Downloaded from rstb.royalsocietypublishing.org

Short-Term Visual Memory [and Discussion]

W. A. Phillips and D. A. Routh

Phil. Trans. R. Soc. Lond. B 1983 **302**, 295-309
doi: 10.1098/rstb.1983.0056

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. Lond. B* go to: **http://rstb.royalsocietypublishing.org/subscriptions**

This journal is © 1983 The Royal Society

Downloaded from rstb.royalsocietypublishing.org

Phil. Trans. R. Soc. Lond. B **302**, 295–309 (1983)
Printed in Great Britain

Short-term visual memory

By W. A. Phillips

Department of Psychology, University of Stirling, Stirling FK9 4LA, U.K.

Serial position effects in visual memory are presented as evidence for a short-term memory for visuo-spatial information that is not plausibly explained in terms of either verbal or sensory representations. This is called visualization, and is distinguished from long-term visual memory. Cases of head injury are reported in which long-term memory is affected but not visualization. In contrast with this, mental arithmetic interferes with visualization but not with long-term memory. Further studies are reported that throw doubt on the earlier explanation of this interference in terms of competition for a central executive or strategic coordinator.

The significance of these findings is discussed in relation to a proposal for classifying the main kinds of information represented in higher visuo-spatial cognition.

Classes of representation in visuo-spatial cognition

The specific topic that I am going to discuss is whether there is any clear distinction beween short-term memory and long-term memory for visuo-spatial representations. My general concern is with functional specialization in visual cognition, and it seems to me that this specialization is primarily semantic. That is, the systems involved are primarily designed to represent different kinds of information about the world. I must therefore begin by saying something about what it is that visuo-spatial representations represent.

The basic distinction I wish to draw is that between the representation of local properties of currently visible surfaces, and the representation of objects and scenes. The various retino-topic maps that occupy most of the occipital lobe, and to which the various kinds of sensory storage relate, represent various aspects of stimulation. They do not represent objects, and they do not provide an adequate basis for affective evaluation, verbal description, or motor control. These all require further descriptions to be computed by relating retinal stimulation to input from other modalities, or to stored schemata of the appearance of objects and scenes. These further visuo-spatial representations are referred to collectively as scene descriptions. There is a very great deal of evidence for the importance of visuo-spatial representations in human cognition (see, for example, Bower 1970; Chase 1973; McGee 1979). Unfortunately the associated debate on the nature of imagery does not seem to me to have led to any clear understanding of this large and diverse field (see, for example, Block 1981). Part of the difficulty might be that while attention has centred on the problem of *how* images represent, too little attention has been paid to *what* they represent.

Four broad classes of information that might be represented in scene descriptions are shown in figure 1. This classification is based on three distinctions. First, there is the distinction between objects and positions. Second, there is the distinction between object class and object structure. Third, there is the distinction between the position of objects relative to various parts of the body and their position relative to other parts of the scene.

There is considerable evidence for the relevance of the distinction between objects and

The Royal Society is collaborating with JSTOR to digitize, preserve, and extend access to
Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences.
www.jstor.org

Downloaded from rstb.royalsocietypublishing.org

positions in human cognition. For example, Allport (1977) found that correct identification of letters or letter groups in arrays of briefly presented words was often combined with incorrect report of their position. An example of evidence that position is represented in memory separately from class and structure is that reported by Mandler & Parker (1976). Neuropsychological evidence indicates that infero-temporal regions are concerned with various aspects of object recognition (see, for example, Mishkin 1982), and that certain parietal structures are concerned with various aspects of spatial position (Hyvarinen 1982).

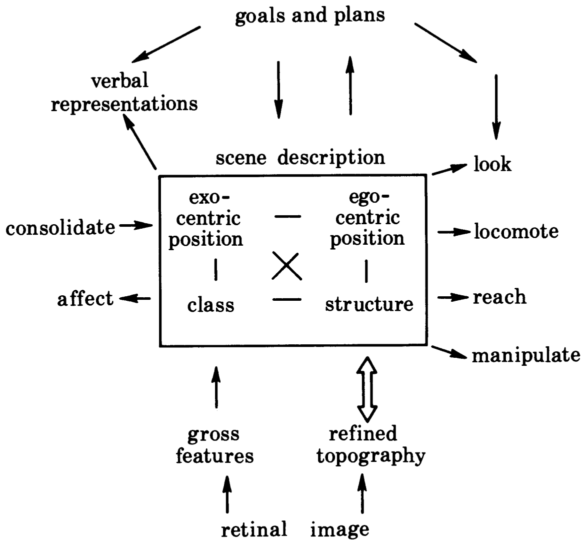

FIGURE 1. Possible classes of representation in visuo-spatial cognition, and their relation
The emphasis is on scene descriptions, and other possible connections

The distinction between class and structure is central to the computational approach to vision developed by David Marr (1982). The particular kind of structural description that he proposes may not be true for human cognition, but it does seem that we can somehow represent the major structural features of objects, such as what parts they have and how these relate. I assume that it is these structural descriptions that are transformed in operations such as mental rotation. To be useful such representations must include more than can be seen from any particular view (e.g. the unseen sides of objects) and even what cannot easily be seen at all (e.g. the insides of opaque objects). Phillips *et al.* (1978) and Neisser & Kerr (1973) provide evidence that higher visuo-spatial representations have such properties. The main evidence that Marr cites for the view that class and structure are separated in human cognition is clinical. Patients with right posterior lesions can recognize objects provided that the view is in some sense simple (Warrington & Taylor 1973). Patients with left posterior lesions are poor at recognizing even simple views but can match different views of objects that they cannot recognize (Warrington 1975). This suggests that structural analysis is more developed in the right hemisphere and classification in the left. Analogous evidence is provided by the ability of people to form descriptions of novel structures that are adequate for them to see analogies or describe how a structure could be decomposed into parts of specific kinds. Indeed, the ability to form schematic but appropriate descriptions of novel objects seems to be a prerequisite for their later recognition.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The distinction between egocentric and exocentric position is, perhaps, less extensively evidenced, but there is a simple argument for it. To look at, move to, or reach for an object its position must be computed relative to the appropriate parts of the body. However, this would be a very inefficient way to specify the location of things in general because their egocentric position will change whenever the person moves. Their position relative to other objects in the scene will change much less often. The extensive neuropsychological evidence that posterior parietal regions are concerned with various aspects of relative spatial position is reviewed by Hyvarinen (1982). It seems clear that area 7 is concerned with some aspects of egocentric position. Regions that deal with the relative positions of objects in the world are not yet positively identified, but they may include the areas that combine visual, auditory and somaesthetic input, because objects are often seen, heard and felt, and shared position in space is perhaps the easiest way to unite such otherwise disparate input.

