

Werner X. Schneider

Visual-spatial working memory, attention, and scene representation: A neuro-cognitive theory

Received: 16 January 1998 / Accepted: 25 July 1998

Abstract This paper addresses the issue of how visual-spatial working memory, attention, and scene representation are related. The first section introduces a modified two-stage conception of visual-spatial processing. “Stage one” refers to low-level visual-spatial processing and computes in parallel for the currently available retinal information “object candidates,” here called “visual-spatial units.” An attentional process called “unit selection” allows access to stage two for one of these units at a time. Stage two contains high-level visual-spatial information that can be used for goal-directions (e.g., verbal report, grasping). It consists of three parallel processing streams. First, the currently selected unit is recognized; second, a spatial-motor program for the selected unit is computed; and third, an “object file” is set up for the selected unit. An object file contains temporary episodic representations of detailed high-level visual-spatial attributes of an “object” plus an “index.” An index acts as a pointer and is bound via temporary connections to the attributes of the file.

Section two of this paper specifies one part of stage two in more detail, namely visual-spatial working memory (VSWM). It can contain up to four object files. A first central claim is that during sensory-based processing for working memory (“access”), one object file is always “on-line,” and up to three other object files are “off-line”. A second central claim is that the process of setting up an object file depends on the number and the activation level of already stored files. Based on the concept of activation-based competition between object files, it is postulated that the more files that are stored and the higher their activation is, the longer it takes for a newly set up object file to reach a sufficient level of ac-

tivation. Activation-based competition is also used to explain “short-term forgetting” by “interference.” A third central claim about VSWM is that a “refreshment” process exists that increases the activation level of an index of an object file in order to prevent forgetting or in order to bring the file back to the state of controlling the current action. Finally, section three gives a selective look at a number of experimental data such as the attentional blink, backward masking, dwell time effects, transsaccadic memory, and change blindness. New explanations are offered and new predictions made.

Introduction

The scientific conception of the visual-spatial representation of the world inside the head has changed considerably during the last decades. Neurobiological research led to counter-intuitive insights – mainly based on single unit recordings in monkeys, lesion data in humans and monkeys, and recently on neuro-imaging data. Consensus emerged that the internal visual-spatial generation of “what is where” in the world occurs in a highly distributed and parallel manner at the several quasi-hierarchically arranged levels involving multiple processing steps and requiring a relatively large part of the cortex (e.g., Felleman & Van Essen, 1991; Zeki, 1993; Crick, 1994; Posner & Raichle, 1994; Milner & Goodale, 1995, for overviews). For instance, within the primary visual cortex there is parallel processing of local motion, color, and orientation signals (e.g., DeYoe & van Essen, 1988; Livingstone & Hubel, 1988), while at higher levels of cortical processing, visual-based recognition within inferior-temporal cortex occurs in parallel with spatial computations within the posterior parietal cortex (e.g., Ungerleider & Haxby, 1994; Milner & Goodale, 1995). Referring to behavioral experimental research, further astonishing characteristics of the visual-spatial scene representation have been revealed recently. The current internal representation of the visual world that can be used for purposeful interaction with

W. X. Schneider
Ludwig-Maximilians-Universität,
Allgemeine und Experimentelle Psychologie,
Leopoldstr. 13, 80802 München,
Germany;
e-mail: wxs@psy.uni-muenchen.de

the environment seems to be highly selective and capacity-limited. It encompasses only a few objects and their spatial relationships (e.g., Irwin, 1992; Henderson, 1994; Rensink, O'Regan, & Clark, 1996; Simons & Levin, 1997). Furthermore, there is consensus about the existence of "visual attentional processes" that determine the current content of such a highly selective capacity-limited scene representation (e.g., Neisser, 1967; Treisman & Gelade, 1980; Duncan & Humphreys, 1989; Bundesen, 1990; Kahneman, Treisman, & Gibbs, 1992). However, it is under ongoing debate how these attentional processes should be conceptualized (see, e.g., Allport, 1993; Schneider, 1993; Pashler, 1997; Styles, 1997; for overviews). Moreover, the internal visual-spatial representation not only contains "on-line" sensory-based information but also has a short-term memory component. For instance, when we move our eyes from one fixation to the next, selected contents of prior fixations are not lost but are temporarily maintained as part of the scene representation (see Pollatsek & Rayner, 1992; Bridgeman, van der Heijden, & Velichkovsky, 1994; Irwin, 1996, for overviews).

These three characteristics of visual-spatial processing in the primate brain – the highly selective internal scene representation that is useful for action, the control of current representation content by the attentional processes, and its short-term memory component – are the main topics of this paper. More precisely, questions such as what is meant by visual attentional processes, how perceptual and short-term memory processes are related, and how these different processes contribute to the current highly selective scene representation will be addressed, and answers at the neuro-cognitive level will be sketched. "Neuro-cognitive level" means, first, that the above mentioned neurobiological data and theoretical concepts will also be taken into account, and, that – second, if possible – relationships between "cognitive" information processing concepts and neurobiological concepts such as neural activation or brain areas will be specified.

Referring to the question of what is meant by "visual attention," a modified version of the two-stage approach first introduced by Neisser (1967) will be advocated here. The basic idea is that visual-spatial processing consists of two major steps, so-called stages. In stage one, which Neisser called the "preattentive stage," elementary visual-spatial representations are computed, and this information is segmented into "object candidates." These object candidates do not contain information about their identity, since no recognition has occurred yet. Only those object candidates that are within the "focus of attention" are allowed access to stage two, which implies recognition and other higher-level visual processes.

An up-to-date neuro-cognitive version of this two-stage conception will be presented in this paper. Figure 1 gives an overview of this conception. Stage one contains the parallelly computed object candidates, here called "visual-spatial units." An attentional process selects one

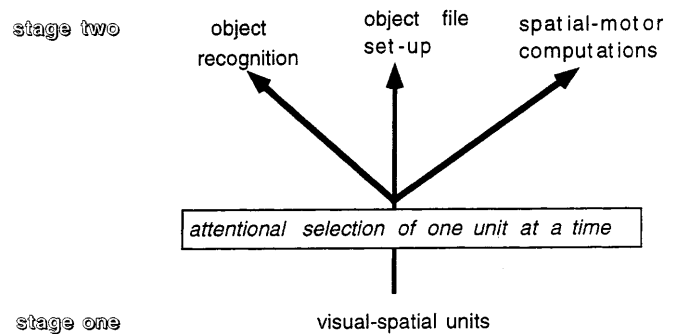


Fig. 1 A flow chart of the modified two-stage theory

of the units at a time for access to stage two. Stage two implies not just recognition of this unit, but also spatial-motor programming for movements towards it (e.g., preparing a grasping movement towards the selected unit) as well as explicit binding of the high-level visual-spatial attributes of the unit into a temporary episodic object representation called an "object file" (Kahneman & Treisman, 1984; Kahneman et al., 1992).

A modified two-stage theory of visual-spatial processing

This section will introduce a modified and "modern" version of Neisser's (1967) two-stage framework for visual-spatial processing. "Modern" means that neuro-cognitive data and theoretical conceptions will also be taken into account that were not available at the time Neisser wrote his book. A brief and selective overview of this neurobiological research on visual-spatial processing should be given now.

The primate visual system¹ is probably one of the best known systems of the brain. The main pathway for processing visual-spatial information runs from the retina via the LGN to the primary visual cortex (V1). Starting from V1, a quasi-hierarchy of at least 30 areas for computing useful visual-spatial information can be distinguished for the macaque (e.g., Felleman & Van Essen, 1991). Basically, these cortical areas compute and represent visual-spatial information about objects in space (Marr, 1982) at multiple, quasi-hierarchical levels in a parallel and distributed fashion (e.g., Zeki, 1993; Crick, 1994; Milner & Goodale, 1995, for overviews). Visual-spatial information refers to dimensions such as color, shape, texture, size, or location (in separate coordinate systems). The concept of "multiple quasi-hierarchical levels" means that lower levels of the hierarchy (e.g., V1 and V2) compute simpler visual-spatial attributes such as orientation of small contours. Higher

¹ The term "primate" visual system refers to humans and monkeys (e.g., macaques). If not stated otherwise in the text, it is assumed that the same principles of neuro-cognitive organization (e.g., parallel and distributed dorsal and ventral processing) hold for both.

levels of the hierarchy – parts of the inferior-temporal cortex (IT) or parts of the posterior parietal cortex (PPC) – compute more complex attributes such as entry-level object recognition categories (e.g., based on categorized shape) or arm-specific spatial coordinates (for overviews, see Oram & Perrett, 1994; Milner & Goodale, 1995). However, feedback connections and short cuts between levels rule out a strict serial hierarchy of processing steps.

Two major pathways, a ventral pathway running from V1 to IT and a dorsal pathway running from V1 to PPC, demonstrate the parallel organization of visual-spatial cortical computation. The ventral pathway computes information useful for entry-level object recognition (the “what” aspect), while the dorsal pathway computes spatial information for motor actions such as grasping and for the perceptual localization of objects in space (the “how” and “where” aspect: see, e.g., Ungerleider & Haxby, 1994; Milner & Goodale, 1995). Visual-spatial information useful for goal-directed, task-dependent interactions with the environment – according to the theory presented here – is only available at higher levels of the visual-spatial processing system, that is, stage two. Referring to the ventral part, output to the phonological (verbal) and semantic system comes from the inferior-temporal cortex, primarily from the entry-level object recognition machinery. Referring to the dorsal part, output to the spatial-motor system (e.g., for grasping and saccading) arises mainly from certain areas of the superior posterior parietal cortex (see, e.g., Milner & Goodale, 1995).

Stage one: Parallel low-level visual-spatial processing for segmented visual-spatial units

Stage one computes and represents within “early” cortical areas (e.g., V1, V2) elementary low-level visual-spatial information in parallel based on the current retina pattern. Elementary visual information means information about color, texture, oriented contours, etc.. Low-level visual processing not only computes visual-spatial information, it also segments the information into “object candidates” (see, e.g., Neisser, 1967; Kahneman, 1973). The results of segmentation processes are low-level object representations which are sometimes called “perceptual units” (e.g., Woodworth & Schlossberg, 1954; Bundesen, 1990). In reference to this tradition, they will be called “visual-spatial units” in this paper.

