

VISUAL ATTENTION AND COGNITION: A SYNOPSIS OF CLINICAL ASPECTS

Wolfgang H. Zangemeister

Neurological University Clinic, Hamburg

Martinistr.52, D 20251 Hamburg, FRG

e-mail: zangemeister@uke.uni-hamburg.de

Abstract

In this synopsis of chapter 3, an introduction and overview is given with regard to the functional anatomical and clinical neurophysiological aspects of visual attention and cognition.

Humans operate in a complex data rich world. Our visual system is constantly confronted with situations that require rapid processing and decision making. Yet, we are able to analyze almost effortlessly, information critical to the task at hand, while ignoring vast amounts of nonessential information. How is this done so effectively? Some have conjectured that top-down processes involving learning and semantic memory account for our ability to process information rapidly (Noton & Stark 1971; Homa et al. 1976; Friedman 1979; Julesz 1991). These processes form higher level associations among the components in a scene. Hence, efficiency in processing is gained by reducing the need for an element-by-element encoding of each item in the scene. How are these associations created? The associations among the components of a scene may be formed, for example, by using their spatial relations (Ullmann 1985), when performing a visual search. A simple example is that a schematic of a face is more readily perceived when the components of the eyes, nose and mouth are normally arranged as opposed to being scrambled. Another means of consolidating information is by its contextual significance based on the relation among the objects in a scene. Context in the natural environment plays an important role. This is because in a natural scene objects have powerful and complex relations, and the association of items in a scene provides the means for rapid and effective processing of visual information. The underlying association of this top-down contextual process with the bottom-up processes in early vision was recently investigated (Hung et al. 1995). It was demonstrated that better performance for simultaneous over sequentially presented items was a general phenomenon, independent of the kind of icons used. They suggested that both a parallel buffer and a serial retrieval mechanism linked to short term memory were involved in the early stages of visual processing. At the connectivity level of visual processing, contextual effects involved higher level processes that incorporated the complex relations among objects. The fact that these relations could be formed rapidly for briefly presented objects indicated that cortical neural connectivity may already be in place to facilitate these interactions.

Visual search has also been a subject of considerable study and interest in the neurosciences as well as operations research. Information as to how humans approach the search task is of great practical interest for applications ranging from air surveillance to manufacturing quality control inspections to search and rescue missions. The mathematical aspects of search theory go back to the probability of detection theory (Koopman 1956, 1957) that assumes that each glance or fixation is independent of the others. Most of these subsequent mathematical models have been based upon the effects of external environmental factors on search performance, ignoring or minimizing the effects of internal cognitive factors. Recent work includes psychophysical and human factors aspects of complex visual search and studies of eye movements during search. Storing and retrieving memories are important components of visual pattern recognition. Hence, the memory system of the brain must contain an internal representation of each pattern that is to be recognized. The process of constructing such representations is obscure. Recognition of a pattern may be viewed as the process of matching it with its stored internal representation. A non Gestalt view suggests that the internal model consists of component features that are matched step by step with the pattern during recognition. This serial recognition process is supported by the findings in several studies on object and pattern recognition that the eyes seem to visit the features of the object or items in the scene almost cyclically, following fairly stereotyped regular scanpaths or searchpaths rather than in random sequences. The scanpath was proposed to be the read-out of the internal representation of pictures, the so-called "cognitive model" (Noton & Stark 1971; Stark & Choi, this book).

