

Research Report

VISUAL SEARCH IS MODULATED BY ACTION INTENTIONS

Harold Bekkering^{1,2} and Sebastiaan F.W. Neggers^{1,3}

¹Department of Cognition and Action, Max Planck Institute for Psychological Research, Munich, Germany;

²Department of Experimental and Work Psychology, University of Groningen, Groningen, the Netherlands; and

³Helmholtz Institute, Unit of Psychonomics, Utrecht University, Utrecht, the Netherlands

Abstract—*The influence of action intentions on visual selection processes was investigated in a visual search paradigm. A predefined target object with a certain orientation and color was presented among distractors, and subjects had to either look and point at the target or look at and grasp the target. Target selection processes prior to the first saccadic eye movement were modulated by the different action intentions. Specifically, fewer saccades to objects with the wrong orientation were made in the grasping condition than in the pointing condition, whereas the number of saccades to an object with the wrong color was the same in the two conditions. Saccadic latencies were similar under the different task conditions, so the results cannot be explained by a speed-accuracy trade-off. The results suggest that a specific action intention, such as grasping, can enhance visual processing of action-relevant features, such as orientation. Together, the findings support the view that visual attention can be best understood as a selection-for-action mechanism.*

Typically, a person who searches for a specific object in the environment also has a specific intention of what he or she wants to do with that object. For instance, someone might search for a big red coffee cup with either the intention to grasp it and subsequently drink from it or the intention to point to it to communicate to a colleague that he or she should use that cup to have some coffee. It is important to note that the physical characteristics of the cup are the same (e.g., big and red) for the two tasks. Nevertheless, information about size and orientation of the cup is irrelevant to pointing to it, whereas such information is very relevant for the grasping action. For example, early movements aimed at grasping an object are influenced by the object's orientation (Mamassian, 1997), but pointing movements are not (Smeets & Brenner, 1999). Therefore, an efficient visual search system should detect action-relevant information in an enhanced way.

Some first evidence for the existence of an action-related attentional search system has recently been reported. For instance, Humphreys and Riddoch (2001) found that a patient with symptoms of unilateral neglect was slow and sometimes unable to find targets when they were defined by their name or even by a salient visual property (such as their color), but he was relatively efficient at finding a target defined by the action it afforded. In addition, Meegan and Tipper (1999) found that interference from a distractor was uniquely related to a reaching task and did not interfere when the same setup was used in a verbal task. The goal of the present research was to investigate further action-related selective visual attention in visual search.

Selective visual attention, the mechanism that enables an organism to select specific information that is registered by the senses for further processing and visually guided behavior, is currently thought to be con-

trolled by two distinct mechanisms. On the one hand, attentional control is said to be *stimulus-driven* when visual selective attention is activated in a bottom-up manner. In this case, the strength of bottom-up activation, which is based on the differences between neighboring items (Wolfe, 1994), guides attention toward distinctive items in the field (see also Duncan & Humphreys, 1989).

On the other hand, attentional control is said to be *user-driven* when visual selective attention is activated in a top-down manner. In this case, the observer is thought to select a subset of the available visual features that corresponds to the recognition task at hand. User-driven visual selection can be best explained by the example of a conjunction search. In a conjunction search, the target might be specified as an object with two specific features; for example, the target might be a green cylinder, and distractors might be orange cylinders and green and orange rectangles. Objects with the same shape or the same color (not both) are more often erroneously looked at than objects that do not share a feature with the saccadic target (Findlay, 1997). It is therefore argued that top-down specification plays a role not only in the programming of saccadic eye movements but also in visual search in general (Cave & Wolfe, 1990).

Remarkably, however, although attentional control is suggested to be user-driven, or *goal-directed* (Gibson, 1996), the interaction between the visual information and the actor has typically been minimized. Goal-directed visual processing has been investigated in cognitive tasks; for example, participants have been asked to report the presence of a certain item in a full display by selecting one of two possible button presses, an act that by itself does not require knowledge of an object property in order to be executed (e.g., Treisman & Gormican, 1988). In other words, the top-down specification is set by the task instructions, and not by the action itself.