How are visual-spatial units represented within the parallel and distributed low-level processing system? They are represented by their low-level visual-spatial attributes (color, orientated contours, etc.) in distributed cortical maps (e.g., within V1, V2 in the macaque; see, e.g., Hubel & Wiesel, 1977; DeYoe & van Essen, 1988; Zeki, 1993) – see Fig. 2 – as well as by “regions.” Regions are the result of the segmentation processes within the low-level visual-spatial processing maps. They con-

tain information about the location, size, and rough shape of an object (see, also Grossberg, Mingolla, & Ross, 1994; Grossberg, 1994), and they are probably computed within an early part of the dorsal path. The idea that “location” representations of objects are derived independently from visual attributes within the corresponding “feature maps” has been forcefully articulated within cognitive psychology by Treisman (1986, 1988). The theory presented here considers each region as a representative of a visual-spatial unit (a low-level “object candidate”). Each region allows access to the low-level visual-spatial attributes of its corresponding visual-spatial unit. This is possible due to the location-based retinotopic coding format of the distributed low-level visual-spatial dimension² maps (color, oriented contours, etc.) and of the map(s) for representing regions. In other words, low-level attribute representations carry (implicit) location information in early areas (V1/V2) – the attributes are not “free-floating” (Treisman & Gelade, 1980). This retinotopic coding scheme therefore allows the attentional selection of one visual-spatial unit via accessing its region. Finally, it has to be mentioned that stage one is restricted to contents of the current eye fixation. It does not refer to a transsaccadic scene representation that is restricted to stage two.

Stage two: High-level visual-spatial processing for recognition, spatial-motor programming and object file set-up

As shown in Fig. 2, and in line with Neisser’s (1967) assumption, stage two implies entry-level recognition of objects, which is located within the higher levels of the ventral pathway, such as within the inferior-temporal cortex (see, e.g., Oram & Perrett, 1994; Tanaka, 1997). However, not just recognition but also complex spatial processing within the dorsal pathway is a characteristic of this second stage. Two kinds of spatial processing streams can be distinguished (Milner & Goodale, 1995), namely, spatial computations for motor actions and spatial computations for perception. The spatial stream for motor actions delivers the required spatial parameters for actions. For instance, when a grasping movement to a certain object has to be made, information (parameters) about the location, size, and rough shape of this object has to be delivered to the motor system (see Allport, 1987; Neumann, 1987). Spatial computations for perception are computations of where objects are located in relation to other objects – the so-called allocentric spatial information.

Further-more, it is postulated that stage two implies a further function besides object recognition and high-level spatial computations, namely, the set-up of an object file (see Kahneman & Treisman, 1984; Kahne-

²The terms “dimension” and “attribute” (“feature”) are distinguished here. For instance, color is a dimension, while red is an attribute of the dimension color.

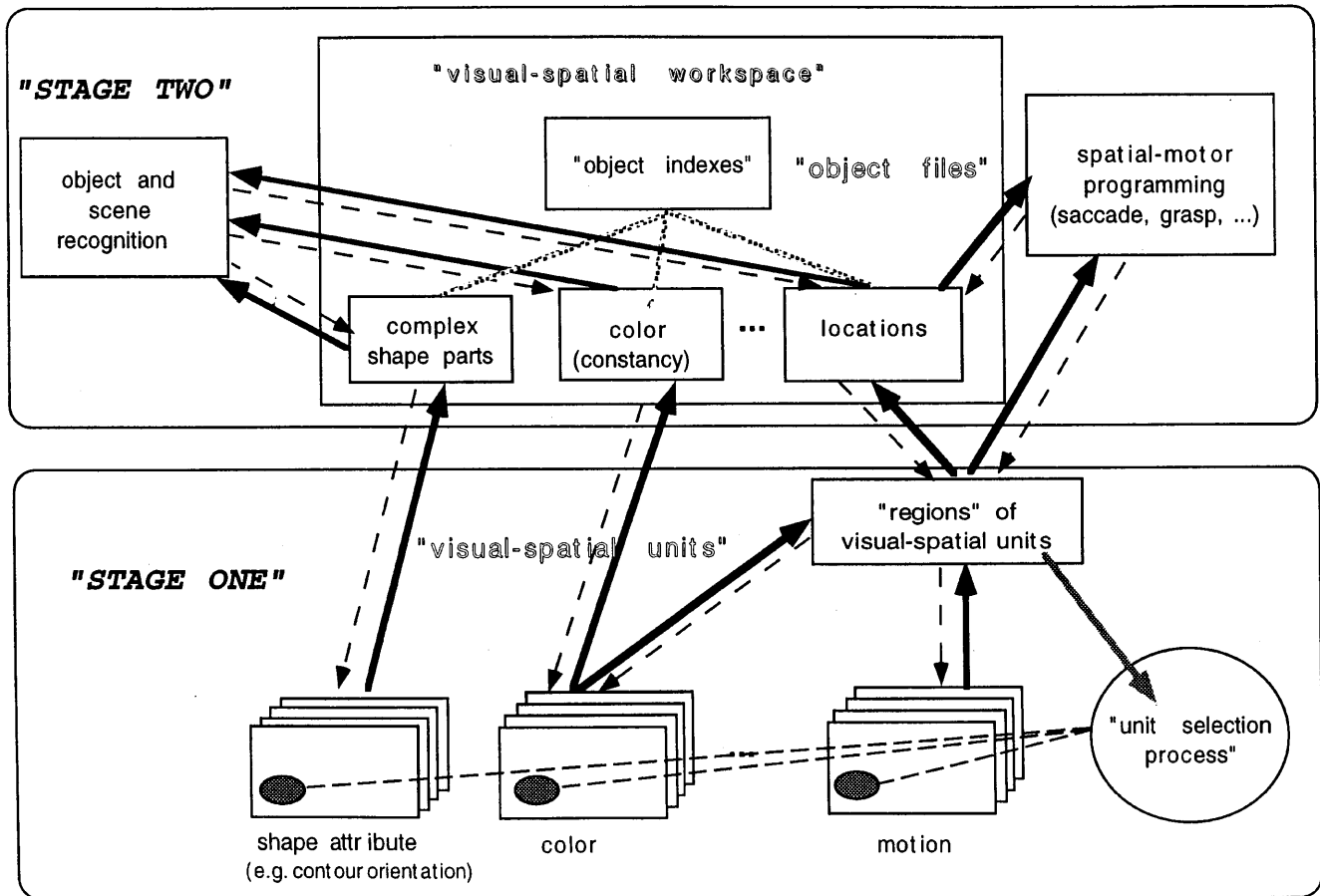


Fig. 2 The architecture of the modified two-stage theory

man, Treisman, & Gibbs, 1992; Irwin & Andrews, 1996; Leslie, Xu, Tremoulet, & Scholl, 1998). An object file refers to temporary episodic information about an object. It contains high-level visual-spatial attributes and it explicitly represents that these attributes belong to an individual object. This allows the internal representation of object-continuity over time and space (see Kahneman et al., 1992), such as in the form of "object permanence" (e.g., Leslie et al., 1998). "Object permanence" refers to the ability to represent an external object internally, even if it is temporarily hidden in the outside world. Furthermore, when an object moves in the outside world, the object-continuity capacity of the primate visual-spatial system allows a representation that the internal visual-spatial attributes still refer to the same entity.

What is the content of an object file? Following Kahneman et al. (1992), spatial-temporal information is the basic content, that is, information about "where something is when." Additionally, high-level attribute representations about visual appearance (shape, color, etc.) are part of the file. How are "recognition" and "spatial-motor programming" related to object files? As displayed in Figs. 1 and 2, it is assumed here that the results of recognition and spatial-motor programming are not part of the file – they are computed in parallel to the

process of the object file set-up. If just recognition output that has lost information about detailed visual appearance (e.g., the shape) of an object were represented, then detailed high-level visual-spatial attributes (e.g., precise shape and color representations) would not be part of an object file. Such detailed information within an object file, however, is necessary for goal-directed actions such as grasping an object or for maintaining the precise visual-spatial characteristics of a hidden object. How this claim is reconciled with "transsaccadic memory" data that seem to suggest only undetailed abstract information as the content of object files (see, e.g., Irwin, 1992, 1996) will be explicated in a later section. Furthermore, stage two and its object files contain information about spatial relationships of the objects (allocentric spatial information) – otherwise, actions that require information about the relative positions of objects would not be possible (e.g., driving a car during rush hour).

How is it encoded that different attributes (shape, color, etc.) belong to one common object ("objectness")? Extending the original object-file concept (Kahneman et al., 1992), it is assumed that "objectness" is represented by an "index" (see Leslie et al., 1998) and by the bindings of this index to the visual-spatial attributes of

the file – see Fig. 2. An index can be considered as a “pointer” that allows one to access the attributes of an object file and to realize spatial-temporal continuity (see also Pylyshyn, 1989). The most important reason for assuming the existence of an index will be introduced in the section of visual-spatial working memory – namely, that an index allows “competition between object files” as a whole. Bindings between an index and its attributes require temporary connection between attribute representations and the index. Such temporary connections in the form of “fast weights” (probably short-term synaptic modifications) could be set up by synchronized firing of the attributes and the index (see von der Malsburg, 1981; Hummel & Biederman, 1992). Analogously to the Hebb (1949) rule, the degree of synchronized firing would determine the strength of the temporary connection. How many object files can be set up? Based on experimental evidence from “whole report” studies in the iconic memory research tradition (e.g., Bundesen, 1990, Irwin, 1992; Irwin & Andrews, 1996), four object files seem to be a reasonable number (see also Luck & Vogel, 1997).