Hochberg (1970) introduced the useful distinction between peripheral search guidance, in which the eye was drawn to move by information in the visual periphery and cognitive search guidance, in which the control of the eye was determined by some central plan as was suggested by the scanpath idea. It is difficult to develop these ideas at a general level because of the variety of forms that might be envisaged for such a central plan. However, the paradigm of visual search usefully constrains the subject's central plan to the task of locating a prespecified target. It is theoretically possible for eye movements during visual search to be totally controlled by cognitive guidance: for example by executing a predetermined systematic scan of the search area until the target is located in the fovea. However, it seems that more frequently, subjects attempt to extract some information from peripheral vision to direct the eyes to the target and thus the pattern of the eye movements during search appears much less systematic. Specification of the size of the object is much less effective and subjects show very little ability to direct their fixation to objects of a prespecified shape. These conclusions in many ways anticipated a current theme in visual search, the feature integration theory (Treisman 1980). This theory states that simple features can be searched for rapidly in parallel, whereas more complex features and combinations of features require a serial search process, involving either covert or overt attentional shifts. Feature integration theory has for the most part relied on indirect measures of attention and paid little attention to the detailed mechanisms of search such as oculomotor control.

The processes underlying shifting of attention from one item to another have recently been studied (Findlay; Seitz; Husain & Kennard, this book). Subjects were required to shift spatial attention in the right or left visual field along foveofugal or foveocentral directions. It was found that the superior parietal and superior lateral frontal cortex were

more active when *shifting* attention compared to *maintaining* attention at the center of gaze. Further, the superior parietal region was active when peripheral locations were selected on the basis of cognitive or sensory cues, while the frontal region was active only when responses were made to stimuli at selected peripheral locations. These observations pointed to a more perceptive role of the parietal activations and motor related activation in the frontal lobe for shifts of attention. Interestingly, these areas of activation were present in both cerebral hemispheres for both visual hemifields in an almost overlapping manner. However, in the parietal lobe of the right hemisphere the visual hemifields were separated with a more posterior location of the right visual hemifield. Since the direction of the moving stimuli did not separate the activation areas, these rCBF increases in the parietal lobe did not simply reflect responses to spatial locations. Rather, they indicated a more widespread involvement of the right parietal lobe for shifting of attention. Raised attention during a first PET scan has been shown to specifically activate the prefrontal cortex, the superior parietal cortex, the cingulate, and the thalamus (Seitz & Roland 1992). Conversely, patients suffering from hemineglect revealed significant metabolic depressions in superior lateral frontal cortex, superior and inferior parietal, and cingulate cortex (von Giesen et al. 1994). These areas included or were located in close neighbourhood to those specifically activated in active focussing and shifting of attention. These positive (activation) and negative (lesion) findings support the concept of a large-scale neuronal system underlying visual attention (Mesulam 1990).

Husain & Kennard start their review with the simple question: What happens during the saccadic reaction time? They review evidence which suggests focal attention normally engages a visual stimulus before a saccade is made to foveate it. Attention can be directed without making an eye movement and, under certain circumstances, in the direction opposite to an eye movement. So, the directing of focal attention does not automatically lead to the generation of a saccade and the two processes can be dissociated. Attention normally appears to shift to the target of a saccade before the eye movement is made. It therefore seems to be an important early step in preparing an eye movement. Exactly what this means in terms of neural operations or representations is yet unclear, but our understanding may improve with further analysis of the contributions of the posterior parietal cortex (PPC), pulvinar and superior colliculus.

The main function of the peripheral part of the retina is that of 'sentinels' which, when beams of light move over them, 'cry: "Who goes there?" and call the fovea to the spot,' remarked William James (1890). Husain & Kennard suggest, there is sufficient evidence to conclude that the parietal cortex, pulvinar and superior colliculus form part of the core neural architecture normally responsible for directing attention and gaze to visual targets. This group of structures serves the role of 'sentinel' in the primate visual system.

Visual search is influenced by multi-item boundary and surface groupings. These may indeed represent the perceptual representations on which the search process is based. The identification of a grouping that includes multiple items speeds search by reducing the total number of candidate visual regions that have serially to be investigated. Factors which influence boundary and surface grouping, such as featural contrast, item spacing, and spatial arrangement alter this number of visual regions to be explored, yielding variations in search time. If bottom-up mechanisms may drive the formation of these emergent perceptual units, then limits must exist on the capacity of semantic or even visual defini-