One final detail of figure 1 needs clarification. The visual input to the scene descriptions is shown divided into two main groups. This is to allow for the possibility that gross features (e.g. low spatial frequencies, or the features computed for the peripheral visual field, or by the superior colliculus) might be related to scene descriptions in a basically hierarchial way, but with more precise or elaborated features being related to scene descriptions in a highly interactive way. This would be consistent with the evidence for fast global processing followed by detailed development (see, for example, Navon 1977). Development by selective enhancement of those particular aspects of stimulation that are relevant to the current interpretation seems likely because objects cannot be defined explicitly in terms of stimulation. The computational feasibility of such an interactive design is suggested by programs such as those of Shirai (1975) and of Hanson & Riseman (1978). The word and object superiority effects (Reicher 1969; Weisstein and Harris 1974) are evidence for some such process in human cognition.

The attempt to specify what visuo-spatial representations represent is clearly important, but for the moment the main point is that scene descriptions are proposed to mediate between representations of stimulation and many of the uses to which they are put. A short-term memory for visuo-spatial information that is not plausibly explained in terms of verbal or iconic representations is part of the evidence for this. I shall argue that short-term visual memory and sensory storage are memories for different classes of representation, and that short-term and long-term visual memory are different kinds of memory for the same class of representation.

Paradigms for studying short-term visual memory

Many paradigms have been developed for studying the role of visual imagery. Many of these involve its role in verbal learning (see, for example, Bower 1970). Most of these do not study short-term memory or visualization directly, but only indirectly through its effects on long-term learning. It cannot be assumed, however, that long-term learning accurately reflects visualization, and results reported later suggest that it may not. More direct techniques are therefore required.

Most prominent among the more direct techniques has been the letter catagorization task devised by Posner & Keele (1967). In this task subjects must categorize two successively appearing letters as the same or different. Comparisons that can be made on the basis of visual appearance (e.g. A followed by A) are made faster than comparisons that cannot (e.g. A followed by a).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

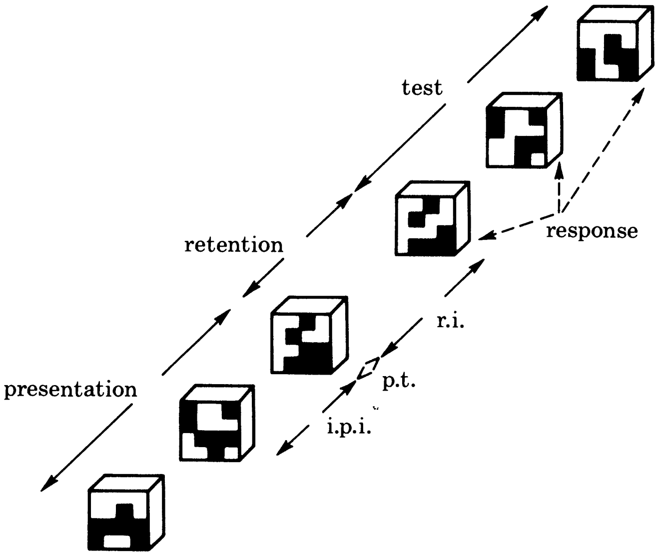

FIGURE 2. Our paradigm for investigating visual memory. A series of patterns is presented, in this case three, and is then tested in reverse order. The times varied are pattern presentation time (p.t.), inter-pattern interval (i.p.i.), and retention interval (r.i.).

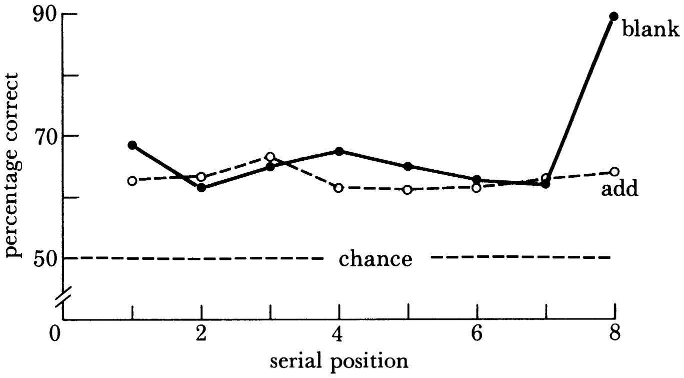

FIGURE 3. The effect of serial position, and intervening activity (p.t. = 1.25 s, i.p.i. = 0.25 s, r.i. = 3 s). During the retention interval the subject was either given no additional task, or required to add five digits. Testing was by recognition. (Modified from Phillips & Christie 1977*a*.)