Attentional control in setting up a visual-spatial scene representation

How visual attention processes work, how they are controlled, and what their effects are has been the subject of intensive research in the last decades; despite an impressive amount of data, no consensus has been reached yet on how to answer these questions (for overviews, Allport, 1993; Schneider, 1993; Duncan, Humphreys, & Ward, 1997; Pashler, 1997; Styles, 1997). For the purpose of this paper, it is not necessary to address all these debated attentional issues. Instead, four assumptions will be made. First, following Schneider (1995), it is assumed that one visual-spatial unit at a time is selected for access to stage two. The corresponding visual attentional process should therefore be called a “unit selection process.” Unit selection is not the only visual attentional process for controlling the activation flow in visual-spatial processing. The visual-spatial system is, within certain limits, also able to define in a flexible and task-dependent way what is computed as an “internal object.” This means, within the two-stage framework, that the system can determine which part of low-level visual-spatial information is segmented into units (for instance, whether a unit encompasses low-level attribute representations of a whole face or parts of the face such as the nose). However, in this paper, the consideration of “visual attention” will be restricted to the unit selection process (see Logan, 1996, for an interesting theory on attentional segmentation).

Second, what does “attentional selection” mean, or more precisely, what are the consequences of unit selection? The general constraint is that unit selection gates the activation flow from stage one to stage two in such a way that among the many units that are com-

puted in parallel in stage one, only one unit at a time will access stage two. For the neurophysiological level, it is assumed that gating the activation flow is realized by a modification of the synchronized firing of neurons. More precisely, those low-level neurons (e.g., V1/V2) that belong to the selected unit should fire within a “time slice” (time window) in a synchronized manner, while neurons representing other units do not fire in synchrony during that time slice. The main advantage of such a gating conception is that modifications in the temporal firing of low-level neurons can be evaluated and “used” much faster at the higher-level stage two than modification of average firing rates; see Singer et al. (1997) for an informative discussion of these issues of synchrony and neural processing speed.

Third how is the attentional process of unit selection controlled? Besides sensory-driven factors such as “pop-out” or “singletons” (see, e.g., Treisman & Gelade, 1980; Yantis, 1998), the current “task set” plays a fundamental role (see, e.g., Posner, 1980, for such a distinction). “Task set” refers in behavioral experiments to the attributes that specify the relevant stimulus and response category (see van der Heijden, Heij, Phaf, Buijs, & van Vleet, 1988; Bundesen, 1990). For instance, if in a reaction-time task a left key should be pressed in response to a green light, then the task set consists of the color “green” and the location attribute “left” (for pressing the left response key). How the task set influences the control of the unit selection process and how this interacts with sensory factors will not be discussed here (see Schneider, 1995, for one suggestion) because it does not directly touch the issues of interest. However, two relevant assumptions about task-dependent control will be stated, since they are required for understanding attentional control of “working memory contents” (next section). One assumption refers to the case when task-dependent control does not directly lead to selection of the task-specified unit. An example is an ineffective visual search task in which the search time increases with the number of distractors (e.g., searching for the target “R” among “Qs” and “Ps”). In such a situation a process of fast “serial scanning,” that is, selecting one region after another, might be used. It is stopped as soon as the relevant unit (target) is detected at stage two (see Treisman & Gelade, 1980; Wolfe, 1994; for such conceptions). Another assumption about control refers to the question of how the task set is implemented at the neural level, how it is fed into the visual-spatial neural processing system. It is postulated that the current task set increases activation of task-relevant attributes within the corresponding parts of high-level visual-spatial processing of stage two (see Schneider, 1995). Referring to the example of the reaction-time task introduced above, representations of the color “green” and the location “left” are supplied with extra neural activation at the corresponding neural sites of stage two.

Fourth, unit selection implies that only one unit at a time can be selected for access to stage two, that is, spatial-motor programming, recognition, and object file

set-up (see also Duncan, 1996). Confirmative evidence for this claim comes from experimental data by Deubel & Schneider (1996; see also Schneider & Deubel, 1995) and Deubel, Schneider, & Paprotta (1998). The data show that during the preparation of a spatial motor action (saccade or pointing movement) to one movement target located among distractor objects, recognition of briefly presented objects is restricted to the movement target object. Distractor objects close to the movement target usually could not be reliably identified and reported. These data support the idea that only one object at a time can be selected during sensory-based processing for spatial-motor action control and recognition (see also Rizzolatti, Riggio, & Sheliga, 1994; Schneider, 1995).

In summary, a modified and extended two-stage conception of visual-spatial processing is postulated here. Stage one refers to low-level vision. It computes elementary visual-spatial information within a single eye fixation and segments this information into visual-spatial units – low-level object representations. The result of this segmentation process are “regions” that make the location, rough shape, and size of an object explicit. Regions are representatives of units and allow access to the attributes of each unit in a location-based way. Stage two refers to high-level visual-spatial processing and delivers the visual-spatial information for goal-directed actions. It consists of three parallel processing streams. One stream leads to entry-level object recognition, the second one to spatial-motor programming (“parameter specification”), and the third stream to the set-up of an object file. An object file contains high-level visual-spatial attributes (e.g., complex shape parts) and an index that allows a representation of the spatial-temporal continuity of an individual object. Stage two can maintain only up to four object files. Decisive for the version of the two-stage approach advocated here is the assumption that only one visual-spatial unit at a time is selected for access to stage two, that is, for recognition, spatial-motor programming, and object file set-up.

Visual-spatial working memory: Functions, architecture, and control of contents

“Working memory” is a central concept in understanding the neuro-cognitive abilities of humans and monkeys. It allows not only the temporary maintenance relevant information, but also its active organization and modification. Often-cited examples of postulated working memory function are: keeping a phone number in mind, constructing and maintaining a visual-spatial image, or comprehending a sentence. Experimental research over the last 20 years has revealed characteristics of working memory at the psychological and neurobiological level (e.g., Baddeley, 1986, 1992; Goldman-Rakic, 1987, 1996; Logie, 1995; Petrides, 1996). One basic characteristic is that two main working memory systems can be distinguished in humans, namely one

system for dealing with visual-spatial information and another system for dealing with phonological information (e.g., Baddeley, 1986, 1992; Logie, 1995). Functional imaging studies validated this distinction by showing that different brain areas are activated when these two working memory systems are involved (e.g., Jonides et al., 1996; Smith & Jonides, 1997). In this paper, visual-spatial working memory is the subject of interest.

What are the functions of visual-spatial working memory (VSWM)? Traditional behavioral working memory research (e.g., Baddeley, 1986; Logie, 1995) has focused, first, on the short-term retention of visual-spatial patterns that were removed from the current view and, second, on “manipulation” functions such as “mental imagery” operations. The modified two-stage conception advocated here implies that the short-term function is required not only when objects disappear from the sensory surface but also when the objects of the scene are still present at the sensory surface. The reason is the following: Only one visual-spatial unit can be selected for access to stage two. If the selection process shifts to another unit, the previously “attended” unit is without bottom-up support of activation flow; activation flow from the retina up to stage two is restricted to one attentionally gated unit at a time. In other words, only one object file is “on-line” in the sense that its attribute representations receive constant activation inflow from the retina. Other object files that were setup during a former attentional allocation are no longer supplied by bottom-up activation flow. A short-term retention function is required for these “off-line” files when they are to be maintained further within stage two. Therefore, the presence of objects on the retina is not sufficient for their representation at stage two these objects have to be “attended” and they have to be temporarily retained by the short-term function. This short-term retention is especially required for tasks that rely on the simultaneous availability of several objects, such as driving a car in a crowded traffic situation. Having visual-spatial information of just one car available would not be sufficient to master the task.

Architecture of visual-spatial working memory:
Relationship to sensory-based processing
and brain areas

How is VSWM, more precisely its short-term memory function, related to sensory-based visual-spatial processing that is conceptualized within the modified two-stage conception? It is suggested that VSWM refers to object files (see Luck & Vogel, 1997), including their spatial relationships (allocentric spatial information) within stage two.³ The part of stage two that contains up

³The so-called “iconic memory” (e.g., Neisser, 1967; Pashler, 1997), often considered a stage prior to VSWM, is assumed to be a part of stage one.

to four object files will be called “visual-spatial workspace”⁴ here (see Fig. 2). As stated above, during sensory-based visual-spatial processing of information within the workspace, one object file is on-line and up to three object files are off-line. On-line means that the attentional unit selection process delivers gated activation flow from stage one to stage two. Every change of a currently “attended” object in the world reflected on the retina (e.g., movement) is directly transferred into the “updating” of the on-line object file; this state should be called “sensory-based object file processing.” “Off-line” means that the object file is cut off from the retinal-based activation flow and has to rely on a temporary storage process within the workspace; this state refers to “short-term memory-based object file retention.” In summary, the current visual-spatial representation of a scene consists of one currently “attended” on-line object file and up to three stored off-line object files; as stated before, the overall capacity limit of the workspace seems to be four (e.g., Sperling, 1960; Shibuya & Bundesen, 1988; Irwin, 1992).

Now the issue will be addressed of how VSWM – more precisely, its object files within the workspace – is related to brain areas of primates. Neuropsychological and neurophysiological data suggest that high-level visual-spatial attributes within the workspace are located within the inferior-temporal and posterior parietal areas (see Zeki, 1993; Milner & Goodale, 1995, for overviews). For instance, lesions of the inferior-temporal cortex cause problems in shape and color perception as well as in the connected function of object recognition (see, e.g., Farah, 1990; Milner & Goodale, 1995). Where in the brain are the indices – central parts of objects files – located? It will be argued that indices are located within the posterior parietal cortex. This claim is based on a neuropsychologically defined Balint syndrome (see, e.g., Rafal, 1997). One symptom of a Balint patient is that he or she reports seeing only one object at a time. Lesions of the posterior parietal cortex in both hemispheres cause such a symptom. This symptom of seeing just one object at a time should be expected if the ability to set up an index, the core of an object file, is hampered. An index allows the short-term maintenance of an object when visual attention (unit selection) is withdrawn and the object file “goes off-line.” However, an index is not needed to represent and maintain the high-level visual-spatial attributes of the currently attended object. Its attentionally gated activation flow to stage two is sufficient to represent the attributes of an object there. Consequently, without indices and off-line object files, the representation of the world would be restricted to one currently attended object at a time. Therefore, the

Balint syndrome is taken here as evidence that indices are represented within the posterior parietal cortex.