tions of target items to exert top-down influence over preattentive grouping mechanisms. The ability of bottom-up processing to accurately distinguish ecological objects depends on a certain amount of autonomy or resistance to top-down interference. Otherwise, it would routinely result in perceptual illusions. Perceptual grouping indeed will often be guided by top-down processes (Stark et al.1992; Desimone 1993). However, some groupings may “emerge” from the structure of the scenic input without the help of top-down influences. Of course, the enforced bottom-up control of viewing is the main domain of our everyday modern life through movies, TV and visual public relation in particular. We learn and we apply these different kinds of top-down control of viewing during our whole life. However, diseases of the eyes, the optical pathways, the visual or motor cortex and its interconnections may cause, that at least one of the three parts of this control becomes disturbed: the sensory, the cognitive, or the motor connection that contribute to the proper functioning of these high levels of visual control. In case of deficiencies of one of these functional parts there is a need to recover from the given deficit, which may be feasible through certain strategies of adaptation. The typical, most frequent deficits that can be found clinically, — and may be simulated experimentally —, are: (1) Motor deficits of one or two eyes with deficits of coordinated eye-movements that may cause doublevision, or slowness and inaccuracy of eye fixation and eye movement; they can be overcome comparatively easily by moving only the healthy eye, and neglecting, i.e. suppressing the information of the eye with the movement deficits; or by helping interocular deficits through adaptive eye- and head- coordination, like in internuclear ophthalmoplegia. (2) More importantly, sensory deficits may disturb top-down control of vision by visual field defects of one eye, or both eyes in case of more centrally located disturbances as is the case in homonymous hemianopia. (3) Most variant difficulties and therefore a whole variety of adaptive strategies may occur with deficits of visual attention and cognition, like visual apraxia and hemineglect. Studies of these effects in hemianopic patients (Zangemeister & Oechsner, this book; Zangemeister et al.1982, 1985, 1995; Schoepf & Zangemeister 1993) demonstrate that it is feasible and quantifiable to observe short term adaptation as an effect of short term training in patients with hemianopic field defects who apply and optimize a high level, top-down visuo-motor strategy to search and scan for targets and sequences of targets in complex visual tasks. This strategy is also evident when patients use very small eye movements, i.e. mini saccades. Evidences for top-down versus bottom-up control are given from that study with respect to the paradox that top-down cognitive models prevail when we see, whereas local stair-steps of bottom-up control prevail, when we are blind. Also, the “complexity” of the picture - “attraction versus distraction”- influences the control of eye movement sequences of fixations in the case of homonymous hemianopia. Evidently, global viewing is the preferred strategy for the healthy subject, who tries to evaluate at the same time both the visual content and the complexity of the picture. Hemianopic patients however, are more busy with developing an optimal sequence of eye movements to detect the overall features of the picture when searching or scanning, since they have primarily to rely on more local and therefore limited picture evaluations that also include more bottom-up control than in the healthy subjects.

What is local scanning? Even though Noton & Stark (1971) and Stark & Ellis (1981) showed that peripheral information can be excluded as the immediate control for the

scanpath, their results relate to local scanpaths. Groner et al. (1984) and also Finke (1983) support their top-down, cognitive model scanpath theory for a global scanpath, but argue in favour of an immediate peripheral bottom-up control of local scanning, although evidence for the latter is not conclusive at the present time. Jeannerod et al. (1968) has argued for an exchange between local and global scanning in free exploration. Evidently, the normal healthy viewer avoids this type of immediate bottom-up control in favour of the top-down controlled global scanpath, whereas the patient when viewing to the side of the blind hemifield relies strongly on such an exchange, that permits him to develop a more and more efficient strategy of searching and scanning with every repetition.

Whether the local scanpath is driven immediately by peripheral, bottom-up information or by small-scale cognitive models is unknown. Locher & Nodine (1987) have claimed immediate bottom up control in symmetry "that catches the eye". Mackworth & Morandi (1967) showed evidence for top-down active selection of informative details through "active looking". This detailed looking is apparently usually applied for realistic images, where anticipation of details may be balanced by a permanent exchange of bottom-up and top-down control (Zangemeister et al. 1995). Hemianopic patients carry this behaviour on to ambiguous and non-realistic images. Obviously with increasing complexity it is more difficult, to apply efficiently a bottom-up control as was shown earlier by Berlyne (1958; 1971), and this result applies also for the blind side of hemianopic patients.