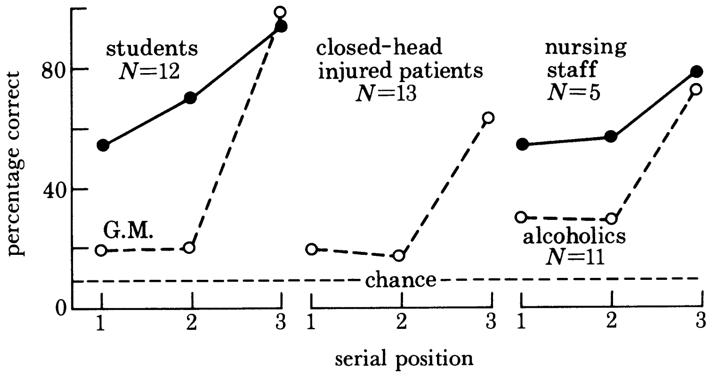

FIGURE 4. Visual memory after closed-head injury and alcoholism (p.t. = 4 s, i.p.i. = 2 s, r.i. = 2 s). Testing was by completion. Closed-head injury data from Wilson, Brooks & Phillips (in prep.). Data on alcoholism from unpublished work by B. McLellan.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Identification latencies for letters that are repeated in close enough succession show related effects (Walker 1978). Results obtained with these techniques have been taken as evidence that letters can be represented either by their names or by a code related to their visual appearance, and that under appropriate conditions this visual code can be actively maintained (Kroll 1975). More recently, however, the possibility has been raised that these letter priming effects may reflect enhanced accessibility of the letter in visual long-term memory (LTM) or of its connection to the appropriate response. Kroll & Parkes (1978) suggest this possibility, and Walker & Marshall (1982) provide a considerable amount of evidence for it. Walker (personal communication does not imply that we cannot actively visualize familiar material such as letters, but that the letter priming effects have not clearly separated active visualization from enhanced accessibility, and are, in fact, more probably due to the latter.

Active visualization can be more clearly implicated in memory for novel patterns. If the subject does not have a ready-made schema for some information, then memory for it cannot be due to enhanced accessibility of that schema. Thus, in verbal memory, digit span is tested by presenting new sequences of digits, not familiar ones. The investigations of visual memory emphasized here are therefore those using novel patterns. These techniques have the added advantage that the learning of new patterns can also be studied, and compared directly with temporary maintenance.

The paradigm that we have used in Stirling for studying visual memory is illustrated in figure 2. A series of patterns is presented and then tested, after a retention interval, in the reverse order of presentation. Thus for a three pattern sequence they are presented in the order 1, 2, 3 and tested in the order, 3, 2, 1. The patterns are novel configurations of simple shapes made by randomly filling cells in a rectangular matrix. Matrix complexity varies, but 4 × 4 cells is most commonly used for studying memory for scene descriptions and very much more complex matrices are used for studying sensory storage (e.g. 10 × 10 cells). The main times controlled are pattern presentation time (p.t.), inter-pattern interval (i.p.i.), and retention interval (r.i.). Testing is by recall, recognition or completion. The similarities in performance across these tests are far more striking than the differences. Recall is tested by presenting empty matrices and asking the subject to recreate the pattern. Recognition is tested by presenting patterns that are either identical to the original or differ by a single randomly selected cell. Completion is tested by deleting a single randomly selected cell from the original and asking the subject to point to the deleted cell. You can work out which cells these are in the examples shown in figure 2. I assume that to do so you most proably form and compare structural descriptions of some kind. The completion technique has proved to be particularly easy for the subject to understand, and is relatively economical of subject time. It has therefore been developed for clinical use (Wilson *et al.* 1982).

The main limitation of this paradigm is that the material is not very interesting. This does not make it unusual, either inside or outside the laboratory. It may nevertheless limit the relevance of the findings; to what extent it does is still unclear.

Serial position effects in visual memory

The basic results obtained with this paradigm are shown in figure 3 (from Phillips & Christie 1977*a*). In this particular experiment eight patterns were presented, with p.t. = 1.25 s, i.p.i. = 0.25 s, and r.i. = 3 s. The solid line shows what happened when the retention interval was blank. The final pattern was remembered very much better than the others, but apart from that the serial position function was flat. Over dozens of such experiments we have never observed a recency effect extending beyond the final item, nor any sign of primacy. This serial position function remained the same when instead of testing all eight patterns on each trial we tested only one selected at random. The serial position effect for reaction time was similar, with subjects responding faster to the final pattern than to the others.

To the subject in the experiment the explanation for these results is straightforward. He can keep the last pattern in mind until it is tested, but the capacity to do this is limited and cannot be extended beyond the final pattern. Support for such an explanation can be seen in the effect of adding five single digits during the interval (the broken line in figure 3). This entirely removes the advantage to the final pattern but has no effect on the others. The causes of this interference are crucial to the understanding of short-term visual memory, and will be discussed in greater detail later.

The main point for the moment is that the last-item advantage and the selective effect of mental arithmetic provide *prima facie* evidence for a process that would normally be called visualization. The next three sections examine the validity of this interpretation by asking whether plausible explanations can be offered in terms of verbal rehearsal, sensory storage, or long-term visual memory.

DISTINCTIVE PROPERTIES OF VISUALIZATION

Differences between visualization and verbal rehearsal

The main reason for claiming that the final item is maintained by visualization rather than by verbal rehearsal is simply that the material is much better adapted to our visual than to our verbal descriptive capacities. Adequate verbal descriptions of patterns of this kind often take a minute or two to produce (Phillips 1974), and rarely less than 15 s. The display time necessary to form a maintainable representation is less than 0.25 s (Avons & Phillips 1980).

Additional support comes from the differences between the performance functions obtained with these pattern tasks and those obtained with verbal or easily verbalized materials. For example, the serial position function observed in these tasks differs from the verbal case in the absence of primacy and in the confinement of recency to the final item.

These observations do not rule out any contribution from verbal processes to the task, but they do make it unlikely that the representation used is either wholly or largely verbal. One severely dysphasic patient that we studied who had very little expressive language and poor comprehension, as well as signs of other deficits, performed quite well on the visualization task provided that only a single pattern was presented. Perhaps verbal processes make a greater contribution to the management of the more complex situation. Further studies of patients with selective deficits are clearly needed. Direct comparison of the performances of human, ape and monkey would also be very revealing.