Next, the question is posed whether VSWM can be restricted to temporal and parietal areas. Based on several lines of evidence, the answer has to be no. At least, frontal – especially prefrontal – areas should be involved in the short-term function of VSWM. One line of evidence comes from neuroimaging studies (see, e.g., Courtney, Ungerleider, Keil, & Haxby, 1997; Smith & Jolles, 1997) that consistently show increased activation of frontal areas when the short-term storage of visual and/or spatial patterns is required. Furthermore, a number of neurophysiological studies, especially single-cell recordings (e.g., Fuster & Alexander, 1971; Miller, Erikson, & Desimone, 1996), suggest a role of frontal areas in such short-term storage (see Fuster, 1997, for an overview). What is the role of these frontal areas? Based on the study by Rao, Rainer, & Miller (1997) that found binding cells in the lateral prefrontal cortex (PFC), it is suggested here that the PFC contains indices, too. These PFC indices are assumed to control parietal indices, that is, the object files; see also, for instance, Desimone, Miller, & Chelazzi (1994) for the idea that frontal areas control more posterior areas. More precisely, it is suggested that a loop between the (lateral) PFC indices and the parietal indices exists. One function of this loop might be to extend the lifetime of parietal indices; a “refreshment” process that increases the activation of an index (see below) might use this loop. If so, then elimination or deactivation of the prefrontal areas should reduce the ability to temporarily maintain visual-spatial material (see Fuster, 1997).

In summary, the short-term storage aspect of VSWM refers to the visual-spatial workspace of stage two. It contains one on-line and up to three off-line object files. Furthermore, attributes and indices of object files are assumed to be stored within inferior-temporal and posterior parietal areas, while prefrontal areas and their indices should exert a control function over these more posterior contents.

Control of visual-spatial workspace content

Temporally-based neural coding and competition of object files

How can up to four object files be represented within the workspace and how can they be distinguished and segregated? This question about segregation refers to one side of the so-called “binding problem” (see Schneider, 1995). Several solutions are conceivable with different degrees of neurophysiological plausibility. First, the workspace could contain representations of each attribute in each dimension, four times in order to allow separate representations of up to four files. For instance, if four yellow object files are set up, then four different and independent neural representations of the same color “yellow” should exist. Up to now, no evidence has

⁴The term “workspace” is introduced in order to distinguish the conception postulated here from two other postulated influential VSWM structures, namely Baddeley’s (1986, 1992) “visual-spatial sketch pad” and Logie’s (1995) “visual cache.” These structures imply neither the object-specific capacity limit of four nor the assumption of one on-line file and up to three off-line files, as well as other characteristics of VSWM specified below.

been found for such a coding scheme within the relevant areas of the inferior-temporal (IT) and posterior-parietal cortex (see, e.g., Tanaka, 1997, for IT). Second, object files within the workspace could be distinguished by four different levels of activation – more precisely, by four different firing rates. This is also not very plausible because it does not allow a representation of several object files that share one attribute, such as two yellow objects and one red object. Given just one (and not four) color representation of yellow, firing rate alone does not allow encoding of the bindings of “what-goes-with-what” in this case. Third, “temporal coding” at the neural level could be used for segregating several object files (see Milner, 1974; von der Malsburg, 1981; Singer, 1989; Goebel, 1991; Schneider, 1995; Singer et al., 1997). More precisely, the neural firing of object files is assumed to be segregated in time. For instance, all patterns of objects file x fire at time slice n ; all patterns of file y at slice $n + 1$, etc., (see, e.g., Goebel, 1991; Hummel & Biederman, 1992). Based on this consideration, it is suggested that the primate brain uses this temporal coding for representing up to four object files.

What are the implications of such a coding scheme? First, time slice segregation implies that the activation of each file depends on the number of currently segregated and maintained files. The more object files there are that have to be retained and segregated, the lower their individual activation values should be. One reason might be that the (synchronized) firing of a file is restricted by the length of its time slice. If there is just one maintained object file, it will fire within every time slice. However, if there are four files, they will fire in four different slices, which in turn reduces the firing rate of each file. A prediction of this conception is that the ability of the VSWM to maintain objects should decrease with the number of objects. The more objects have to be retained – the limit is probably four – the lower the individual neural activation of each file, and therefore, the less the probability of recalling files after a short-time retention period (e.g., a few seconds) – see also the later subsection on short-term forgetting.

How is the activation value of an object file as a whole represented? It is suggesting that the activation of an object file as a whole is represented by the activation of its index. Only as long as the index has sufficient activation is the file represented within the workspace and can be accessed. If the index falls below a critical activation value, the object file is lost, is “forgotten” (see next subsection). The activation of an index should be distinguished from the activation of file attributes. How are both activations related within an object file? The current hypothesis is that the highest attribute activation of a file determines the activation of a file determines the activation level of the index. For instance, if the task set specifies the color red, then the color red of an object file (that is actually red) will have relatively high activation level compared to the shape and location attribute. Consequently, the activation of the color red will determine the index activation level. If index activation

and activation of attributes of an object file fire in a synchronized (correlated) manner – which is required to set up temporary connections between the attribute and the index – then the assumption that the highest attribute activation will determine the index activation makes sense.

As stated before, the conception of a temporal segregation of object files implies that the activation level of each file depends on the number of stored files. The more slices there are that are occupied by different files, the less the activation of an individual file should be. However, such a “static” relationship between object file activation can only be part of the story; activation of object files change over time. Thus, a concept of dynamic competitive coupling is needed that specifies how the current activation level of each file changes over time dependent on the activation value of other files. The simplest assumption is that activation of files is directly coupled in an inhibitory manner. When the activation of one file increases, the activation of other files decreases (see, e.g. Phaf, Van der Heijden, & Hudson, 1990; Duncan, 1996). The higher the activation of one file is, the stronger the “inhibitory,” competitive influence on other files.⁵ An implication is that the set-up time for a new object file should depend on the number and activation level of already stored off-line files. If several files are stored, then a strong competitive influence on the currently setup file should be exerted. Therefore, the activation of the currently set up file should rise more slowly compared to a case where no files are stored and no competition is exerted. Consequently, the set-up time of an object – until a sufficient activation level for temporary maintenance is reached – should depend on the number of stored files. Data from a study by Shibuya & Bundesen (1988) deliver confirmative evidence for this claim. In one experimental condition, briefly presented and masked target stimuli (digits) were to be reported. The SOA between stimulus display and the mask varied trialwise. The data show that a 35-ms presentation time of the display was sufficient for the reporting of one stimulus, 70 ms for two stimuli, and 150 ms for three stimuli, which means that there is a non-linear increase of required presentation time with the number of reported stimuli – a result that is to be expected if the set-up time for each object file increases with the number of stored files.

Short-term maintenance and forgetting object files

The main questions of this subsection are: How are object files maintained over time? How does their activation value behave over time? How does short-term forgetting occur? First, as stated before, object file

⁵ How the conception of time slice segregation fits the suggestion of dynamic competitive coupling cannot be properly stated within a pure verbally articulated theory; neural network simulation studies are required for this purpose.

existence depends on its activation value. If an object file falls below a critical activation value, it is lost from workspace and therefore forgotten. Only long-term storage is able to retain the file content for a longer time scale. Second, the concept of dynamic competitive coupling introduced above implies that during the set-up of a new file, its activation rises and, consequently, the activation of other stored files decreases. Third, what does dynamic competitive coupling imply about forgetting during a retention interval? If no further object file should be set-up during the interval, no activation reduction for the stored files should occur. However, this would presuppose that attentional unit selection could be shut off during the retention interval. Given that the eyes are usually open during this interval (e.g., several seconds), this does not seem to be a plausible option. Furthermore, the visual-spatial system of primates should have a built-in bias to scan the current environment permanently by attentional processes. Given this bias, the system should not be able to shutoff the visual attentional process of unit selection for a few seconds. If so, then each newly attended object (unit) during the retention interval will imply the set-up of the corresponding object file within the workspace and therefore competitive interaction with stored files. The set up of a new object file will decrease the activations of the stored files, so that forgetting will occur during the retention interval; additionally, there is the limit of four files within the workspace. Fourth, the offered explanation of short-term forgetting stands in a tradition of “interference” theories of forgetting⁶ (see, e.g., Baddeley, 1990). This kind of explanation assumes that forgetting occurs because memory traces of items disrupt each other. A special role is given to items that were maintained during former trials.

Finally, a number of mixed issues about forgetting and maintenance should be addressed. First, what happens if the workspace is filled with four object files and a new object file is set up by the unit selection process? An implication of the concept of competitive coupling is that the file with the lowest activation is replaced and “thrown out” of the workspace. Second, how does the task set influence the maintenance of object files? As already stated for the control of attentional processes such as unit selection, the task set influences the processing by the application of extra activation to those visual-spatial attributes that are specified within the task set. For instance, if red letters are to be reported among black digits, then the “color red” and the “category letter” is supplied with extra activation at the workspace level. When the unit selection process has actually selected a red letter (e.g., by the serial scanning mechanism), this task-dependent extra activation to the

attributes “color red” and “category letter” flows via the temporary connection directly to the index and causes an additional activation increase. Therefore, the index of a task-set compatible file will have a higher activation value than the index of a file that does not fit the task set. Thus, task-set compatible object files should have a longer short-term storage duration and should be better protected against replacement than files that are not specified by the set.⁷

In summary, short-term forgetting is assumed to occur due to activation-based competition of object files. If the activation of one file increases, the activation of other files simultaneously decreases; therefore, a version of the “interference” explanation of forgetting is postulated here. Furthermore, task-relevant object files reach a higher activation value than non-relevant files.