The aim of this study was to investigate whether visual selective processes are modulated by the action intentions of the actor. Specifically, the question was whether processing of task-relevant information is enhanced compared with processing of task-irrelevant information. In one condition, subjects were asked to saccade and point to a target object as quickly and accurately as possible; in the other condition, they were asked to saccade to and grasp the predefined target object. Orientation, but not color, is known to influence the way an object is grasped. In contrast, pointing to a target is known to be insensitive to both object orientation and color (Smeets & Brenner, 1999). Therefore, fewer orientation errors were expected in the saccade-and-grasp condition than in the saccade-and-point condition, but the number of color errors was expected to be the same in the two conditions.

METHOD

Participants

Participants had normal or corrected-to-normal vision and were tested to ensure they were not color-blind. The 8 participants, 4 males and 4 females, had a mean age of 25 years and were paid 15 Deutsch Marks.

Address correspondence to Harold Bekkering, Experimental and Work Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands; e-mail: h.bekkering@ppsw.rug.nl.

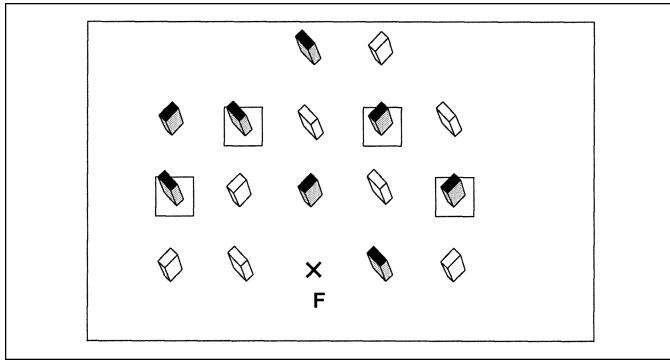


Fig. 1. Illustration of the object array, showing the 16 possible locations of targets (black upper surfaces) and distractors. The outline squares mark the four target positions that were equidistant from the fixation dot (F) and used for the error analysis.

Apparatus

A typical visual conjunction search task was used. Subjects needed to select a predefined target having a specific orientation (45° or 135°) and color (green or orange). The objects in the display were rectangular blocks (20 mm long \times 10 mm wide \times 20 mm high), which were placed in the holes in a Plexiglas surface, above a 12×8 array of LEDs (see also Neggers & Bekkering, 2000). An object was visible only when illuminated by the LED immediately below it. Sixteen objects were placed on the board and served as potential targets and distractors. Of these 16 objects, 8 were possible targets across trials (shown in black in Fig. 1): Four were oriented 45° and 4 were oriented 135° to the right. Four of these possible target objects (marked with squares in Fig. 1) were located at the same distance from the fixation point, and only trials with these target objects were analyzed. Orientations were interchanged (between 45° and 135°) after each block of trials in order to prevent memorization, and color varied constantly from trial to trial. On each trial, the positions of the distractors were randomly chosen from the 16 possible locations.

Four display sizes were used, one containing the target object only, and the others containing the target object and an additional 3, 6, or 9 distractor objects. Distractors differed from the target with respect to orientation (45° or 135°), color (green or orange), or both, according to the following rule: One third of the distractor objects had the same color as the target but a different orientation, one third had the same orientation as the target but a different color, and one third differed from the target with respect to both orientation and color. Number of distractors was randomized throughout the experiment.

Eye and head movements were measured with the SMI™ Eye Link, and the positions of the hand and the LED board were measured with the OPTOTRAK Northern Digital™.

Procedure

Subjects were seated in a totally dark room and instructed to point to or grasp the target object as accurately and naturally as possible. Specifically, they were required to point to the center of the top surface of the target or to grasp the object between the thumb and index finger. Also, they were instructed to gaze at the fixation dot until they thought they knew where the target was located. No other instructions