Differences between visualization and sensory storage

One of the advantages of this paradigm is that by increasing pattern complexity and reducing the retention interval it can be used to study sensory storage. This allows direct comparison of the two uncontaminated by the many confounding changes that are inevitably involved in a change of paradigm. When this is done the differences between the two are clear. The experiments reported by Phillips (1974) show the following:

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

1. *Capacity*. Visualization is of severely limited capacity; with favourable conditions errors average about 10% for patterns in 4 × 4 arrays (an average of 8 filled cells), and are very much more frequent for patterns in 5 × 5 arrays (an average of 12.5 filled cells). Under conditions where sensory storage operates, e.g. retention intervals of less than 100 ms and no masking, errors are rare even in arrays as complex as 32 × 32 (Wilson 1979). Similar tasks also implying very high capacity for sensory storage are reported by Wilson (1981).

2. *Reaction time and pattern complexity*. Reaction time under conditions of visualization increases with pattern complexity. Under conditions of sensory storage very large changes in complexity have little or no effect on reaction time.

3. *Durability*. Patterns can be visualized for at least 15 s without loss, and probably much longer, provided that subjects are free from interfering tasks. Sensory storage only endures beyond 0.5 s in special conditions, and usually, as in our conditions, is even shorter than that.

4. *Masking*. Visual masks have little or no effect on visualization, but integrate with or interrupt information in sensory storage.

5. *Spatial restriction*. Visualized representations refer to structure and position separately. Thus visualization of structural information can be used for comparison with input from many different positions of the field. The information in sensory storage refers jointly to features-in-position, and the ways in which successive representations relate is constrained by this fact. Thus the separate representation of structure and position discussed in the first section is one of the things that differentiates refined topography from scene descriptions.

The differences between visualization and sensory storage are all clear and large. Whatever visual images are like they are not like icons.

Differences between visualization and long-term visual memory

The paradigm also allows direct comparison of visualization with long-term memory. The following are examples of the differences that emerge (see, for example, Phillips & Christie 1977*a*; Christie & Phillips 1979; Avons & Phillips 1980):

1. *Capacity*. Visualization is limited to one pattern or scene at a time (although possibly with dynamic updating to enable description of temporal change), and is also limited in the complexity of the scene description that can be maintained. Visual LTM applies to an indefinitely large number of patterns or scenes. No limit on the complexity of the descriptions that can be held is yet established.

2. *Duration*. Information can be visualized for at least 15 s, but is lost very soon after the onset of interference. Visual LTM deteriorates little or not at all over the time periods we have studied (up to about 10 min so far).

3. *Mental arithmetic*. Simple digit tasks, such as adding five single digits, during the retention interval interfere greatly with visualization but have little or no effect on visual LTM. This is also true for many other tasks, e.g. digit span, but as will be discussed later it is not true for all.

4. *Presentation time*. This has the opposite effect to mental arithmetic. Reducing pattern presentation time from 2 to 0.5 s impairs visual LTM but has little or no effect on visualization. This is because the visualized representation is constructed within about 100 ms, whereas visual learning continues slowly and erratically over many seconds.

5. *Clinical evidence*. Results obtained within the past few months from two kinds of neuropsychological patient give further support to this distinction. If the presentation time per

pattern is increased, performance improves such that the last-item advantage looks less convincing. The curve for students in figure 4*a* shows the results of an experiment with p.t. = 4 s, i.p.i. = 2 s, and r.i. = 2 s. The distinction between the final item and the others can still apply to this situation, however, because it reappears clearly with the patients we tested. G.M. in figure 4*a* is a 21 year old man who had been a college student until a severe closed-head injury received in a motor-cycle acident. He performed well on the final pattern but very poorly on the others. The results in figure 4*b*, *c* suggest that this pattern of performance is common. It therefore seems that visualization can remain good while visual learning is very poor.

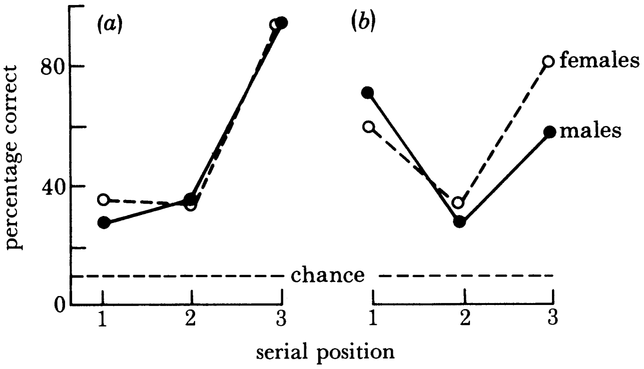

FIGURE 5. Voluntary control of visualization (p.t. = 2 s, i.p.i. = 1 s, r.i. = 4 s). In one session (*a*) subjects were asked to visualize the final pattern of a series of three, and in another session (*b*) they were asked to visualize the first. Testing was by completion.

6. *Relations between visualization and long-term visual memory.* It does not seem plausible to interpret this evidence for visualization in terms of either activated LTM or enhanced accessibility to LTM. The subjects accurately visualized patterns that they had not yet learned. This involves specifying what parts there are, what properties they have, and how they are related to each other. Although this presumably involves the use of well established descriptors, simply selecting such descriptors, whether by activation or some other way, is not adequate because how the descriptors are related must also be specified, and that is what is novel.

The results also warn against equating visualization with visual learning. Longer periods of visualization need not lead to better learning (Avons & Phillips 1980), and, in addition, the patient data suggest that the two can be dissociated.