Output streams to action, an index activation increase process, and refreshment of object files

How are the object files within workspace related to “output” processing for action, that is, to spatial motor programming and to the recognition of objects and scenes? Object files are represented by attribute and index patterns of a sufficient activation level within the workspace. The output activation flows of these files are distinguished by their activation level. What does this output activation flow imply for spatial-motor programming and for object and scene recognition?

First, for the system of programming spatial-motor actions such as grasping or saccading, it implies that up to four output streams coexist, distinguished by time slices and activation level. For systems such as the saccadic eye movement system that can only move to one object at a time, a selection mechanism is needed in order to guarantee that the location parameter of only one file accesses the low-level motor structure (for controlling the muscles). A simple winner-take-all mechanism (e.g., within the superior colliculus; see Dominey & Arbib, 1992) might be sufficient for such a selection mechanism; among the segregated activation streams it would select the stream of the object file with the highest activation. For spatial-motor systems such as the grasping system that allows a manipulation of more than one object (e.g., by two arms), it is possible that spatial parameters of up to four files might be used by motor systems.

Second, the output activation flow of up to four files for object and scene recognition allows the possibility of recognizing a complex object or scene by their decomposition into up to four parts (“sub-objects” or “geons”; see Biederman, 1987; Hummel & Biederman, 1992), including

⁶ Besides interference, some researchers have suggested passive decay as an explanation for forgetting (Baddeley, 1990). By postulating the interference mechanisms of dynamic competitive coupling, it is not claimed that decay of activation does not exist. Current data do not allow us to rule this possibility.

⁷ If an object does not fit the task set in all aspects, it could nevertheless be supplied with extra activation. For instance, in a “conjunction search” task (see, e.g., Treisman & Gelade, 1980), distractors are not compatible with the whole task set, but one attribute usually fits. Thus, distractors should be supplied with extra activation of the attribute that fits.

their spatial relationships. In such cases, the sub-objects build a “chunk” or “supra-object.” Hummel & Biederman have shown that a connectionist model including time slice segregation of up to four object parts can be successfully used to simulate object recognition data.

Third, once an object or a scene is recognized, programming of the corresponding phonological patterns for verbal report might take place, if the recognized patterns are available for a sufficient time. However, programming phonological patterns or spatial-motor patterns does not lead in an obligatory way to overt execution; something like a “go-signal” controls the conversion into overt action (e.g., Bullock & Grossberg, 1988; Schneider, 1995). Moreover, complex visual-spatial action planning processes also rely on information about several objects and need an output activation flow of several object files from the workspace – for instance, imaging an offense play in a soccer game in which it has to be specified where the ball should be kicked.

In summary, up to four files have an output activation flow to action-related structures such as the recognition machinery (followed by a verbal report) and the spatial-motor programming machinery. These files are distinguished not only by different time slices but also by different activation levels. When the files form a supra-object consisting of sub-objects, then the output stream of these chunked files is used for recognition of the overall configuration. When files cannot be chunked, then the file with the highest activation will dominate the output processing – that is, it will generate spatial-motor programs and a recognition-based verbal report pattern.

As stated above, when object files within the workspace cannot be chunked into a supra-object, then the file with the highest activation controls the current action output. How can these maintained files with a lower activation level be “brought back to action control” so that they deliver the parameters for spatial-motor programming and recognition (verbal report)? Given that object files have indices, an obvious suggestion is that the index activation of the file determining the action should be increased up to the highest level. Once the index of the file is supplied with extra activation, this activation flows via the temporary connections to its attributes and increases their activation. The process of supplying the index with extra activation is called here the “index activation increase process.” Only one index at a time should receive the extra activation – otherwise, several files with a similar output activation flow level would emerge with would not be functional for action control.

How does refreshment fit into this picture? Refreshment is a process that allows the prevention or delay of the forgetting of short-term stored information. The basic suggestion is that refreshment is realized by the index activation increase process. The index selection process increases the index activation up to the relatively highest level so that the storage duration of the corresponding file is increased. The higher the activation of a file is, the less are the chances that a newly set up file will

lead – via competitive coupling (interference) – to forgetting, that is, elimination of the stored file from the workspace. Furthermore, the result of the refreshment process is not the same for all attributes of a file and depends strongly on the current refreshment task. For instance, if locations of objects should be temporarily maintained, then this task-dependent extra activation to the dimension “location” will increase the activation level of the location attribute, but not of other attributes. Task-dependent activation differences found in neuroimaging studies that require either the short-term retention of spatial or visual attributes (see, e.g., Joindes et al., 1996) are assumed to reflect such a task-specific object file set-up and refreshment processes. In each task condition (spatial or visual), a whole object file is maintained, but the activation level of attributes should differ considerably, depending on the task demand.

Executive control processes and their brain sites

In this subsection the issue of “executive control operations” within VSWM should be briefly touched upon – operations responsible for the “working” (or “manipulatory”) part of working memory, in addition to its short-term storage function. What is meant here by executive operations within VSWM? Executive operations are viewed as optional operations that change the content of the visual-spatial workspace, namely, the object files. One class of executive operations refers to specialized processes for modifying the contents of object files. Another class refers to processes that allow the generation of new object files, that is files that are not based on sensory input via the unit selection process. Examples of the modification type of executive control operations are all kinds of sub-processes that are involved in “mental imagery” and “mental rotation” tasks (see, e.g., Kosslyn & Koenig, 1992). The index activation increase process (implying refreshment) is also an example of an executive process of the modification type; it modifies the activation level of the object file. An example of an executive operation of the type “generation of new information” within the workspace refers to the antisaccade task (see, e.g. Guitton Buchtel, & Douglas, 1985). In such a task, the saccade target location is not given by the sensory stimulus but has to be generated according to the instructions (“Saccade to the opposite side of the sensory stimulus”). Therefore, the antisaccade target location has to be generated internally according to this rule. “Inhibition of prepotent responses; which is also involved in the antisaccade task (“Do not make a saccade to the suddenly appearing sensory stimulus”), might be a further executive process (see, e.g., Roberts, Hager, & Heron, 1994); however, it is not clear whether this kind of inhibition uses the visual-spatial workspace and can thus be considered an executive operation of VSWM.

If modality-specific executive operations within the visual-spatial domain exist – for instance, visual execu-

tive processes (e.g., those involved in changing the shape of an object, see Kosslyn and Koenig, 1992) or spatial executive processes (e.g., those involved in generating an antisaccade target location) – then these operations should probably be hosted in different brain areas. If this assumption of modality-specific executive processes in different brain areas is combined with the formerly made assumption that indices are also represented within certain parts of the prefrontal cortex (PFC), then a new look at the current issue of how the frontal lobe is organized can be taken (see, e.g., Rushworth & Owen, 1998, for an overview). Basically, two positions in this issue can be contrasted: Goldman-Rakic (1996) postulates a modality-specific organization of working memory processes in the dorso- and ventro-lateral prefrontal cortex, namely a segregation of “what”-aspects (visual attributes such as color and shape) from “where”-aspects (spatial attributes). Petrides (1996) and Owen (1997), on the other hand, argue against such a modality-specific organization and instead propose two types of executive processes (working memory processes) in different areas of the prefrontal cortex: first, processes for the temporary maintenance and organization of information, and second, processes for the active manipulation and monitoring of information. The position taken here is that the short-term representation of visual and spatial attributes is located within posterior areas such as IT and PPC but not within the PFC, in line with Petrides and Owen and against Goldman-Rakic. However, the PFC is assumed to contain indices (see Rao et al., 1997) and areas for specialized modality-specific executive operations – in this sense, the segregation of “what” and “where” is continued – in line with Goldman-Rakic and against Petrides and Owen.

The last remark on executive processes suggests a distinction between first- and second-order executive processes. First-order executive processes refer to the modality-specific executive operations that have been described above – operations that modify visual and/or spatial attributes of object files and that also allow the generation of new information. Second-order executive processes determine and shift the current task set. For instance, when in trial n red letters are to be reported and trial $n+1$ requires subjects to report black digits, then preparation and maybe execution of trial $n+1$ requires the second-order executive operation of task set shifting; see Allport, Styles, & Hsieh (1994) and Monsell (1996) for experiments on this type of control process. Both levels of executive operations, first- and second-order, are essential for understanding the control of visual-spatial processing inside and outside working memory.

A new look at the attentional blink, backward masking, change blindness, and transsaccadic memory

In this section, recent experimental data on “attention” and “scene perception” will be discussed from the perspective of the theory specified up to now. First, a new

look at the “attentional blink” (AB) will be taken. Basically, it is suggested that AB arises because Target 1 is stored as an object file with a high activation level within VSWM and thus competes strongly with the set-up of Target 2. Consequently, the activation of Target 2 does not rise to a level that would allow it to survive the competition of the item following Target 2. Second, “backward masking” and “dwell time effects” are explained in a similar fashion. Finally, the experimentally investigated phenomena of “transsaccadic memory” and “change blindness” are reconsidered. A modified view on transsaccadic memory is given that emphasized the storage of detailed high-level visual-spatial information in the form of a few object files.

The attentional blink

The attentional blink (Raymond, Shapiro, & Arnell, 1992) refers to an experimental situation where two target objects have to be reported⁸ within a rapid serial visual presentation (RSVP) of distractor objects. For instance, the subject is instructed to report two target letters among distractor digits (see, e.g., Chun & Potter, 1995). Usually, each item (distractor or target) is presented for a short duration (e.g., 100 ms), replaced at the same position by the next item, and so forth. The basic finding is that the first target (T1) can be reported with a relatively high level of accuracy (e.g., 90%). However, the second target (T2) is only reported with low accuracy when it appears in close temporal succession after T1 (e.g., as a third item after the T1 item). This failure to report T2 is called the attentional blink (AB).