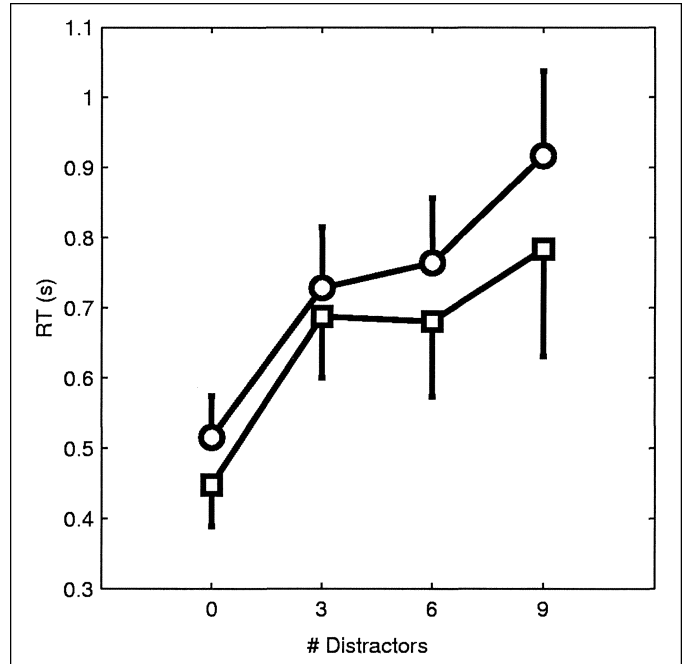


Fig. 2. Mean reaction times (RTs) and standard errors for saccadic eye movements in the saccade-and-point condition (squares) and saccade-and-grasp condition (circles) as a function of the number of distractors.

about the eye movements were given, because it is well known that subjects naturally look at a target before aiming (Abrams, Meyer, & Kornblum, 1990; Bekkering, Adam, Kingma, Huson, & Whiting, 1994; Bekkering, Adam, van den Aarssen, Kingma, & Whiting, 1995; Neggers & Bekkering, 1999).

A trial started when the fixation point lit up and a stable ocular fixation was reached, which was checked by the eyetracker. Then, a 2000- or 1000-Hz tone indicated whether the target object had an orientation of 45° or 135° , respectively. When the tone sounded, the fixation point turned from red to green or orange to indicate the color of the target object. After 1,500 ms, the object set lit up. Subjects participated in four blocks of trials, two that required grasping movements and two that required pointing movements. The order of blocks was balanced (ABAB design); for half the subjects, the saccade-and-grasp task was first, and for the other half, the saccade-and-point task was first. Each block consisted of 64 trials (4 possible numbers of distractors \times 2 target colors \times 2 target orientations \times 4 repetitions).

RESULTS

Hits

The probability of hitting the correct target with the first saccade decreased significantly when the number of distractors increased, $F(2, 7) = 20.2$, $p < .01$. There was no main effect for task condition (pointing vs. grasping), $F(1, 7) < 1$. Also, the interaction effect was nonsignificant, $F(2, 14) = 1.3$, $p > .2$.

Saccadic reaction times showed a main effect for the number of distractors, $F(1, 7) = 8.67$, $p < .01$ (see Fig. 2). No main effect for

Action-Modulated Visual Search

task condition (pointing vs. grasping) was present, $F(1, 7) = 1.62, p > .23$.¹ Also, no significant interaction effect was found, $F(2, 14) < 1$.

Error Analysis

The error analysis included data from those trials (on average 67%) on which participants' saccades landed within a 40-mm-square area centered on an object. (See Fig. 3 for an illustration of typical saccadic color and orientation errors.) A 3 (number of distractors: 3, 6, or 9) \times 2 (task: grasping or pointing) \times 2 (feature type: color or orientation) analysis of variance was performed. Errors increased with the number of distractors, $F(2, 14) = 12.3, p < .01$. Also, subjects made more orientation errors than color errors, $F(1, 7) = 16.4, p < .01$. No main effect for task was found, $F(1, 7) = 1.3, p > .2$. However, an interaction effect between feature and task was present, $F(1, 7) = 7.9, p < .05$, indicating that orientation errors were less frequent in the grasping than in the pointing condition, whereas the number of color errors was similar in the two conditions (see Fig. 4). Also, a three-way interaction was present, $F(2, 14) = 5.4, p < .05$, indicating that the number of orientation errors in the two conditions differed particularly when the number of distractors was high.

GENERAL DISCUSSION

The goal of this experiment was to examine whether a specific action intention can enhance visual processing of action-relevant features. To this end, participants had to search for, and saccade to, a target object among distractors in two conditions, a saccade-and-point condition and a saccade-and-grasp condition. Clearly, orientation selection, but not color selection, was better in the saccade-and-grasp condition than in the saccade-and-point condition. That is, the first eye movement was more accurate in selecting objects with the correct orientation when the object needed to be grasped afterward than when the object needed to be pointed to. Given that object orientation is relevant for a grasping movement whereas color discrimination is not, the present findings support the view that the planning of motor action depending on a particular object feature modulates the visual processing of that feature.

