Postma and De Haan (1996) claimed that memory for positions and memory for item-position associations could be distinguished in two respects. Firstly with respect to capacity limitations, with item-position but not position memory showing tight limitations. Secondly with respect to interference from articulatory suppression, item-position associative memory being affected much more than position memory.

However Postma and De Haan (1996) failed to effectively distinguish memory for items from memory for item-position associations. Their test of item-position memory was not a pure test, participants would fail this task if they failed to store item information. In the current experiment an independent test of item memory was introduced, which allowed item memory to be distinguished from item-position memory. The probability that a participant selected the correct item was taken as an index of access to item memory. The conditional probability that a participant placed the correct item in the correct position was taken as an index of item-position memory. While both dependent measures showed an effect of memory load, only item access and not item-position memory was sensitive to articulatory suppression. The results show that whereas access to visual form memory is aided by elaborative verbal recoding, such recoding makes no additional independent contribution to memory for form-position associations. Therefore sensitivity to disruption by articulatory suppression may not be a dimension which differentiates item-position memory from other forms of memory.

The results of Experiment 2 further show what can be achieved from a detailed consideration of the errors in the positional reconstruction task. This is the first time, to the best of the author's knowledge that such an analysis has been attempted. When participants make an error in this task they do not just make any old error. When they make an incorrect item-position assignment they tend to do so on the basis of visual similarity. This finding validates the competitive matching framework presented earlier and demonstrates that incorrect responses may result from reasons other than simple representational loss.

Postma and De Haan (1996) suggested that memory for form-position associations is mediated by a system of categorical spatial relations, which specify the positions of items in an explicitly relational fashion. According to Postma and De Haan an item would be coded as "above" and "to the left of" a second item, and so on for a whole set. It is important to note that Postma and De Haan (1996) do not explicitly draw out the link between such a categorical network of relations and explicit verbal-linguistic recoding. However it would seem that an explicit verbal linguistic specification would provide a natural way to capture such a set of categorical spatial relations. It would seem that such categorical relations might provide important constraints on the remembered positions of items, which may themselves be coded using visual representations. The current results show that if such relations do constrain item-position memory in this way, then their representational format must be non-verbal. However given the natural coupling between such a propositional specification of the locations of a set of items and explicit verbal processes, one may begin to question whether such a categorical constraints do contribute to item-position memory.

Although the current paper has shown that verbal coding does not make an independent contribution to form-position memory over and above the contribution made to form memory, this does not rule out the possibility that verbal coding does make such a contribution under other circumstances. Firstly our results indicated that memory for form-position associations was tightly capacity limited with the majority of participants managing to hold only three associations. Verbal recoding may be a useful way to by-pass some of the limitations imposed by visual representations. When presented with large sets of items, such that the capacity of a visual representation system is well over stretched and reliance on such a system would result in the storage of an unacceptably low proportion of form-location units, resort may be made to verbal recoding. Larger arrays may also encourage verbal recoding since they present participants with a greater number of inter-object categorical relations, which may be well captured by verbal-propositional descriptions such as "*X* is above *Y*, but to the left of *Z*". Whether verbal representations are more important in the object relocation task under conditions of high memory load must await future research.

We were able to show that storing just five form-location associations is beyond the capacity of most individuals. Yet participants were much better at accessing form information per se. The majority of errors in the task consisted of occasions where a correctly selected item was placed in an incorrect location. Most individuals succeeded in relocating a correctly selected item about 60% of the time when presented with five items. Assuming a representation of the visual form of each of these five items is stored in memory, this corresponds to a capacity of about three item-position associations. This is consistent both with the finding from the current study showing that most people performed extremely well with three items and previous estimates of the capacity of visual object memory.

Irwin (Irwin, 1992; Irwin & Andrews, 1996) also showed that people were able to store approximately three item-position units in memory. He recruited the notion of the object file (Kahneman & Treisman, 1984; Kahneman et al., 1992) in order to explain this finding. According to Irwin visual memory may contain a small number of object file representations in which all the properties of the remembered objects are integrated. The current results may also be captured by an object file hypothesis, with only around three object-position associations being supported by three object files. However this cannot be the whole story.

