Visual Neuroscience (1990), 5, 353–369. Printed in the USA.

Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: A case study

LUCIA M. VAINA,[1,2] MARJORIE LEMAY,[3] DON C. BIENFANG,[3]
ALBERT Y. CHOI,[1] AND KEN NAKAYAMA[4]
[1]Intelligent Systems Laboratory, College of Engineering and the Department of Neurology, School of Medicine, Boston University
[2]Harvard–Massachusetts Institute of Technology, Division of Health Sciences and Technology
[3]Brigham and Women's Hospital, Harvard Medical School
[4]Smith Kettlewell Eye Research Institute

(RECEIVED April 17, 1990; ACCEPTED June 18, 1990)

Abstract

A series of psychophysical tests examining early and later aspects of image-motion processing were conducted in a patient with bilateral lesions involving the posterior visual pathways, affecting the lateral parietal-temporal-occipital cortex and the underlying white matter (as shown by magnetic resonance imaging studies and confirmed by neuro-ophthalmological and neuropsychological examinations). Visual acuity, form discrimination, color, and contrast-sensitivity discrimination were normal whereas spatial localization, line bisection, depth, and binocular stereopsis were severely impaired. Performance on early motion tasks was very poor. These include seeing coherent motion in random noise (Newsome & Paré, 1988), speed discrimination, and seeing two-dimensional form from relative speed of motion. However, on higher-order motion tasks the patient was able to identify actions from the evolving pattern of dots placed at the joints of a human actor (Johansson, 1973) as well as discriminating three-dimensional structure of a cylinder from motion in a dynamic random-dot field. The pattern of these results is at odds with the hypothesis that precise metrical comparison of early motion measurements is necessary for higher-order "structure from motion" tasks.

Keywords: Motion perception, Middle temporal area (MT), Clinical psychophysics, Structure from motion, Motion coherence

Introduction

The visual system of nonhuman primates is divided into several multistage cortical pathways that originate in the primary visual cortex. The most studied of these pathways involves the analysis of visual motion. A major component of the motion pathway begins in the primary visual cortex and progresses by way of the middle temporal area (MT) situated in the posterior bank of the superior temporal sulcus to the medial superior temporal area (MST), terminating in area 7a of the posterior parietal lobe. Anatomical and physiological studies suggest that MT must play a critical role in the cortical analysis of visual motion and that it may constitute a more abstract stage of processing (Maunsell & Van Essen, 1983*a*,*b*; Newsome et al., 1985; Zeki, 1974). Furthermore, there is strong experimental evidence that several behavioral tasks may rely on its outputs (Newsome & Paré, 1988; Siegel & Anderson, 1988; Adelson & Movshon, 1985). In the macaque, MT is closely associated with the occipito-parietal system (Ungerleider & Desimone, 1986) and it provides a major link between the striate cortex and the parietal lobe.

Very little is known about the anatomical structures that might mediate different aspects of visual-motion analysis in humans. The most significant functional–anatomical correlation of visual-motion processes in humans comes from neurological reports of selective visual deficits in patients with focal brain lesions.

Most notably, Zihl et al. (1983) reported findings from a patient, L.M., who presented with a dramatic deficit on discriminating motion, while visual acuity, stereopsis, color discrimination, saccadic localization, and visual fields were all normal. More recently, Zihl's initial findings in this patient have been augmented by extensive psychophysical studies (Hess et al., 1989). These studies revealed that L.M. had some residual motion perception, perhaps corresponding to an impaired "short-range" motion mechanism. On the basis of these deficits and from the initial computer tomography (CT) study showing

---

Reprint requests to: Lucia M. Vaina, Intelligent Systems Laboratory, ERB-331, Boston University, 44 Cummington Street, Boston, MA 02215, USA.

Present address of Ken Nakayama: Department of Psychology, Harvard University, Cambridge, MA 02138.

involvement of the extra striate cortex in her otherwise extensive lesion. Zihl et al. (1983) conjectured that L.M.'s lesion included the human homologue of MT. The CT study was recently supplemented with magnetic resonance imaging (MRI) studies which delineated more precisely the anatomical extent of the lesion showing tissue loss in the lateral occipital gyrus on both sides, extending bilaterally into the posterior portions of the middle and posterior temporal gyri (including the posterior superior temporal sulcus) (Hess et al., 1989).