What does the modified two-stage theory including VSWM state about an experimental situation that generates an AB? Each time an item appears within an RSVP stream it accesses stage two, that is, recognition and object file set-up starts (spatial-motor programming, too, but this is not relevant here). Due to the short time an item is available before the next item appears (around 100 ms) – either on the screen or in “iconic memory” after its disappearance – and due to competition between object files, the newly set-up object file (index) of each item reaches only a relatively low activation level. Therefore, when T1 is set up as an object file in such an RSVP stream, its activation level will be low, too. Low-level activation of T1 implies the danger that T1 will be “thrown out” of VSWM. This danger of elimination from VSWM is there because the current activation of T1 is permanently reduced (in the end, below the activation threshold) by the ongoing competition of items following T1 (interference). There is only one option to maintain T1 within such an RSVP stream, namely the

⁸ In some experiments (e.g., Raymond et al., 1992), the second target (often called “probe”) does not have to be reported, but only its presence or absence has to be signalled. For instance, subjects have to tell whether the letter “X” (second target) was present or absent. However, in both cases – report and detection – identification of the second target is required.

index activation increase process (refreshment). This executive operation increases T1 activation up to a level that allows it to be retained within the workspace until the end of the trial.

However, there is a cost for this T1 activation increase. The high activation of the object file of T1 implies a strong competitive influence on T2 during its set-up (see the section on “object file forgetting”). When T2 is set up as an object file, its activation value will rise slowly due to this competitive coupling with other files, especially with T1. Consequently, T2 will reach a lower activation value than T1 that was not subject to such strong competition of stored object files with high activation. The stronger the activation of T1 is, the stronger the competition and the lower the final activation level of T2 will be. A very low activation value of T2 – lower than T1 before refreshment – implies that (due to competitive coupling) an item following T2 will reduce the already very low activation of T2 to a level below the activation threshold of the workspace – in other words, the item which follows T2 causes its elimination from VSWM.

There are a number of basic observations on the AB. Each of the observations should first be described, then explained, and finally supported by prediction. Some of the unanswered questions about the AB recently formulated by Shapiro, Arnell, & Raymond (1997, p. 295) will also be addressed.

First, T2 performance strongly depends on the relative serial position after T1. If T2 appears immediately after T1, many studies show good performance for T2, that is, no AB. T2 performance, however, shows a relatively strong drop when there is one distractor item between T1 and T2. This drop is often most pronounced when two distractors appear between T1 and T2. After the second serial position of T2, performance often improves continuously with the number of intervening items between T1 and T2 until the AB is eliminated⁹ after 5–8 intervening items, that is, after 500–800 ms (see, e.g. Raymond et al., 1992; Chun & Potter, 1995; Shapiro, Driver, Ward, & Sorensen, 1997; Shapiro, Arnell, & Raymond, 1997). Why is AB reduced when more items intervene between T1 and T2? The more items intervene between T1 and T2, the more T1 activation should be reduced, due to the competitive influence of these intervening items. Lower T1 activation implies a less strong competitive influence on T2, so that T2 will rise to a higher final activation value compared to the case of T1 with high activation. This relatively high final level of T2 prevents the post-T2 item from reducing the T2 activation to a sub-threshold level.

Second, why is T2 in some studies reported with a minimal loss (no AB) when T2 appears immediately after T1? The initiation of the refreshment process of T1 takes time. Consequently, items immediately following

T1 enter stage two and are set up as object files as long as the refreshment process is not initiated.

Third, a well-established observation about the AB is that T2 items that cannot be reported are nevertheless recognized (“identified”) at a level prior to report. A study by Luck, Vogel, & Shapiro (1996) shows that electrophysiological activity associated with the processing of the meaning of T2 can be observed even when T2 cannot be reported (an AB situation). Making the same point, priming effects of T2 that could not be reported have been observed by Shapiro, Driver, et al. (1997). The explanation of the AB given above implies that implicit recognition but no report can be observed. T2 is set up in each trial as an object file, and recognition begins in parallel. The AB arises because the item following T2 reduces its already low activation to a sub-threshold level, so that T2 cannot be reported. Nevertheless, the recognition result was available for a few tens of milliseconds – namely, during the presentation time – and this was sufficient for implicit processing, as observed by Luck et al. (1996) and Shapiro, Driver et al. (1997).

Fourth, AB depends on target distractor discriminability. For instance, if targets are letters and distractors are keyboard symbols (e.g., > or *) instead of digits (Chun & Potter, 1995, Exp. 5), then AB will be strongly reduced. Keyboard symbols have a much lower featural overlap with the target description (letters) than digits. Therefore, keyboard symbols should hardly be supplied with extra task-specific activation at the workspace activation at the workspace level. Therefore, due to their relatively low activation value, they will be weak competitors compared to digit distractors; digits have some overlap with the target description (letters) and should therefore receive some task-specific extra activation.

What predictions follow from the explanation of the AB made here? First, the duration of T2 should be critical for the amount of AB. Increasing T2 duration by 100 ms (resulting in 200 ms overall presentation) should eliminate a 500–800 ms AB almost completely. A time of 200 ms for T2 should be enough to set up an object file at a sufficient activation level to protect it against a following item; evidence for the important role of the item following T2 can be found in a recent study by Giesbrecht & DiLollo (1998). They have shown that “interruption masking” of the T2+1 item is more effective than “integration masking” (see, e.g., Scheerer, 1973; and the next subsection on this issue). Second, the presentation time for each item (RSVP rate) should also be critical for the AB. Decreasing the rate should strongly reduce the AB. Third, the magnitude of AB should increase with the number of targets that have to be reported. Three targets should cause a stronger AB than two targets: this has already been shown by Chun & Potter (1995, Exp. 2). The AB should be even stronger when four targets have to be reported, which should still be within the capacity limits of VSWM. Fourth, a primary task that requires the visual-spatial material in working memory to be retained should strongly

⁹ Elimination of the AB occurs when T2 performance is the same as within a condition when only T2 (and no T1) has to be reported.

influence the report of a single target in an RSVP stream (secondary task). The more object items have to be retained, the stronger their competitive influence on the set-up of a new file, and thus the lower the report performance for the RSVP target should be.

Backward masking and dwell time effects

Backward masking refers to a situation when a target stimulus that should be reported is followed by another pattern stimulus (the mask) at the same location within a short interval. If the SOA between target and mask is sufficiently short (e.g., 40 ms), then the target will only be reported with a low probability compared to a condition with a relatively long interval or no mask at all ("interruption masking," see Scheerer, 1973). If the mask is represented simultaneously with the target ("integration masking"), a much weaker performance drop occurs compared to the interruption masking condition (see, e.g., Scheerer, 1973; Breitmeyer, 1984; Giesbrecht & DiLollo, 1998).

Basically, the explanation for backward masking is the same as for the AB. Due to the short presentation time to the target, only a low level of activation will be reached for the target. The immediately following mask reduces the already low activation of the target to a sub-threshold level. A failure of masking occurs if sufficient activation is still present when the refreshment process is initiated. Integration masking is less effective because the workspace does not treat the input of target and mask as two successive events and thus does not set up two files that compete. Instead, it attempts to recognize one object based on the degraded input.

Recently, Duncan, Ward, & Shapiro (1994) introduced an experimental paradigm – the "dwell time" paradigm – that also generated a failure to report a second target. Basically, two target stimuli were presented at different locations, and each target was backward masked (presentation time: 45–60 ms). The SOA between the two targets was varied. Similar to results of AB experiments, Duncan et al. (1994) found that the ability to identify the second target (T2) was strongly impaired after successful identification of the first target (T1) when the SOA between T1 and T2 was within the range of 0–600 ms. Again, the suggested explanation of the so-called "attentional dwell time" effect is analogous to the explanation of AB and backward masking. T1 is presented and followed by the mask, which reduces the already low activation of T1. Therefore, the refreshment process (index-activation-increase process) is required in order to prevent the forgetting of T1 by increasing its activation. Due to the increased T1 activation and its competitive interaction with T2, the final activation level of T2 is lower than the final activation level of T1 (prior to refreshment). When the T2-mask appears, the very low activation of T2 is reduced even further (close to the threshold), which in turn strongly reduces the chance

that the refreshment process can work on existing activation patterns of T2.

A prediction for the suggested "dwell time" effect is that it should be strongly reduced when the SOA between T2 and the mask is increased (e.g., from 45–60 up to 120 ms). This manipulation should decrease the dwell time of several hundred milliseconds to less than a hundred. Furthermore, the claim of Duncan et al. (1994) that these data reflect attentional dwell time of several hundred milliseconds, which is inconsistent with fast serial scanning models of visual attention, has to be reconsidered. The model presented here includes a fast serial scanning mechanism (unit selection), and dwell time effects are explained by a second component of the system, namely, the activation-based competition between object files.

Change blindness and transsaccadic memory

This subsection presents the implications of the theory presented in this paper for "transsaccadic memory: (see, e.g., Irwin, 1992, 1996; Pollatsek & Raynor, 1992; Henderson, 1994) and "change blindness" (see, e.g., Simons & Levin, 1997). What is the content of the transsaccadic visual-spatial scene representation within the primate brain? "Transsaccadic" refers to the content of the scene representation across several eye fixations. Based on the modified two-stage theory, the evident claim is made that the transsaccadic scene representation refers to the visual-spatial workspace representations of stage two. This representation consists of representations of objects in the form of object files. Up to four files are assumed to be maintained within the visual-spatial workspace. Given this conception, there is no real difference between a scene representation of a single fixation and a transsaccadic scene representation referring to several fixations. In each case, only one object file is on-line and up to three files are off-line. Only the on-line file is supplied with current bottom-up information from the retina, while the off-line files rely on short-term maintenance processes.

How are attentional processes, eye movements, and transsaccadic scene representation process related? Let us begin the analysis with a situation when a person has just opened his/her eyes and begins to explore the multi-object environment in front of him/her. Immediately after the eye opening the unit selection process modifies the activation flow of one visual-spatial unit, and this unit is set up as a first object file in workspace. Next, a second unit is selected and the file of the first unit is shifted towards the short-term storage mode (stored object file). After this selection of a second unit, a third unit is selected. Let us assume that the unit is presented far in the periphery of the fixation. Due to this peripheral presentation, the selected information of the third unit is not sufficient for recognition. Therefore, the decision is made to improve the retinal resolution by making a saccadic eye movement towards the third unit.