It is important to note that saccadic eye movements in general precede goal-directed aiming movements, whether one grasps or points to a target. Therefore, the visual enhancement observed can only be ascribed to action-planning mechanisms and not to action-execution mechanisms. In other words, it is the intention of an action that modulates the visual processing of object features. In addition, the absence of a significant increase in saccadic reaction times in the grasping condition compared with the pointing condition indicates that no speed-accuracy trade-off was present. That is, the increased accuracy of the first saccadic fixation in the grasping condition did not result in longer reaction time latencies compared with the pointing condition.

One could argue that fewer color than orientation errors were made in general, and that this might explain why reductions in orientation errors only were found in the saccade-and-grasp condition compared with the saccade-and-point condition. More color errors were made,

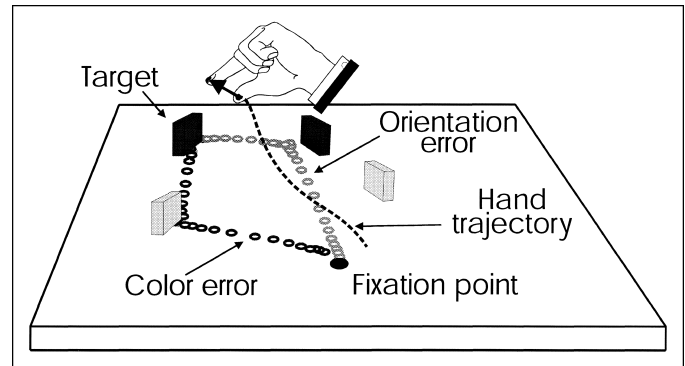


Fig. 3. Spatial overview of two possible incorrect saccadic trajectories (shown by circles), one to a distractor with the wrong orientation but the correct color (orientation error) and one to a distractor with the wrong color but the correct orientation (color error). The correct hand trajectory to the target is also shown.

however, with greater numbers of distractors (six or nine vs. three); thus, improvements in the number of color errors should have been possible in trials with six or nine distractors. It is therefore highly unlikely that a floor effect explains the results.

Another important theoretical implication of the present study stems from the fact that the specific action intention (grasp vs. point) cannot directly modify the components of the search template. That is, regardless of the task condition, subjects had to select the target object with both correct features (orientation and color) to avoid errors. Thus, the top-down modification of visual processing occurred within the specific dimension of orientation, and was not due to selection of the task-relevant dimension per se. Accordingly, a working hypothesis to explain the present results is that action intentions can modify the tuning of neural channels representing specific features of the outer world. For instance, the brain may represent action-relevant features with a higher resolution than action-irrelevant features in order to optimize action planning.

What might be the underlying mechanism of this action-related enhancement in visual search? One possible candidate could be biased competition (Chelazzi, Duncan, Miller, & Desimone, 1998; Desimone, 1996; Desimone, Chelazzi, Miller, & Duncan, 1995; Luck, Chelazzi, Hillyard, & Desimone, 1997). Biased competition is based on the assumption that objects in the visual field activate neural representations in the cortex in a parallel fashion, and that the cells participating in these representations engage in competitive interactions. Frequently, these interactions are evidenced by suppressive effects of one stimulus on the response to another (Rolls & Tovee, 1995), and such competition has been posited for both stimulus-driven mechanisms and top-down (often called attentional-driven) mechanisms (Luck et al., 1997; Moran & Desimone, 1985). None of the studies reporting such competitive interactions have reported motor-related effects, however. For instance, Chelazzi et al. found qualitatively similar effects of attention in the inferior temporal cortex for both the saccade and the lever-release version of a visual search task. In that study, however, the action by itself did not require knowledge of an object property; in other words, the target was relevant to get a reward, but not to execute the action.

Therefore, to explain the present modulation in visual search performance, we propose a more specific action-related view of visual attention. Several authors have proposed that visual attention might be

1. Although there was a nonsignificant trend for saccades to be slower in the grasping condition than in the pointing condition, there was no difference in search slopes. Because search slopes are better estimates of visual processing rate than overall reaction time, any possible intercept effect in reaction time would not hold any theoretical importance.