Another line of evidence for the existence of selective motion deficits in humans comes from psychophysical studies in our laboratory of patients with focal lesions involving or disconnecting the cortical visual pathways (Vaina, 1988*a*, *b*; Vaina, 1989; Vaina et al., 1988, 1989). We found a double dissociation of function with respect to the perception of form and motion when related to the side and site of the brain lesion. Patients with lesions involving the right occipito-parietal area, but with the primary visual cortex spared, were severely impaired on motion tasks but they showed no deficits on form, color, and contrast-sensitivity discriminations. Form discrimination, however, was selectively impaired in patients with lesions in the right occipital-temporal area. On a variety of motion tasks, these patients' performance was as good as that of the normal subjects' group.

In this paper, we report the results on a variety of psychophysical motion tasks from a stroke patient, A.F., in whom MRI studies documented a lesion in the extrastriate visual areas bilaterally, extending into the posterior parietal and temporal lobes. We chose motion tasks that involve higher and lower aspects of motion processing, ones that might be differentially impaired by the presence of lesions.

The paper is organized in two main parts. In the first part, the Background Examination, we describe A.F.'s clinical, neurological, neuro-ophthalmological, and neuropsychological background and we document his performance on several psychophysical nonmotion visual tasks.

In the second part, we relate A.F.'s performance on motion tasks. In particular, we examined his ability to achieve segregation by motion contrast in three types of random-dot kinematograms. We also report on A.F.'s ability to compute and use average velocity measurements in a task of global speed discrimination and to perceive a global motion field in a task of discrimination of direction of motion in a dynamic random-dot display (Newsome & Paré, 1988). Finally, we examined the subject's ability to perceive rigid three-dimensional structure and nonrigid "biological motion" (Johansson, 1973) in dynamic random-dot patterns.

Background examination

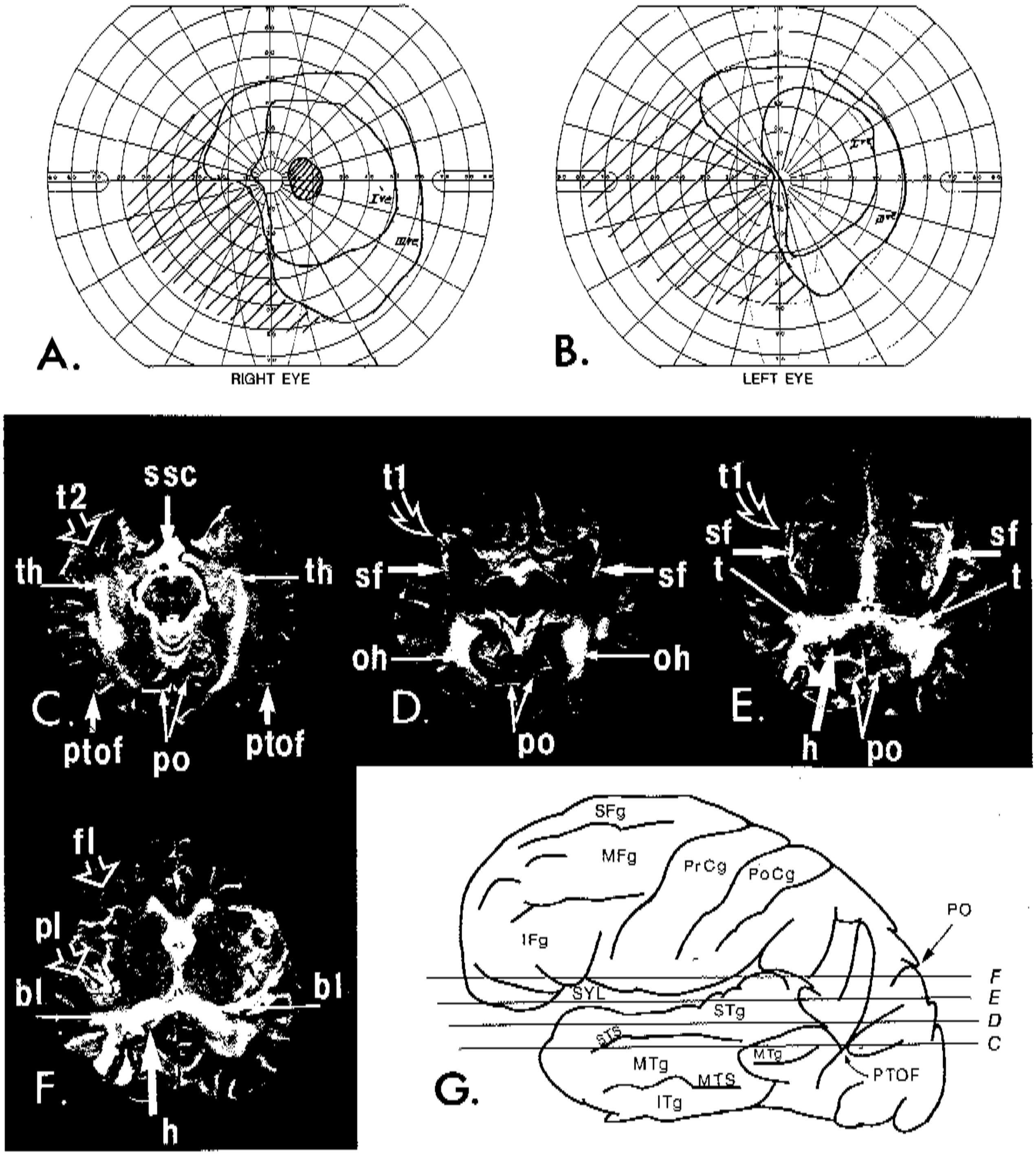
A 60-year-old left-handed man, A.F., was admitted to the New England Rehabilitation Hospital for recuperation and for the evaluation of visual and cognitive deficits following a new acute hemorrhage in the posterior right hemisphere which required emergency hospitalization and subsequent evacuation of a large hematoma. While in the rehabilitation unit, A.F. underwent detailed neurological, neuro-ophthalmological, and neuropsychological evaluations. A CT scan performed in April 1988, at the time of his recent strokes, showed a hemorrhagic infarction involving the medial margin of the occipital horn of the right lateral ventricle and extending into the parietal lobe between the proximal portion of the occipital horn of the lateral ventricle and the margin of the tentorium. The hemorrhage extended superficially into the medial parietal region. There was considerable low density about the area of hemorrhage consistent with edema.