Voluntary control of visualization

Visualization should not be operationally defined as performance on the final pattern of a series. It seems likely to be at least to some extent a voluntary process dependent upon the strategy adopted (Phillips & Christie 1977*b*, expt 5). To make this point more clearly we ran an experiment in which subjects saw three patterns, with p.t. = 2 s, i.p.i. = 1 s, and r.i. = 4 s. In one session subjects were instructed to keep thinking about the final pattern during the retention interval, and in another session to keep thinking about the first pattern from its offset until its test, while still looking at and responding to the intervening patterns but with as little attention as possible. The results are shown in figure 5. Subjects were able to switch attention to a considerable extent, but as expected found it more difficult to visualize the first pattern than the last. Performance for pattern 2 was not worsened by looking at it while trying to

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

visualize the previous pattern. This further supports the view that under these conditions visualization and visual learning are not closely connected.

The results for males and females are shown separately because this is the only time we have found a sex difference. The only difference here is a significant interaction between sex, instruction and position. Even that must be interpreted with caution as there were only six males and six females, and the experimenter was male throughout. Other than this there seem to be few, if any, sex differences in this task.

If visualization is under voluntary control the question that arises is 'control by what?' This is the main question to which studies of interference with visualization relate.

INTERFERENCE WITH VISUALIZATION

Dissociation of deficits, and double dissociation in particular, is central to the strategy for using neuropsychological evidence to reveal functional specialization (Teuber 1955; Kinsbourne 1971). Selective interference between tasks is an analogous method in cognitive psychology. It has often been misused or misinterpreted, however, and has been much criticized. Before describing our recent selective interference experiments (Phillips & Wilson, in prep.), therefore, I shall make a few comments on the methodological issues involved.

Design and interpretation of dual task experiments

These issues have been much discussed (see, for example, Clayton & Warren 1976) but no fully satisfactory analysis has yet emerged. The issues are too complex to discuss adequately here, but fortunately Allport (1980*a*, *b*) has provided a very incisive critique. I think that his analysis is largely correct, but I should like to add four brief comments. First, the logic of the situation is not restricted just to interference between activities. Selective enhancement could be just as revealing but has been unduly ignored. Second, independence can only be claimed if both tasks are properly assessed. It is very easy to slip from saying 'A does not interfere with B', into saying 'There is no interference between A and B'. Third, double dissociation (e.g. as in Brooks (1967, 1968) and Baddeley & Leiberman (1980)) is evidence for two distinct systems. It does not, as is so often assumed, imply three distinct systems. Fourth, and most important, investigations should be designed so as to isolate the particular component processes that are interacting. This is necessary in order to be able to say what it is that the distinct systems do, instead of simply being able to say that there must be at least two of them. This could be seen as a combination of the subtractive method with the selective interference method. If both tasks are designed so that they have two versions that differ in a limited and specified way, and if the interaction applies to only one version of each task, then the component processes that are interacting can be given a principled specification. The experiments that follow and those of Phillips & Christie (1977*b*, expts 3, 4 and 5) are examples of attempts to meet these requirements.

The interaction of visual memory with other activities

1. *Phonemic categorization*

Although many activities do not interfere with visualization, many do; for example, adding digits, counting backwards by threes, repeating strings of more than five digits, and actively looking at a new matrix pattern so as to be able to remember it. Even randomizing or varying conditions across trials rather than blocking them seems to cause interference. Furthermore,

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

visualization is a voluntary process dependent on the strategy being used. It seemed to us that the most plausible way to summarize these effects was by assuming that visualization requires a continuous contribution from a limited-capacity system that is also involved in a wide range of other processes. This system would therefore be something like the central executive of Baddeley & Hitch (1974). Allport (1980*b*) raises strong objections to this idea. To some extent the force of his criticisms may be reduced if the system is described as a strategic coordinator, with the connotation that tactics may sometimes be part of a strategy, and sometimes not.

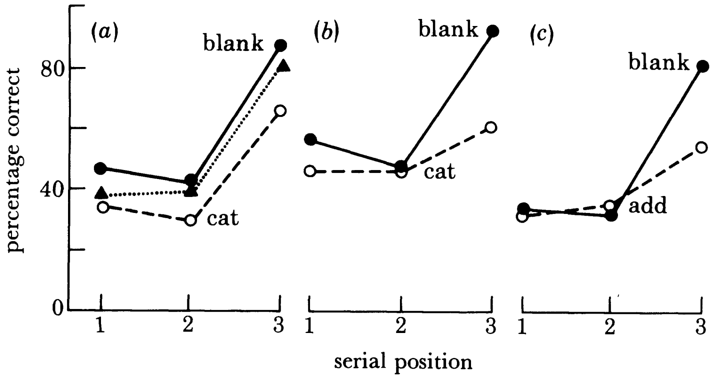

FIGURE 6. Effects of interference on visual memory. (*a*) Phonemic categorization; (*b*) graphemic categorization; (*c*) adding. For all three experiments p.t. = 2 s, i.p.i. = 1 s, r.i. = 9 s. The dotted line in (*a*) shows the effects of just repeating the nonsense word. Testing was by completion.

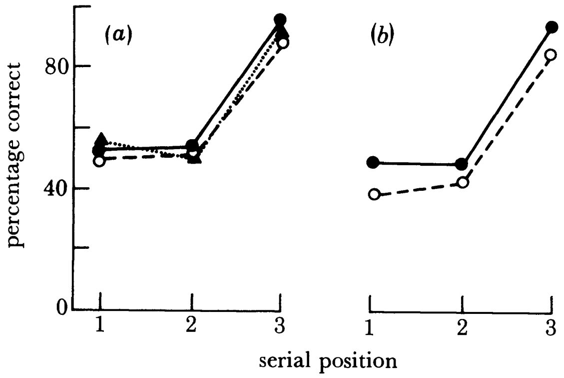

FIGURE 7. The effects of (*a*) looking at matrix arrays (p.t. = 2 s, i.p.i. = 1 s, r.i. = 12 s), and (*b*) eye movements (p.t. = 1 s, i.p.i. = 0.5 s, r.i. = 6 s). In (*a*), ●, matrix empty; ▲, eyes closed; ○, matrix chequered; in (*b*), ●, eyes still; ○, eyes moving.