This requires the unit selection process to be allocated to the third unit for saccade programming (see e.g., Deubel & Schneider, 1996). During the saccade, the transsaccadic scene representation within the visual-spatial workspace does not change. After the eye has landed, the object file of the saccade target is still represented. However, this postsaccadic file will be updated by the unit selection process that is allocated to the expected position of the saccade target immediately after the saccade. If the saccade target object is at the expected position, then the presaccadically computed object file content (shape, color, etc.) will be replaced by the new postsaccadic input, and recognition occurs due to the improved retinal resolution. Next, either the other stored presaccadic object files are updated via the unit selection process, or new units are selected. Which of both options is realized depends on the current task. If the task requires a scene to be inspected for changes (see next section), then presaccadic object files have to be updated. If a visual search task has to be carried out and none of the presaccadically selected units was recognized as a target, then new postsaccadic units will be selected.

A number of recent studies have shown our inability to detect dramatic changes within natural or artificial scenes (see Simons, 1996; Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1997). This change blindness occurs, for instance, when scene changes (e.g., the color of an object) are made while the eye is moving (intrasaccadic changes: e.g., McConkie & Currie, 1996; Deubel, Schneider, & Bridgeman, 1996). However, change blindness also occurs when natural scenes are viewed and changed during fixation (see Simons, 1996; Rensink, O'Regan, & Clark, 1996; Simons & Levin, 1997). What does the modified two-stage approach imply for this phenomenon of change blindness? First, the current scene representation cannot contain more than up to four object files; therefore, one cannot expect changes of all objects (e.g., shape change) to be detected, if the environment contains many objects. If for some reason an object has been attended, that is, set up as an object file, then a modification of the corresponding object could be detected during the updating process of its object file. Second, one prerequisite for change blindness is that the object attribute change (e.g., in color) is not marked by a unique transient. For instance, if the colors of all objects stay the same and only one object changes its color, then this will easily be detected. The reason for this is that such a color transient attracts the unit selection process involuntarily. Therefore, scene changes are often investigated with the "flicker" paradigm (Rensink, O'Regan, & Clark, 1996) that introduces a masking blank field between alternating displays of a scene without a change and with a change.

Furthermore, change detection across the saccade is a special case. For instance, let us consider an experimental situation in which the location of the saccade target is shifted during a saccade (e.g., 10% of the saccade amplitude while the direction is maintained). Subjects usually show a low performance level for detecting

this location change (see, e.g., Deubel, Schneider, & Bridgeman, 1996). Within a fixation, such a location shift is easily detected. What is the reason for this saccade-related change blindness? First, the detection of a location change should be easy if a presaccadic object file (which is set up parallel to the programming of the saccade location) is updated by the new postsaccadic information. The only thing needed as a mechanism that monitors the updating and signals the change. However, it is known that a "saccadic suppression" mechanism acts shortly before, during, and immediately after the saccade. As implied by its name, this mechanism suppresses the processing of visual-spatial contents, especially of magnocellular (dorsal) processing contents such as motion and location (see, e.g., Burr, Morrone, & Ross, 1994). Therefore, if an object change is made during the saccade – for instance, its location is shifted – then this saccadic suppression mechanism should prevent computation of this change during the object file updating phase – that is, usually shortly after the eye has landed and the saccade target is attended again (by the unit selection process).

If this explanation is correct, then changes that are made outside the saccadic suppression range should be detected at a much better performance level, which is indeed the case, as Deubel & Schneider (1996) and Deubel, Schneider, & Bridgeman (1996) have recently shown. The basic experimental paradigm requires a saccade to a stimulus in the periphery. The stimulus, the saccade target, was removed during the saccade. It reappeared modified, that is, at a shifted location, after a postsaccadic blank (e.g., 250 ms after the end of the saccade – outside saccadic suppression). Therefore, the eye had already landed when the changed saccade target appeared after this blank. At first glance, this blanking manipulation has an astonishing result. The ability to detect intrasaccadic change (here, a location change) is dramatically increased compared to the no-blank case. The reason for this difference is that the blanking manipulation allows an update of the presaccadic object file outside the saccadic suppression range so that the change detection mechanism can do its job. In further experiments (Schneider & Deubel, 1996) it has been shown that this blanking effect holds only for certain visual attributes (e.g., location or size changes) and not for others (e.g., color). This is not surprising, because saccadic suppression works most effectively for detecting "magnocellular" changes such as location or size (looming).

The view of transsaccadic memory advocated here stands in contrast to the currently dominating view (see, e.g., Irwin, 1996; Henderson, 1997) which assumes "that there is no internal, highly detailed, spatially integrated stimulus representation" (Irwin, 1996, p. 96; see also Pollatsek & Raynor, 1992; Henderson, 1997) in transsaccadic memory. Evidence for this claim is data such as mentioned above, namely, the inability to detect large location changes across the saccade or other changes in the visual appearance of objects (e.g., a change from

upper to lower case during reading; see Irwin, 1996, for an overview on these data). The above reported data of the blanking manipulation by Deubel, Schneider, & Bridgeman (1996) that indeed show the storage of detailed information are problematic for the dominant view. How can these conflicting data be reconciled, and how should the dominant view on transsaccadic memory be modified?

First, the dominant view does not have to be changed in reference to the claim that not all information about the scene is stored in a detailed way across saccades. However, it should be corrected in reference to the short-term storage of single objects; their information is assumed to be stored with high precision within object files. The inability to detect these dramatic changes across a saccade should not be interpreted as a failure of veridical storage of presaccadic contents (the dominant view). Instead, a failure to compute the changes between pre- and postsaccadic contents during the updating process due to the saccadic suppression mechanism is the suggested cause. This mechanism explains why changes within object files (e.g., a location change) that can be easily detected within fixation cannot be detected across the saccade.

Outlook

Finally, a brief evaluation of the suggested modified two-stage theory proposed should be given. First of all, it has to be conceded that many important questions about VSWM, scene representation, and attention were not addressed. Referring to attentional control, many issues were not touched, for instance, how task-set and sensory-driven factors control attentional processes such as unit selection (see, e.g., Yantis, 1998, for a recent overview). Referring to visual-spatial working memory, the same statement has to be made. For instance, only a very vague conception of “executive processes” was given that are responsible for the “working part” of VSWM. Furthermore, it has been left open how the spatial-motor system is related to the current contents of VSWM, that is, to the object files within the workspace. Can the spatial-motor system configure programs for action with several objects without involving the workspace? Also, referring to scene perception, a number of issues were not addressed, for instance, how stored knowledge (e.g., in the form schemas) influences the scene representation (see, e.g., Biederman, 1981; Intraub, 1997).

The decisive plus of the modified two-stage theory is that it attempts to be explicit at the “mechanistic level” within one conceptual framework about what is meant by such terms as attention, visual-spatial working memory, and scene representation, and how these terms are related. If it turns out that most of the suggested theoretical concepts (e.g., competition between object files or index-based access of object files) were wrong but a number of new and interesting experimental results were generated, then the main purpose of this manuscript has been fulfilled.

Acknowledgements I wish to thank Claus Bundesen, Heiner Deubel, Leonardo Chelazzi, Rudi Kerschreiter, Andy Mojzisch, and an anonymous reviewer for helpful comments, as well as Wolf Singer and his group for constructive criticisms on my binding conception, and, finally, Alexandra Tins for her help in improving the English of this paper.

This research was supported by the Deutsche Forschungsgemeinschaft (Forschergruppe “Wahrnehmungsplastizität,” PR 118/19-1, Teilprojekt “Aufgabenabhängiges Sehen”).

References

- Allport, A., Styles, E.A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance: XV. Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: Bradford, MIT Press.
- Allport, D.A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A.F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395–419). Hillsdale, NJ: Erlbaum.
- Allport, D.A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In D.E. Meyer & S. Komblum (Eds.), *Attention and performance: XIV. Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 183–218). Cambridge, MA: MIT Press.
- Baddeley, A. (1986). *Working memory*. Oxford: Clarendon Press.
- Baddeley, A. (1990). *Human memory: Theory and practice*. Hove and London: Erlbaum.
- Baddeley, A. (1992). Is Working Memory Working? The Fifteenth Bartlett Lecture. *The Quarterly Journal of Experimental Psychology, 44A (1)*, 1–31.
- Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J.R. Pomerantz (Eds.), *Perceptual organization* (pp. 213–253). Hillsdale, NJ: Erlbaum.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115–147.
- Breitmeyer, B.G. (1984). *Visual masking: An interactive approach*. New York: Oxford UP.
- Bridgeman, B., Van der Heijden, A.H.C., & Velichkovsky, B.M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences, 17*, 247–292.
- Bullock, D., & Grossberg, S. (1988). Neural dynamics of planned arm movements: Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review, 95*, 49–90.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523–547.
- Burr, D.C., Morrone, M.C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature, 371*, 511–513.
- Chun, M.M., & Potter, M.C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 109–127.
- Courtney, S.M., Ungerleider, L.G., Keil, K., & Haxby, J.V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature, 386*, 608–611.
- Crick, F. (1994). *The astonishing hypothesis: The scientific search for the soul*. New York: Maxwell Macmillan.
- Desimone, R., Miller, E.K., & Chelazzi, L. (1994). The interaction of neural systems for attention and memory. In C. Koch & J.L. Davis (Eds.), *Large-scale neuronal theories of the brain* (pp. 75–91). Cambridge, MA: MIT Press.
- Deubel, H., & Schneider, W.X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*, 1827–1837.
- Deubel, H., Schneider, W.X., & Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research, 36*, 985–996.