An MRI study performed three months later (July, 1988) showed areas of increased signal on T2 weighted images, consistent with lesions involving the temporal-parietal-occipital junction bilaterally. The right-hemisphere lesion was larger than the left but both extended dorsally into the posterior parietal lobes. Figure 1C,D,E,F document in detail the area of bilateral tissue loss. The left-hemisphere lesion shows tissue loss from an old infarct extending along the lateral margin of the lateral ventricle and around the margins of the occipital horn, going upward posterior to the body of the lateral ventricle and into the

---

FACING PAGE

Fig. 1. Visual fields and magnetic resonance imaging of the patient's brain three months following a hemorrhagic stroke and the localization of the axial scans on the lateral view of a human brain. Figures A and B show A.F.'s visual fields by perimetry. A.F. has a bilateral congruous loss of the left inferior visual field and a minimal loss in the upper visual field. C–F are the relevant slices in axial view of T2 weighted magnetic resonance imaging studies (Tr/2000 ms/TE80 ms). Labeling; bl: body of the lateral ventricle; h: hemosiderin; fl: frontal lobe; oh: occipital horn of the lateral ventricle; pl: parietal lobe; po: parietal-occipital sulcus; sf: Sylvian fissure; ssc: supersellar cistern; t: trigone of the lateral ventricles; t1: superior temporal gyrus; t2: medial temporal gyrus; ptof: parietotemporo-occipital fossa (see G). C: The picture shows the temporal horns of the lateral ventricles and some patchy hyperintensities at their margins, more on the right. D: The picture shows local tissue loss in the right hemisphere at the sites of the hemorrhagic stroke in the parietal lobe just medial to the occipital horn of the lateral ventricle and anterior to the parietal occipital sulcus. There is also some patchy hyperintensity in the temporal lobe along the lateral margin of the right occipital horn. Localized hyperintense areas are seen in the left temporal lobe adjacent to the lateral proximal margin of the occipital horn and in the margins of the occipital horn. Images E and F show bilaterally larger patchy areas of hyperintensity posterior, medial, lateral and above of the trigones, and at the margins of the bodies of the lateral ventricles. Tortuous narrow bands of hemosiderin (labeled h) are seen in both images at the site of the recent hemorrhage. G: A schematic drawing of adult human brain showing the major gyri and sulci in lateral view. Labeling: IFg: inferior frontal gyrus; ITg: inferior temporal gyrus; Mfg: middle frontal gyrus; MTg: middle temporal gyrus; PoCg: postcentral gyrus; PrCg: precentral gyrus; SFg: superior frontal gyrus; STg: superior temporal gyrus; STS: superior temporal sulcus; MTS: middle temporal sulcus; PO: parietal occipital sulcus; SYL: Sylvian fissure; and PTOF: the parietotemporo-occipital fossa from Polyak (Fig. 271, Polyak, 1957) which several authors suggested to correspond to the human homologue of the macaque MT; th: temporal horn. The horizontal lines correspond to the levels of the axial scans corresponding to C–F above.