Our assumption that mental arithmetic interferes with visualization because both compete for a strategic coordinator was challenged by Gerhard Steiner and Philip Seymour. They suggested that mental arithmetic may involve visuo-spatial representations and compete because of that. To test this we tried to find a task that would be very unlikely to involve visuo-spatial representations yet make heavy demands on a strategic coordinator. This was not easy to do, but the task we eventually designed was phonemic categorization.

This task was as follows. During the 9 s retention interval the subject heard a CVCVC nonsense word in which the vowels were always ə as in bird, and the consonants were selected from 16 plosives, fricatives and affricates, eight of these voiced and eight being unvoiced. The

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

subject's task was to categorize the voicing for each of the three consonants. Thus for the nonsense word gəbəs he or she should say 'yes, yes, no'. In addition we included a condition in which the subject's task was simply to repeat the nonsense word. This was to enable a comparison between the effects produced by a novel use of the voicing information, and those produced by a habitual use. The main questions at issue were whether phonemic categorization interferes with visualization, and if so whether it interrupts it entirely. If there is some autonomous buffer capacity as suggested by Baddeley & Lieberman (1980), the latter should not occur.

The results are shown in figure 6. The effects of repeating the nonsense word were not significant. Phonemic catagorization had a highly significant detrimental effect on performance, but to our surprise there was no sign of an interaction with the serial position effect. Our earlier assumption therefore seems to be false. If there is a strategic coordinator, then either phonemic categorization does not make heavy demands on it, or visualization can continue despite such demands. The first of these two possibilities seems unlikely because phonemic categorization does interfere with the visual memory task as a whole.

2. *Adding*

The difference between the interference caused by phonemic categorization and that caused by the other tasks we had studied was so great that we first checked whether any of the various small differences in procedure were responsible. We therefore repeated the procedure of the phonemic categorization experiment except that adding five orally presented digits was the intervening task. The results are shown in figure 6. Adding has a highly significant effect on the final item but no effect on the others. Procedural differences are therefore not responsible for the different effects of adding and phonemic categorization.

3. *Graphemic categorization*

In this task, the subject heard three letter names and had to say whether in his or her own cursive script that letter would contain a fully enclosed region or not. Input and output were thus very nearly identical to that in phonemic catagorization. This time, however, the processes of categorization required reference to visuo-spatial representations.

These results are also shown in figure 6. There is a significant overall effect of intervening activity, and a highly significant interaction with serial position. Graphemic categorization interferes with visualization. It seems as though there may also be interference with the non-final items but this effect does not approach significance.

4. *Perception and matrix arrays*

In this experiment the retention interval was 12 s. During that interval the subject closed his or her eyes, or looked at an empty matrix, or looked at a matrix filled as a chequer-board. The results are shown in figure 7. There is no significant effect of intevening activity. Other experiments of this sort suggest that there may be small effects both of enhancement with empty arrays and of interference with chequered arrays. The main point that these experiments make, however, is that there is very little competition between visualization and whatever visual input processing happens automatically.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

5. *Eye movements*

In this experiment the retention interval was 6 s and during that time subjects were required to make a fixed pattern of eye movements. After presentation of the final pattern they looked up to the top left-hand corner of the video monitor, right round the outside of the monitor in a clockwise motion, and then back to where the final pattern would reappear for test. The results are shown in figure 7. The effect of eye movement was not significant, although it does seem that there may be a small overall effect. There was no sign of any decrease in the last-item advantage. The main point of this experiment, therefore, is that visualization seems to be quite compatible with carrying out a fixed and unrelated sequence of eye movements.

Patterns of interaction with visual memory

The various effects on visual memory discussed above can be summarized in terms of three different patterns of interaction.

1. Effects selective to the final pattern. These include the interference caused by adding digits, graphemic categorization, and many others. This pattern of interaction is evidence for selective effects upon visualization.

2. Effects selective to non-final patterns. Decreasing the pattern presentation time down to about 0.25 s has a large selective effect on non-final items. Similar effects seem to be caused by head injury, alcoholism and probably ageing. These effects are interpreted as a decrement in the acquisition of new cognitive schemata.

3. Effects that are not selective to serial position. The only clear example of this pattern of interference at present is the effect of phonemic categorization. Its interpretation is not yet clear. Perhaps the most obvious possibility is to propose that a strategic coordinator is involved in processing during presentation and test and that this is interfered with by the phonemic task. The effects of mental arithmetic, etc., would then have to be accounted for in terms of specialized interference with visualization. Other interpretations are possible, however. For example, there could be separate specialization for the strategic coordination of linguistic and non-linguistic activities. The compatibility of language with most other activities makes this a possibility. The effect of phonemic categorization would then reflect linguistic contributions to the management of the visual memory task, but this would not imply that visualization was independent of strategic coordination. Alternatively, this result could be taken as support for Allport's view than all interaction is specific, and that there is no central executive or strategic coordinator.

THE ROLE OF VISUALIZATION IN HIGHER VISUO-SPATIAL PROCESSING

There is thus good evidence for a short-term memory for visuo-spatial information that cannot be easily interpreted in terms of either verbal or iconic representations. It is probably closely allied to our ability for visual thought. In relation to the classification suggested in figure 1, information about structure seems to be what is primarily involved. Memory for the positions of patterns was not needed, and pattern class could contribute little because of their novelty. The long-term learning in this task could be interpreted as the early stages in the acquisition of classifying schema allowing later recognition and recall. If this is so, what of short-term memory for position? There are suggestions that there might be various forms, quite different from that which we have studied. For example, Thomson (1980) reports evidence for a short-term visual memory that guides locomotion but is limited to just 8 s. A better validated division of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

scene descriptions into the major classes of information represented and of the various kinds of memory associated with them is one of the major goals for the future.