Hemianopic patients may have lost one half of their central vision. But they are still capable, and often highly efficient, in integrating what they have "looked at": such that they may even "see" what is in their blind hemifield through applying preview control, prediction and closely connected visual integration. By far the most interesting examples of a failure of visual integration are to be found in patients with visual agnosia. Neurologists commonly speak of such patients as if they are 'form blind', suffering from object agnosia. But opinion on the subject has been divided, at least in part because the syndrome itself is complex and manifests itself with variations in different patients. Patients may be able to recognize some objects, but not others; they may not recognize an object at one examination and yet be able to do so at a subsequent one. Some may be able to read while others cannot. The lesions are commonly large, often associated with scotomas and some, but not all, patients suffer from problems of amnesia, aphasia and general mental deterioration.

All this makes it difficult to relate a specific impairment to a specific cerebral defect. Indeed, some neurologists have put forward the view that visual agnosia is nothing more than the consequence of a failing visual apparatus. Yet there exists a sufficient number of patients whose eyes are normal, who are not aphasic and who do not suffer from mental deterioration to testify to the fact that there is a syndrome in which patients can apparently see objects, or at least parts of objects, and yet be unable to recognize what the objects are. Since integration itself is a multistage process, one should not be surprised to find that there are degrees of agnosia, ranging from the severe effects of carbon monoxide poisoning, due to damage of VI itself, to the relatively mild ones due to damage of more central visual areas. Of course, integration can also operate in the opposite direction, i.e. top-down, and patients as well as healthy people can be made to see things once they have understood them, but not until then. So, this is another way of looking at these defects, and the result of such an enquiry leads us to the view that seeing and understanding

merge into one another, and are not discrete activities localizable to different parts of the cerebral cortex. Contrariwise, clinical evidence suggests that many examples of visual agnosia can be considered to be failures of the integrative mechanisms in the brain, leading the patient to both see and understand only in relation to the capabilities of the *intact* parts of the brain.

References

- Berlyne, D.E. (1958), 'The influence of complexity and novelty in visual figures on orienting responses', *J. Exp. Psychol.* **55**, 289–296.
- Berlyne, D.E., McDonnell, P. (1951), 'Effects of complexity and prechoice stimulation on exploratory choice', *Perception and Psychophysics* **10**, 241 – 246.
- Finke, R.A. (1983), 'Directional scanning of remembered visual patterns', *J. Exp. Psychol.* **9**, 398 – 410.
- Friedman, A. (1979), 'Framing pictures: the role of knowledge in automatized encoding and memory for gist'. *J. Exp. Psychol. Gen.* **108**, 316–355.
- Groner, R., Walder, F., Groner, M. (1984), 'Looking at faces: Local versus global aspects of scanpaths', in A.G. Gale and F. Johnson, eds, 'Theoretical and Applied Aspects of Scanpaths', North Holland Publ. Company, Amsterdam, pp. 58 – 59.
- Hochberg, J. (1968), 'In the Mind's Eye'. in R.N. Haber, ed, 'Contemporary Theory and Research in Visual Perception', Holt, Rinehart and Winston Publ., New York, pp. 309 – 312.
- Hochberg J. (1970), 'Components of literacy: speculations and exploratory research', in H. Levin and J.P. Williams, eds, 'Basic studies on reading', New York, Basic Books.
- Homa D., Haver B., Schwartz T. (1976), 'Perceptibility of schematic face stimuli: evidence for a perceptual gestalt', *Mem. Cognit.* **4**, 176–185.
- Hung G.K., Wilder J., Curry R., Julesz B. (1995), 'Simultaneous better than sequential for brief presentations', *J. Opt. Soc. Am. A* **12**, 441–449.
- James, W. (1890), 'The Principles of Psychology', New York, Holt.
- Jeannerod, M., Gerin, P., Pernier, J. (1968), 'Déplacements et fixation du regard dans l'exploration libre d'une scène visuelle', *Vision Res.*, **8**, 81–97.
- Julesz, B. (1991), 'Early vision and focal attention', *Rev. Modern Physics* **63**, 735–772.
- Koopman, B.O. (1956), 'The theory of search Pt. I: Kinematic bases', *Oper. Res.* **4**, 324–246.
- Koopman, B.O. (1956), 'The theory of search Pt. II: Target detection', *Oper. Res.* **4**, 503–531.
- Koopman, B.O. (1957), 'The theory of search Pt. III: The optimum distribution of searching effort', *Oper. Res.* **5**, 613–626.
- Locher, P.J., Nodine, C.F. (1987), 'Symmetry catches the eye', in J.K. O'Reagan and A. Levy-Schoen, eds, 'From Physiology to Cognition', North Holland Publ. Company, Amsterdam, p. 353.
- Mackworth, N.H., Morandi, A.Y. (1967), 'The gaze selects informative details within pictures', *Perception and Psychophysics* **2**, 547–552.