A.
RIGHT EYE
I ve
III ve
B.
LEFT EYE
I ve
III ve
ssc
t2
th
th
C.
ptof
po
ptof
t1
sf
sf
oh
oh
D.
po
t1
sf
sf
t
t
E.
h
po
fl
pl
bl
bl
F.
h
SFg
MFg
PrCg
PoCg
IFg
PO
SYL
STg
STS
MTg
MTg
MTS
ITg
PTOF
F
E
D
C
G.

medial portion of the parietal lobe. This lesion was chronic and resulted from a possibly silent infarct that was not readily seen in the earlier CT scan.

In the right hemisphere, there were irregular narrow areas of decreased signal, consistent with hemosiderin from the recent hemorrhage (April, 1988) extending into the medial and anterior portions of the parietal lobe. Large patchy areas of increased signal were consistent with tissue loss lateral, posterior, and medial to the trigone of the lateral ventricle which extended upward behind the body of the lateral ventricle.

In addition to these large areas, there were also multiple smaller foci of increased signal in the deep white matter by the bodies of the lateral ventricles, in the white matter of the frontal lobes, the basal ganglia, and a few small foci in the pons. These are consistent with small-vessel ischemic changes.

The patient's medical history was remarkable for long-standing untreated hypertension. He had no paresis or significant weakness. Although arm movements were intact and he could perceive clearly and recognize objects in the visual field, he was hesitant and moderately inaccurate on touching objects placed in his reach (more in the left visual field than in the right) which, however, he could clearly see and recognize visually. Reading was difficult and he was slow and impaired on written, but not oral, simple calculations. These difficulties were entirely derived from his inability to locate letters or numbers on the page under visual control. Thus, for example, printing in upper case letters was very disorganized (Fig. 2E) when he wrote looking at the page, but he wrote without difficulties when blindfolded (Fig. 2F), indicating that the ability to write *per se* was not impaired. As expected, writing cursively with lower case letters was good, since this does not require specifically visual positioning of each individual letter. Copying (Fig. 2C) and drawing (Fig. 2B) were poor and lacked perspective. He was impaired on judging lengths of lines and he was unable to bisect a line in the middle. Informal testing showed that monocular depth was not grossly impaired.

The neuro-ophthalmological examinations included the measurements of visual fields, eye movements, letter and vernier acuity, contrast sensitivity, and color discrimination. Visual field charted by Goldmann perimetry revealed a dense congruous loss of the left inferior visual field bilaterally (Fig. 1A and 1B: visual fields) consistent with the anatomical locus of right-hemisphere lesion. There was also a minimal loss in the upper visual field. Letter acuity with correction glasses was 20/30. Vernier acuity as tested was excellent. He was able to discriminate Vernier displacement of 17 arc s with 80% success. On the examination of ocular motility, there was no strabismus and saccadic eye movements to static targets were normal both to the right and left. He substituted saccades for smooth pursuit when following a smoothly moving target. This occured to the right and to the left. He was unable to hold fixation in the left lateral gaze and tended to drift back towards the central fix-