Many members of the Psychology Department, University of Stirling, contributed to the research reported in this paper. Particularly valuable contributions were made by Jim Anderson, Ranald Macdonald, Gerry Orchard, Francis Pratt and Lindsay Wilson.

REFERENCES

Allport, D. A. 1977 On knowing the meaning of words we are unable to report: the effects of visual masking. In *Attention and performance*, vol. 6 (ed. S. Dornic), pp. 505–533. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Allport, D. A. 1980*a* Patterns and actions: cognitive mechanisms are content-specific. In *Cognitive psychology new directions* (ed. G. Claxton), pp. 26–64. London: Routledge & Kegan Paul.

Allport, D. A. 1980*b* Attention and performance. In *Cognitive psychology new directions* (ed. G. Claxton), pp. 112–153. London: Routledge & Kegan Paul.

Avons, S. E. & Phillips, W. A. 1980 Visualization and memorization as a function of display time and poststimulus processing time. *J. exp. Psychol.: hum. Learn. Memory* **6**, 407–420.

Baddeley, A. D. & Hitch, G. 1974 Working memory. In *The psychology of learning and motivation*, vol. 8 (ed. G. Bower), pp. 47–89. London: Academic Press.

Baddeley, A. D. & Lieberman, K. 1980 Spatial working memory. In *Attention and performance*, vol. 8 (ed. R. S. Nickerson), pp. 521–540. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Block, N. 1981 *Imagery*. Cambridge, Massachusetts: M.I.T. Press.

Bower, G. H. 1970 Analysis of a mnemonic device. *Am. Scient.* **58**, 496–510.

Brooks, L. R. 1967 The suppression of visualization by reading. *Q. Jl exp. Psychol.* **19**, 289–299.

Brooks, L. R. 1968 Spatial and verbal components of the act of recall. *Can. J. Psychol.* **22**, 349–368.

Chase, W. G. 1973 *Visual information processing*. New York: Academic Press.

Christie, D. F. M. & Phillips, W. A. 1979 Simple drawing and pattern completion techniques for studying visualization and long-term visual knowledge. *Memory Cogn.* **7**, 360–367.

Clayton, K. N. & Warren, M. W. 1976 Methodological problems with the use of the retroactive interference paradigm to infer what is stored. *Memory Cogn.* **4**, 237–243.

Hanson, A. R. & Riseman, E. M. (eds) 1978 *Computer vision systems*. New York: Academic Press.

Hyvarinen, J. 1982 Posterior parietal lobe of the primate brain. *Physiol. Rev.* **62**, 1060–1129.

Kinsbourne, M. 1971 Cognitive deficit: experimental analysis. In *Psychobiology* (ed. J. L. McGaugh), pp. 285–348. New York: Academic Press.

Kroll, N. E. A. 1975 Visual short-term memory. In *Short-term Memory* (ed. D. Deutsch & J. A. Deutsch), pp. 154–181. New York: Academic Press.

Kroll, N. E. A. & Parks, T. E. 1978 Interference with short-term visual memory produced by concurrent central processing. *J. exp. Psychol.: hum. Learn. Memory* **4**, 111–120.

Navon, D. 1977 Forest before trees: the precedence of global features in visual perception. *Cogn. Psychol.* **9**, 353–383.

Neisser, V. & Keer, N. 1973 Spatial and mnemonic properties of visual images. *Cogn. Psychol.* **5**, 138–150.

McGee, M. G. 1979 Human spatial abilities. *Psychol. Bull.* **86**, 889–918.

Mandler, J. E. & Parker, R. E. 1976 Memory for descriptive and spatial information in complex pictures. *J. exp. Psychol.: hum. Learn. Memory* **2**, 38–48.

Marr, D. 1982 *Vision*. San Francisco: W. H. Freeman.

Mishkin, M. 1982 A memory system in the monkey. *Phil. Trans. R. Soc. Lond.* B **298**, 85–96.

Phillips, W. A. 1974 On the distinction between sensory storage and short-term visual memory. *Percept. Psychophys.* **16**, 283–290.

Phillips, W. A. & Christie, D. F. M. 1977*a* Components of visual memory. *Q. Jl exp. Psychol.* **29**, 117–133.

Phillips, W. A. & Christie, D. F. M. 1977*b* Interference with visualization. *Q. Jl exp. Psychol.* **29**, 637–650.

Phillips, W. A., Hobbs, S. B. & Pratt, F. R. 1978 Intellectual realism in childrens' drawings of cubes. *Cognition* **6**, 15–33.

Posner, M. I. & Keele, S. W. 1967 Decay of information from a single letter. *Science, Wash.* **158**, 137–139.

Reicher, G. M. 1969 Perceptual recognition as a function of meaningfulness of stimulus material. *J. exp. Psychol.* **81**, 275–280.

Shirai, Y. 1975 Analyzing intensity arrays using knowledge about scenes. In *The psychology of computer vision* (ed. P. H. Winston), pp. 93–114. New York: McGraw-Hill.

Teuber, H. L. 1955 Physiological psychology. *A. Rev. Psychol.* **6**, 267–296.

Thomson, J. A. 1980 How do we use visual information to control locomotion? *Trends Neurosci.*, October, pp. 247–250.