Object files in the sense used by Irwin cannot be the only form of representation contributing to memory in this task. If this were the case then there should be no form information without associated position information. A closer inspection of the notion of the object file as presented by Khaneman (Kahneman & Treisman, 1984; Kahneman et al., 1992) shows that in the object file not all properties are equal, and helps us to understand how form and position information may be dissociated in memory.

The object file as discussed by Khaneman has a two component structure. On the one hand there is the content of the file which specifies the multiple properties of an object. On the other hand there is a spatial tag which specifies the place of occurrence of the object. The spatial tag is crucial since it allows an object to be treated as a particular individual conferring a "numerical identity" on the object. The spatial tag allows the object to be treated as the same persisting individual overtime despite changes. New properties which appear at the same place or at places consistent with a spatial transformation of an existing object, are added to the same file. If the contents of an object file were to be dissociated from the spatial tag, then form information would be stored without position information.

We wish to propose that this is indeed what happens in the current task. A small number of object files with tags may be stored along with a larger number of files without tags. According to this framework what is limited in VSSTM is not the ability to store object representations or the ability to code the overall spatial configuration of a set of objects, but the availability of discrete spatial tags. These discrete spatial tags allow non-spatial object properties to be associated with a particular independent spatial location.

Support for the idea that spatial tagging of independent locations is tightly capacity limited comes from the work of Pylyshyn and colleagues (see Pylyshyn, 2001 for a review). Pylyshyn claims that the visual system only has a small number of visual indexes at its disposal. These indexes are assigned to locations and allow these locations to be treated as individuals, permitting visual properties to be correctly assigned to locations. We suggest a similar process gives rise to limitations in associating form and position in VSSTM. This spatial tagging hypothesis predicts that, in contrast to the effects of articulatory suppression, any secondary task which also demands the use of independent spatial tags will selectively impair memory for object-position associations, whilst sparing memory for objects and positions per se. It would be interesting in this regard to explore the effects of interposed visual enumeration (e.g. Trick & Pylyshyn, 1994) and multiple object tracking (e.g. Pylyshyn & Storm, 1988) tasks, in the context of object relocation, since according to Pylyshyn (2001) these tasks also require the deployment of discrete spatial indexes or tags.

5. Conclusions

Contrary to previous reports, articulatory suppression does not selectively impair memory for form-position associations. These associations are likely to be held using a non-verbal visual representation. The ability to store such associations in memory is tightly limited, whereas the ability to store form information per se is not. These results are consistent with the hypothesis that limitations on the storage of form position associations result from the limited availability of spatial tags which serve to locate object representations in space.

Acknowledgments

The experiments reported here also form part of a doctoral dissertation completed by the first author at Lancaster University under the supervision of Mary M. Smyth. The research was supported by an ESRC studentship awarded to the first author, and the writing of this paper was supported by a grant from the Leverhulme Trust. Thanks are due to Glyn Humphreys, Albert Postma and two anonymous reviewers for helpful comments on an earlier draft.

Appendix

The complete set of character stimuli.

男	古	山	女	紙	黄	木	円	夏	画
春	牛	遠	森	時	虫	才	手	右	水
入	生	社	早	弱	町	子	車	寺	室
出	空	汽	百	書	青	夕	四	行	力
左	村	近	記	先	強	九	気	外	
帰	見	名	原	戸	力	月	今	土	
語	火	黒	考	作	正	雲	年	五	
元	赤	家	引	思	糸	校	雨	何	
計	市	国	科	光	石	会	色	心	
場	貝	八	京	海	休	林	学	少	
止	花	足	形	交	間	字	午	中	

References

Averbach, E., & Coriell, E. (1961). Short-term memory in vision. *Bell System Technical Journal, 40*, 309–328.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D., Lewis, V. J., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology, 36*, 233–252.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.