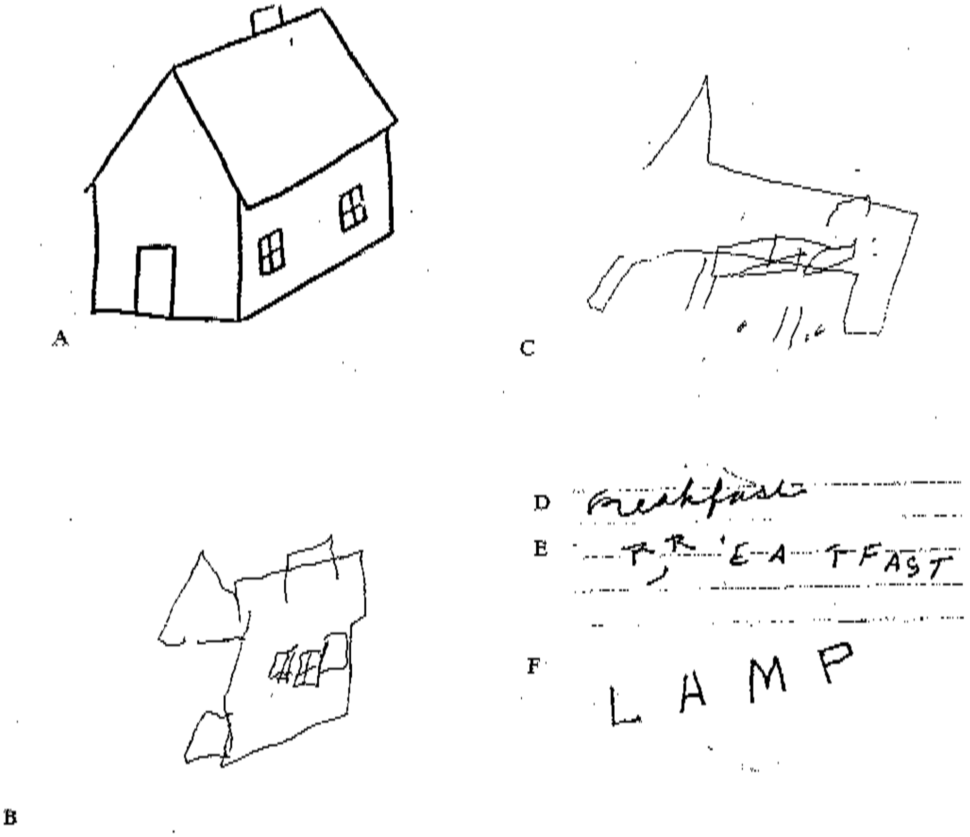


Fig. 2. Drawing and writing. A: The stimulus is a simple house. B: Picture of A.F.'s drawing of a house from memory. C: Pictures of A.F.'s copying of a house (presented in A). D: Examples of A.F.'s writing. On the first row is an example of his cursive writing, and on the second row is an example of printing on a page under visual guidance. The third row shows printing again in blindfolded condition.

ation. When making saccades to the right, he could hold fixation in the right field of gaze and appropriately make fast saccades back to central fixation.

The following tasks of visual discrimination were performed: contrast sensitivity, color and shape discrimination, spatial location, assessment of binocular disparity, and judgement of depth.

Visual-discrimination tasks

Contrast sensitivity, evaluated with the Vistech 6500 chart[*] (Ginsburg, 1968,1983) at spatial frequencies of 1.5, 3, 6, 12, and 18 cycles/deg, was normal (Fig. 3A).

Color and grey levels discrimination. The ability to discriminate color was evaluated with the Farnsworth-Munsell 100-Hue test. This test has only 88 hues in four groups of 22. A group of 22 is presented in random but predetermined order in a single row. The task is to sort them so that there is an orderly progression of hues along the row between a predetermined anchor hue at each end. A.F.'s performance on this task was normal (Fig. 3B). He scored at a normal level on a task of discrimination of neutral grey values. This is a standard scale of neutral greys (Munsell Color Corporation, 1987) ranging between absolute black (0% reflectance) and absolute white (100% reflectance) in steps that look equal to the eye in standard viewing conditions.[†]

Shape discrimination. This task closely replicated the Efron's "square" task (Efron, 1968; Warrington & James, 1988). The observer is instructed to keep his eyes fixed on a small fixation mark displayed either to the right or left of the stimulus. The target stimulus was a square 5 × 5-deg and the distractors were oblongs with an area chosen from one of the following dimensions: 5.25 × 4.77-deg, 4.6 × 5.5-deg, and 6.5 × 4-deg (Fig. 3C). For each of these dimensions, ten squares and ten oblongs were presented for 0.5 s, singly in pseudorandom order.