- Deubel, H., Schneider, W.X., & Paprotta, I. (1998). Selective dorsal and ventral processing: Evidence for a common attentional mechanism in reaching and perception. *Visual Cognition*, 5, 81–107.
- DeYoe, E.A., & van Essen, D.C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, 11, 219–226.
- Dominey, P.F., & Arbib, A. (1992). A cortico-subcortical model for generation of spatially accurate sequential saccades. *Cerebral Cortex*, 2, 155–172.
- Duncan, J. (1996). Cooperating brain systems in selective perception and action. In T. Inui & J.L. McClelland (Eds.), *Attention and performance: XVI. Information integration in perception and communication* (pp. 549–578). Cambridge, MA: MIT Press.
- Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*, 7, 255–261.
- Duncan, J., Ward, R., & Shapiro, K.L. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, 369, 313–315.
- Farah, M.J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision*. Cambridge, MA: MIT Press.
- Felleman, D., & Van Essen, D. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1–47.
- Fuster, J.M. (1997). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe*. Philadelphia, PA: Lippincott-Raven.
- Fuster, J.M., & Alexander, G.E. (1971). Neuron activity related to short-term memory. *Science*, 173, 652–654.
- Giesbrecht, B.L., & DiLollo (1998). Masking in the attentional blink: Decoupling integration- from interruption masking. Manuscript submitted for publication.
- Goebel, R. (1991). The role of visual perception, selective attention, and short term memory for symbol manipulation: A neural network model that learns to evaluate simple LISP expressions. In K.F. Wender, F. Schmalhofer, & H.D. Boecker (Eds.). *Cognition and computer programming*. University of Braunschweig, Department of Psychology: Ablex Publishing Corporation.
- Goldman-Rakic, P.S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology* (Vol. 5, pp. 373–417). Bethesda, MD: American Physiological Society.
- Goldman-Rakic, P.S. (1996). Regional and cellular fractionation of working memory. *Proceedings of the National Academy of Sciences, USA*, 93, 13473–13480.
- Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception & Psychophysics*, 55, 48–121.
- Grossberg, S., Mingolla, E., & Ross, W.D. (1994). A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations. *Psychological Review*, 101, 470–489.
- Guitton, D., Bachtel, H.A., & Douglas, R.M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research*, 58, 455–472.
- Hebb, D.O. (1949). *Organization of behavior*. New York, NY: John Wiley & Sons.
- Henderson, J.M. (1994). Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General*, 123, 410–426.
- Henderson, J.M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, 8, 51–55.
- Hubel, D.H., & Wiesel, T.;N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London*, 198, 1–59.
- Hummel, J.E., & Biedermann, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, 99, 480–517.
- Intraub, H. (1997). The representation of visual scenes. *Trends in Cognitive Sciences*, 1, 217–222.
- Irwin, D.E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 307–317.
- Irwin, D.E. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science*, 5, 94–100.
- Irwin, D.E., & Andrews, R.V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J.L. McClelland (Eds.), *Attention and performance: XVI. Information integration in perception and communication* (pp. 125–155). Cambridge, MA: MIT Press.
- Jonides, J., Reuter-Lorenz, P.A., Smith, E.E., Awh, E., Bames, L.L., Drain, M., Glass, J., Lauber, E.J., Patalano, A.L., & Schumacher, E.H. (1996). Verbal and spatial working memory in humans. In D.L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 35, pp. 43–88). San Diego: Academic Press.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kahneman, D., & Treisman, A.M. (1984). Changing views of attention and automatically. In R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention* (pp. 28–61). New York: Academic Press.
- Kahneman, D., Treisman, A.M., & Gibbs, B.J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219.
- Kosslyn, S.M., & Koenig, O. (1992). *Wet mind: The new cognitive neuroscience*. New York, NY: The Free Press.
- Leslie, A.M., Xu, F., Tremoulet, P.D., & Scholl, B.J. (1988). Indexing and the object concept: Developing “what” and “where” systems. *Trends in Cognitive Sciences*, 2, 10–18.
- Livingstone, M.S., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240, 740–749.
- Logan, G.D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, 103, 603–649.
- Logie, R.H. (1995). Visuo-spatial working memory. In R.H. Logie (Ed.), *Visuo-spatial working memory*. Hove, UK: Erlbaum.
- Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.
- Luck, S.J., Vogel, E.K., & Shapiro, K.L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, 383, 616–618.
- Marr, D. (1982). *Vision*. New York: W.H. Freeman.
- McConkie, G.W., & Currie, C.B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 563–581.
- Miller, E.K., Erickson, C.A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *The Journal of Neuroscience*, 16, 5154–5167.
- Milner, A.D., & Goodale, M.A. (1995). *The visual brain in action*. New York: Oxford University Press.
- Milner, P. (1974). A model for visual shape recognition. *Psychological Review*, 81, 521–535.
- Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind* (pp. 93–148). Stirling, UK: Erlbaum.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A.F. Sanders (Eds.) *Perspectives on perception and action* (pp. 361–394). Hillsdale, NJ: Erlbaum.
- Oram, M.W., & Perrett, D.I. (1994). Modeling visual recognition from neurobiological constraints. *Neural Networks*, 7, 945–972.
- Owen, A.M. (1997). Tuning in to the temporal dynamics of brain activation using functional magnetic resonance imaging (fMRI). *Trends in Cognitive Sciences*, 1, 123–125.

- Pashler, H. (1997). *The psychology of attention*. Cambridge, MA: MIT Press.
- Petrides, M. (1996). Specialized systems for the processing of mnemonic information within the primate frontal cortex. *Philosophical Transactions of the Royal Society of London: Biological Sciences*, 351, 1455–1462.
- Phaf, R.H., Van der Heijden, A.H.C., & Hudson, P.T.W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology*, 22, 273–341.
- Pollatsek, A., & Rayner, K. (1992). What is integrated across fixations? In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 166–191). Berlin: Springer Verlag.
- Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.
- Posner, M.I., & Raichle, M.E. (1994). *Constructing mental images*. New York: Scientific American Library.
- Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, 32, 65–97.
- Rafal, R.D. (1997). Balint syndrome. In T.E. Feinberg & M.J. Farah (Eds.), *Behavioral neurology and neuropsychology* (pp. 337–356). Berkshire, UK: McGraw-Hill.
- Rao, S.C., Rainer, G., & Miller, E.K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, 276, 821–824.
- Raymond, J.E., Shapiro, K.L., & Arnell, K.M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849–860.
- Rensink, R.A., O'Regan, J.K., & Clark, J.J. (1996). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8, 368–373.
- Rizzolatti, G., Riggio, L., & Sheliga, B.M. (1994). Space and selective attention. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance: XV. Conscious and nonconscious information processing* (pp. 231–265). Cambridge, MA: MIT Press.
- Roberts, R.J., Jr., Hager, L.D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, 123, 374–393.
- Rushworth, M.F.S., & Owen, A.M. (1998). The functional organization of the lateral frontal cortex: Conjecture or conjuncture in the electrophysiology literature? *Trends in Cognitive Sciences*, 2, 46–53.
- Scheerer, E. (1973). Integration, interruption and processing rate in visual backward masking. *Psychologische Forschung*, 36, 71–93.
- Schneider, W.X. (1993). Space-based visual attention models and object selection: Constraints, problems, and possible solutions. *Psychological Research*, 56, 35–43.
- Schneider, W.X. (1995). VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cognition*, 2, 331–375.
- Schneider, W.X., & Deubel, H. (1995). Visual attention and saccadic eye movements: Evidence for obligatory and selective spatial coupling. In J.M. Findlay, R. Walker, & R.W. Kentridge (Eds.), *Eye movement research* (pp. 317–324). Amsterdam: Elsevier.
- Schneider, W.X., & Deubel, H. (1996). Transsaccadic memory: A postsaccadic gap enhances the perception of intrasaccadic object changes differently for dorsal and ventral representations. *Perception*, 25, (Suppl. 9–10).
- Shapiro, K., Driver, J., Ward, R., & Sorensen, R.E. (1997). Priming from the attentional blink: A failure to extract visual tokens but not visual types. *Psychological Science*, 8, 95–100.
- Shapiro, K.L., Arnell, K.M. & Raymond, J.E. (1997). The attentional blink. *Trends in Cognitive Sciences*, 1, 291–296.
- Shibuya, H., & Bundesen, C. (1988). Visual selection from multiple displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 591–600.
- Simons, D.J. (1996). In sight, out of mind: When object representations fail. *Psychological Science*, 7, 301–305.
- Simons, D.J., & Levin, D.T. (1998). Change blindness. *Trends in Cognitive Sciences*, 1, 261–267.
- Singer, W. (1989). Search for coherence: A basic principle of cortical self-organization. *Concepts in Neuroscience*, 1, 1–26.
- Singer, W., Engel, A.K., Kreiter, A.K., Munk, M.H.J., Neuenschwander, S., & Roelfsema, P.R. (1997). Neuronal assemblies: Necessity, signature, and detectability. *Trends in Cognitive Sciences*, 1, 252–261.
- Smith, E.E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, 33, 5–42.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74, 1–29.
- Styles, E.A. (1997). *The psychology of attention*. Hove, UK: Psychology Press.
- Tanaka, K. (1997). Mechanisms of visual object recognition: Monkey and human studies. *Current Opinion in Neurobiology*, 7, 523–529.
- Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, 255, 114B–125.
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology*, 40, 201–237.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Ungerleider, L.G., & Haxby, J.V. (1994). “What” and “where” in the human brain. *Current Biology*, 4, 157–165.
- van der Heijden, A.H., Heij, W.L., Phaf, R.H., Buijs, D.A. & van Vleet, E.C. (1988). Response competition and condition competition in visual selective attention. *Acta Psychologica*, 67, 259–277.
- von der Malsburg, C. (1981). *The correlation theory of brain function. Internal Report*. Göttingen, Germany: Max Planck Institute for Biophysical Chemistry, 1–38.
- Wolfe, J.M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1, 202–238.
- Woodworth, R.S., & Schlossberg, H. (1954). *Experimental Psychology*. New York: Henry Holt.
- Yantis, S. (1996). Control of visual attention. In H. Pashier (Ed.), *Attention* (pp. 223–256). Hove, UK: Psychology Press.
- Zeki, S.M. (1993). *A vision of the brain*. Oxford: Blackwell Scientific Publications.