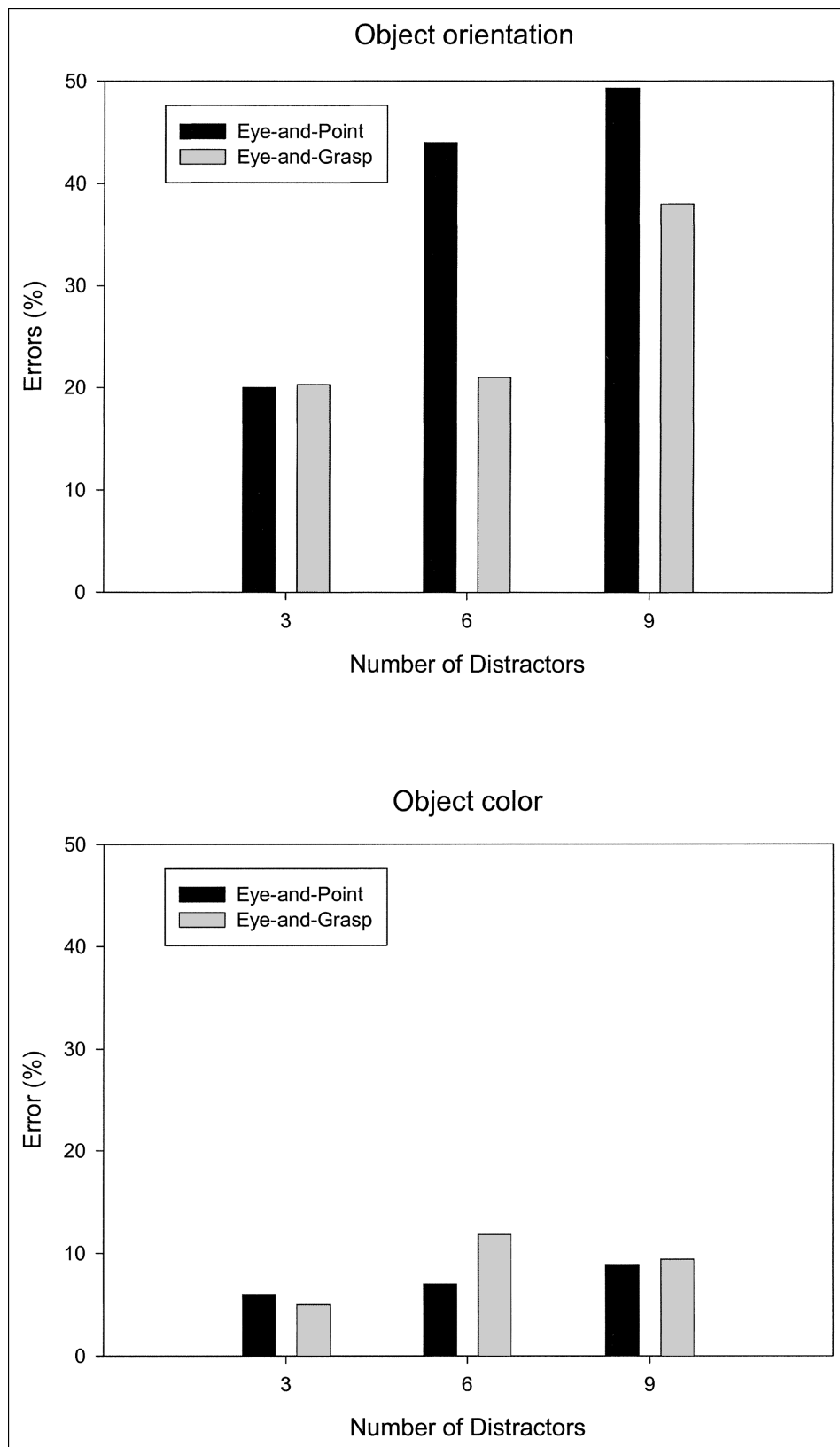


Fig. 4. Percentage of first saccadic eye movements that were made to objects with the wrong orientation (top panel) and the wrong color (bottom panel) as a function of the number of distractors and task.

Action-Modulated Visual Search

seen as a selection-for-action mechanism (Allport, 1987; Deubel & Schneider, 1996; Schneider, 1995). In such a view, the same cognitive control mechanism contributes to both efficient, task-related perception of a visual scene and the specification of action parameters (Neumann, 1987). That is, the intention to perform an action will result in a top-down modulation of visual processes that favors object features (e.g., orientation) that are directly related to the ongoing specification of parameters for action control (e.g., characteristics of the grasping movement). In this view, the perception of specific object features and the planning of an action to interact with an object to some extent tap into the same mechanism (Prinz, 1997).