Walker, P. 1978 Short-term visual memory: the importance of temporal and spatial separation of successive stimuli. *Q. Jl exp. Psychol.* **30**, 665–679.
Walker, P. & Marshall, E. 1982 Visual memory and stimulus repetition effects. *J. exp. Psychol.: gen.* **111**, 348–368.
Warrington, E. K. 1975 Selective impairment of semantic memory. *Q. Jl exp. Psychol.* **27**, 635–657.
Warrington, E. K. & Taylor, A. M. 1973 The contribution of the right parietal lobe to object recognition. *Cortex* **9**, 152–164.
Weisstein, N. & Harris, C. S. 1974 Visual detection of line segments: an object superiority effect. *Science, Wash.* **186**, 752–755.
Wilson, J. T. L. 1979 Event perception and sensory storage. Ph.D. thesis, University of Stirling.
Wilson, J. T. L. 1981 Visual persistence at both onset and offset of stimulation, *Percept. Psychophys.* **30**, 353–356.
Wilson, J. T. L., Brooks, N. & Phillips, W. A. 1982 Using a microcomputer to study perception, memory and attention after head injury. Poster presentation at 5th International Neuropsychology Society European Conference, Deuville, France.

Discussion

D. A. ROUTH (*Department of Psychology, University of Bristol, U.K.*). I am generally quite enthusiastic about Dr Phillips's general orientation to this area in terms of structural descriptions. However, given the conceptual overlap between this approach to associative networks and a propositional one I am deeply puzzled as to why his serial position curves for supraspan sequences of patterns should be so different from the ubiquitous form observed for sequences of verbal events. It occurs to me that his preferred measure of retention (percentage correct recognition) for the same–different paradigm must be, in terms of signal detection theory, the average of the hit rate together with the complement of the false alarm rate. Now, in terms of the unit square, it is fairly easy to see that the locus of points for which such a measure remains constant must be a line parallel to the diagonal (0, 0) to (1, 1). Obviously, such a line must cut through a family of r.o.c. curves, each associated with a different value of d'. Accordingly, therefore, is it not conceivable that Dr Phillips would have obtained a more conventional serial position curve if he had assessed performance in terms of d' or some such related index? If so, might there not be greater continuity between the verbal and non-verbal domains that he has suggested?

W. A. PHILLIPS. It is conceivable but I do not think that it is so. In the early experiments we used signal detection theory measures, but they never told us anything not already apparent in the total percentage correct, and in the separate hit and false alarm rates. As the latter are more comprehensible, and do not require unlikely assumptions, those are what we use. The shape of the serial position function shown is robust across many different conditions, and across recognition, recall and completion tests. Explanations in terms of specific measurement artefacts are not very plausible. In any case the absence of primacy can be easily explained. One main cause of primacy in verbal memory is that early items are rehearsed more than later items. As I said, in our experiments visualization produces little or no improvement in visual LTM. This will therefore reduced or remove primacy.

I find Dr Routh's assumption that reference to structural description implies acceptance of a generalized propositional approach interesting and important. As I implied in an earlier section I am very sceptical of 'pictorial' approaches to imagery. This is because images are very different from icons, and icons are 'pictorial' in clearly definable ways. However, this in no way commits me to a generalized propositional position: I believe in functional specialization. Scene descriptions are specifically designed to represent visuo-spatial information; the properties they have follow primarily from that. Verbal representations are designed to represent linguistic entities, and it is because these differ in fundamental ways from non-linguistic entities that language gives distinct intellectual abilities to those that have it.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

P. T. **Smith** (*Department of Psychology, Univeristy of Reading, U.K.*). Am I right in supposing that Dr Phillips sees the earliest stages of visual processing as being concerned with identifying visual properties in particular spatial locations? If so, how should such ideas be reconciled with the work of Anne Treisman and her colleagues (e.g. Treisman & Gelade 1980; Treisman & Schmidt 1982), who show that in pre-attentive processing (equals sensory memory?) visual properties such as colour and shape may be identified without accurate specification of spatial location? Might it not be that Dr Phillips's tasks, involving the identification of presence or absence of specific squares in a chessboard array, because of the premiums that they place on subjects' accurate specification of spatial location, are missing functionally important aspects of early visual processing (i.e. parallel processing of visual features without spatial location information)?

References

Treisman, A. & Gelade, 1980 A feature integration theory of attention. *Cogn. Psychol.* **12**, 97–136.
Treisman, A. & Schmidt, H. 1982 Illusory conjunctions in the perception of objects. *Cogn. Psychol.* **14**, 107–141.

W. A. **Phillips**. I do not know Treisman's experiments in detail, but as far as I can see there is no difficulty in interpreting her results in terms of criteria specified upon scene descriptions. The extensive evidence for automatic access to scene descriptions, e.g. the various Stroop effects, makes it implausible to equate the boundary between sensory representations and scene descriptions with pre- and post-attentive processing, and I certainly intend no such equation.

The physiological evidence for spatially localized feature representation in early visual processing is clear, and in good agreement with psychophysics. I don't think Treisman claims that her results are in conflict with all that evidence.

I think that what Dr Smith says is right in an important way. The spirit of our experiments is analytic. We are not trying to study everything at once, even within the domain of scene descriptions. However, as I said at the end of the paper, I think that these particular experiments emphasize structure, and not, as Dr Smith assumes, position. Recognition tests, e.g. figure 3 and many other experiments, give essentially the same results as completion, so locating the position of the change is not critical. Even when completion is used, pointing does not have to be very precise because the arrays are divided into 4 × 4 cells, and subjects simply indicate which cell has been deleted. The results would probably be quite different if the task was to remember the position of a single element in a large blank screen.

The results I have discussed, therefore, seem to me to emphasize memory for the structure of novel configurations, and to tell us little about memory for position. They also tell us nothing about memory for surface properties, such as texture and colour, and these would also have to be included in a more complete account. Perhaps information about such properties is included in scene descriptions by being associated with the surfaces represented in the structural schema.

In my opinion the most important omission concerns the ancient problem of universals. Stored schemata are assumed to be used for recognition and description, but no indication is given as to how this is done. The problem of how abstract ideas gain sustenance from and lend direction to the realities of the moment remains the central mystery of cognition.