Brandimonte, M., Hitch, G. J., & Bishop, D. (1992). Influence of short-term memory codes on visual image processing: Evidence from image transformation tasks. *Journal of Experimental Psychology: Learning Memory and Cognition, 18*, 157–165.

Hecker, B., & Mapperson, B. (1997). Dissociation of visual and spatial processing in working memory. *Neuropsychologia, 35*, 599–603.

Henshall, K. G. (1988). *A guide to remembering Japanese characters*. Rutland Vt: Tuttle.

Hitch, G. J., Woodin, M. E., & Baker, S. (1989). Visual and phonological components of working memory in children. *Memory and Cognition, 17*, 175–185.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5*, 67–103.

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning Memory and Cognition, 18*(2), 307–317.

Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inu & J. L. McClelland (Eds.). *Attention and performance* (XVI). Cambridge, MA: MIT Press.

Jones, D., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*, 1008–1018.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29–61). New York: Academic Press.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175–219.

Kessels, R. P. C., & Postma, A. (2002). Verbal interference during encoding and maintenance of spatial information in working memory. *Current Psychology Letters: Behaviour, Brain and Cognition, 9*, 39–46.

Kessels, R. P. C., Postma, A., & De Haan, E. H. F. (1999). P and M channel specific interference in the what and where pathway. *Neuroreport, 10*, 3765–3767.

Klauer, K. C., & Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology: General, 133*(3), 355–381.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access. *Behavioral and Brain Sciences, 22*, 1+.

Logie, R. H. (1995). *Visuo-spatial working memory*. Hove: LEA.

Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive. In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 105–115). Amsterdam: North Holland Press.

Murray, D. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679–684.

Musen, G., & Treisman, A. (1990). Implicit and explicit memory for visual patterns. *Journal of Experimental Psychology: Learning Memory and Cognition, 16*, 127–137.

Olson, I. R., & Jiang, Y. (2002). Is short-term memory object based? A rejection of the strong object hypothesis. *Perception and Psychophysics, 64*, 1055–1067.

Pickering, S. (2001). Cognitive approaches to the fractionation of visuo-spatial working memory. *Cortex, 37*, 457–473.

Postma, A., & De Haan, E. H. F. (1996). What was where: Memory for object locations. *Quarterly Journal of Experimental Psychology, 49*(A), 178–199.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects and situated vision. *Cognition, 80*(1–2), 127–158.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*, 179–197.

Schumann-Hengsteller, R. (1992). The development of visuo-spatial memory: How to remember location. *International Journal of Behavioural Development, 15*(4), 455–471.

Schweickert, R. (1985). Separable effects of factors on speed and accuracy: Memory scanning, lexical decision, and choice tasks. *Psychological Bulletin, 97*(3), 530–546.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science, 7*, 301–305.

Smyth, M. M., & Scholey, K. A. (1994). Interference in spatial immediate memory. *Memory and Cognition, 22*, 1–13.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315.
Treisman, A. M. (1988). Features and objects: The fourteenth annual Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology, 40A*, 201–237.
Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97–136.
Tresch, M. C., Sinnamon, H. M., & Seamon, J. G. (1993). Double dissociation of spatial and object visual memory: Evidence from selective interference in intact human subjects. *Neuropsychologia, 31*, 211–219.
Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited capacity preattentive stage in vision. *Psychological Review, 10*, 1–23.
Ungerleider, L. G., & Miskin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. Mansfield (Eds.), *Analysis of visual behaviour*. Cambridge, MA: MIT Press.
Walker, P., & Cuthbert, L. (1998). Remembering visual feature conjunctions: Visual memory for shape colour associations is object based. *Visual Cognition, 5*(4), 409–455.
Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131*(1), 48–64.
Zeki, S. (1993). *A vision of the brain*. Oxford: Blackwell.
Zimmer, H. D., Speiser, H. R., & Seidler, B. (2003). Spatio-temporal working memory and short-term object location tasks use different memory mechanisms. *Acta Psychologica, 114*, 41–65.