Neurophysiological support for such a close coupling between action and perception has been widely reported. For example, Moore (1999) described a connection between early visual stimulus coding and guidance of the eyes to prominent features of an object. Remarkably, if certain cells in area V4 showed high presaccadic activity, eye movements landed accurately along the target orientation axis, but when these cells showed low presaccadic activity, eye movements landed randomly with respect to the orientation axis. This finding suggests that neurons in V4 not only participate in the passive representation of visual stimuli but also contribute to the guidance of saccadic eye movements. Moreover, Mushiaki, Tanatsugu, and Tanji (1997) found differential presaccadic activation of neurons in the supplementary eye field during a saccade-only condition versus a saccade-and-reach condition, suggesting that these neurons are involved in signaling whether the motor task is oculomotor or combines eye and arm movements.

Together, these findings suggest that action intentions influence visual attentive processes at a very early stage. Clearly, more experiments are needed to investigate this essential but widely neglected interaction between action intentions and the manner in which people search for objects in the visual environment.

Acknowledgments—This article reflects equal contributions from the two authors. We would like to acknowledge Diana Romero, Jay Pratt, Mieke Donk, Bruce Bridgeman, Rob Ward, James Enns, and an anonymous reviewer for helpful comments on an earlier version of this manuscript. We also thank Fiorello Banci and Karl-Heinz Honsberg for technical support and Julian Garbotz and Anne Häberle for running the experiments.

REFERENCES

- Abrams, R.A., Meyer, D.E., & Kornblum, S. (1990). Eye-hand coordination: Oculomotor control in rapid aimed limb movements. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 248–267.
- Allport, 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.
- Bekkering, H., Adam, J.J., Kingma, H., Huson, A., & Whiting, H.T. (1994). Reaction time latencies of eye and hand movements in single- and dual-task conditions. *Experimental Brain Research*, 97, 471–476.
- Bekkering, H., Adam, J.J., van den Aarsen, A., Kingma, H., & Whiting, H.T. (1995). Interference between saccadic eye and goal-directed hand movements. *Experimental Brain Research*, 106, 475–484.
- Cave, K.R., & Wolfe, J.M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology*, 22, 225–271.
- Chelazzi, L., Duncan, J., Miller, E.K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, 80, 2918–2940.
- Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences, USA*, 93, 13494–13499.
- Desimone, R., Chelazzi, L., Miller, E.K., & Duncan, J. (1995). Neuronal mechanisms of visual attention. In T.V. Papathomas & C. Chubb (Eds.), *Early vision and beyond* (pp. 219–226). 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.
- Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Findlay, J.M. (1997). Saccade target selection during visual search. *Vision Research*, 37, 617–631.
- Gibson, B.S. (1996). Visual quality and attentional capture: A challenge to the special role of abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1496–1504.
- Humphreys, G.W., & Riddoch, M.J. (2001). Detection by action: Neuropsychological evidence for action-defined templates in search. *Nature Neuroscience*, 4(1), 84–88.
- Luck, S.J., Chelazzi, L., Hillyard, S.A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, 77, 24–42.
- Mamassian, P. (1997). Prehension of objects oriented in three-dimensional space. *Experimental Brain Research*, 114, 235–245.
- Meegan, D.V., & Tipper, S.P. (1999). Visual search and target-directed action. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1347–1362.
- Moore, T. (1999). Shape representations & visual guidance of saccadic eye movements. *Science*, 285, 1914–1917.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782–784.
- Mushiaki, H., Tanatsugu, Y., & Tanji, J. (1997). Neuronal activity in the ventral part of premotor cortex during target-reach movement is modulated by direction of gaze. *Journal of Neurophysiology*, 78, 567–571.
- Neggers, S.F., & Bekkering, H. (2000). Ocular gaze is anchored to the target of an ongoing pointing movement. *Journal of Neurophysiology*, 83, 639–651.
- Neggers, S.F.W., & Bekkering, H. (1999). Integration of visual and somatosensory target information in goal-directed eye and arm movements. *Experimental Brain Research*, 125, 97–107.
- 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.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129–154.
- Rolls, E.T., & Tovee, M.J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, 73, 713–726.
- 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, 331–376.
- Smeets, J.B., & Brenner, E. (1999). A new view on grasping. *Motor Control*, 3, 237–271.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15–48.
- Wolfe, J.M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1, 202–238.

(RECEIVED 12/21/00; REVISION ACCEPTED 8/17/01)