- Mesulam, M.M. (1990), 'Large-scale neurocognitive networks and distributed processing for attention, language, and memory', *Ann. Neurol.* **28**, 597–613.
- Noton, D., Stark, L. (1971), 'Scanpaths in eye movements during pattern perception', *Science* **171**, 308–311.
- Schoepf D., Zangemeister, W.H. (1993), 'Correlation of Coordinated Gaze Strategies to the Status of Adaptation in Patients with Hemianopic Visual Field Defects', *Ann. N.Y. Acad. Sci.* **682**, 404–409.
- Stark, L., Ellis, S. (1981), 'Scanpaths revisited: Cognitive Models in Active Looking', in B. Fisher, C. Monty and M. Sanders, eds, 'Eye Movements, Cognition and Visual Perception', Erlbaum Press, New Jersey, pp. 193–226.
- Stark, L., Yamashita, I., Tharp, G., Ngo, H.X. (1992), 'Searchpatterns and searchpaths', in D. Brogan and K. Carr, eds, 'Visual Search II', Taylor & Francis, pp. 37–58.
- Seitz, R.J., Roland, P.E. (1992), 'Variability of the rCBF measured with (11C)-fluoromethane and positron emission tomography (PET) in rest', *Comp. Med. Imaging. Graphics.* **5**, 311–322.
- Treisman, A. (1980), 'A feature integration theory of attention', *Cognit. Psychol.* **12**, 97–136.
- Ullman, S. (1985), 'Visual Cognition, Visual Routines', Cambridge, MA, MIT Press.
- von Giesen, H.J., Schlaug, G., Steinmetz, H., Benecke, R., Freund, H.J., Seitz, R.J., 1994, Cerebral network underlying unilateral motor neglect: evidence from positron emission tomography. *J. Neurol. Sci.* **125**, 29–38.
- Zangemeister, W.H., Meienberg, O., Stark, L., Hoyt, W.F. (1982), 'Eye-Head Coordination in Homonymous Hemianopia', *J. Neurol.* **225**, 243 – 254.
- Zangemeister, W.H., Dannheim, F. (1985), 'Adaptation of Gaze to Eccentric Fixation in Homonymous Hemianopia', in E.L. Keller and D. Zee, eds, 'Adaptive Processes in Visual and Oculomotor Systems', *Adv. in Bio. Sci.* **57**, 247 – 252.
- Zangemeister, W.H., Sherman, K., Stark, L. (1995), 'Evidence for global Scanpath strategy in viewing abstract compared to realistic images', *Neuropsychologia* **33**, 1009–1025.
- Zeki, S., Shipp, S. (1988), 'The functional logic of cortical connections', *Nature* **335**, 311–317.