This was a two-alternative forced-choice task and chance was considered at 50% correct responses. The task is to judge whether a shape is a square or an oblong. A.F.'s performance on this task was normal (19/20 correct responses) for the most difficult discrimination (Fig. 3C) and 20/20 choices for the other two.

Spatial localization. The stimuli consists of two squares (Fig. 3D), one with randomly placed symbols (single-digit numbers and letters) and the other with a black square-shaped mark subtending 5.6-arc-min square (MacQuarrie, 1953). The subject's task is to identify the symbol in the top square corresponding to the position of the black mark in the bottom square. A.F. scored 50% correct (10/20), which was at a severely impaired level (Fig. 3D) (his scores were below the 5% percentile scores obtained in an age-matched control group (Warrington & James, 1988)).

Binocular (global) stereopsis. The stimuli consist of a series of random-dot stereograms from the series devised by Julesz (Julesz, 1971). When viewed monocularly, each target appears as a random array of small light and dark squares in which no form or depth is apparent. The two random-dot patterns are identical except that in one pattern a contiguous cluster of dots in the central region has been displaced laterally with respect to the same region in the other pattern. Two disparities were used corresponding to 250″ and 500″. There are no monocular cues to depth in these patterns but with binocular fusion a central figure stands out in front of the surround. As in Julesz (1971), the patterns were printed in pale red and green ink. When viewed with a green filter over one eye and a red filter over the other, the central figure is seen in front of (or behind) the surround. We used five random-dot stereograms. The subject's task was to say whether he saw a central figure standing away from the surround, and then to identify the shape of the figure. All of the figures were simple geometric forms. For subjects unable to identify the forms verbally, silhouettes of four possible matching choices were shown, and they were asked to choose the shape that corresponded to the figure presented stereoscopically.

A.F. lacked binocular stereopsis at both disparities tested. He reported seeing in the random-dot stereograms "just dots, no pattern or form at all." Interestingly, A.F. commented that prior to his recent stroke, he enjoyed resolving anaglyphs and that years ago he used to watch three-dimensional movies.

Depth perception (local stereopsis). The depth-perception apparatus (Howard, 1919) consists of a box 60 cm long which had inside two illuminated parallel-vertical rods about 1 cm thick and with a visible length of about 8 cm. The rods, positioned 4 deg apart, were viewed through a slit that extended the widths of the exposure field. One of the rods was stationary and the other was set in prearranged order of 3, 7, 10, 15, 22, 30, 35, 40 mm in front or behind the stationary rod. The subject was required to judge whether the two rods were equally distant from the viewer. A pseudorandom schedule determined whether the right or left rod was nearer on each trial. There were 24 trials, three trials for each of the eight distances.

A.F. was unable to discriminate distances between the rods smaller than 35 mm, which is at a grossly impaired level since the normal controls showed no deficits in any of the trials that are in approximate size 3 mm.

We do not think that A.F.'s loss in stereopsis can be attributed to disordered vergence eye movements or strabismus because, as described earlier, neuro-ophthalmic examination

---

[*] The Vision Contrast Test System (VCTS) consists of six rows of 1-cm-diameter patches of gratings obtained from photographing computer-generated sine-wave gratings, each row having patches of different spatial frequency: 1.5, 3, 6, 12, and 18 cycles/deg (Ginsburg, 1983). Each row contains, from left to right, one sample patch and eight test patches. The contrast of these patches ranges from zero contrast to contrast above and below threshold in roughly 0.1 log unit steps. The gratings are tilted in one of the three orientations: −15, 0, +15 deg. The contrast and orientation of test patches are randomized for each row to control for guessing. The observer holds the chart at 56-cm distance positioned vertically at eye level and contrast sensitivity is checked for each eye separately. The observer is shown a high-contrast sample grating at the three possible orientations. The task is to determine whether each sequential grating patch on a specific row was blank or at a particular orientation. The procedure is repeated for the remaining rows of grating patches. The test method is a four-alternative forced-choice procedure. Contrast-sensitivity threshold is determined from an answer key provided by VCTS on the basis of an extensive data base of results.

[†] The Munsell Neutral Values test is presented as 31 matt patches of paper $5 \times 5\ \mathrm{deg}^2$ whose reflectance differed by maximum 6%. The patches were viewed through an oblong white mask with an aperture of 2 × 3-deg which provided a constant matt background. The task was to sequence the patches of gray in order, from the most white to the most black.

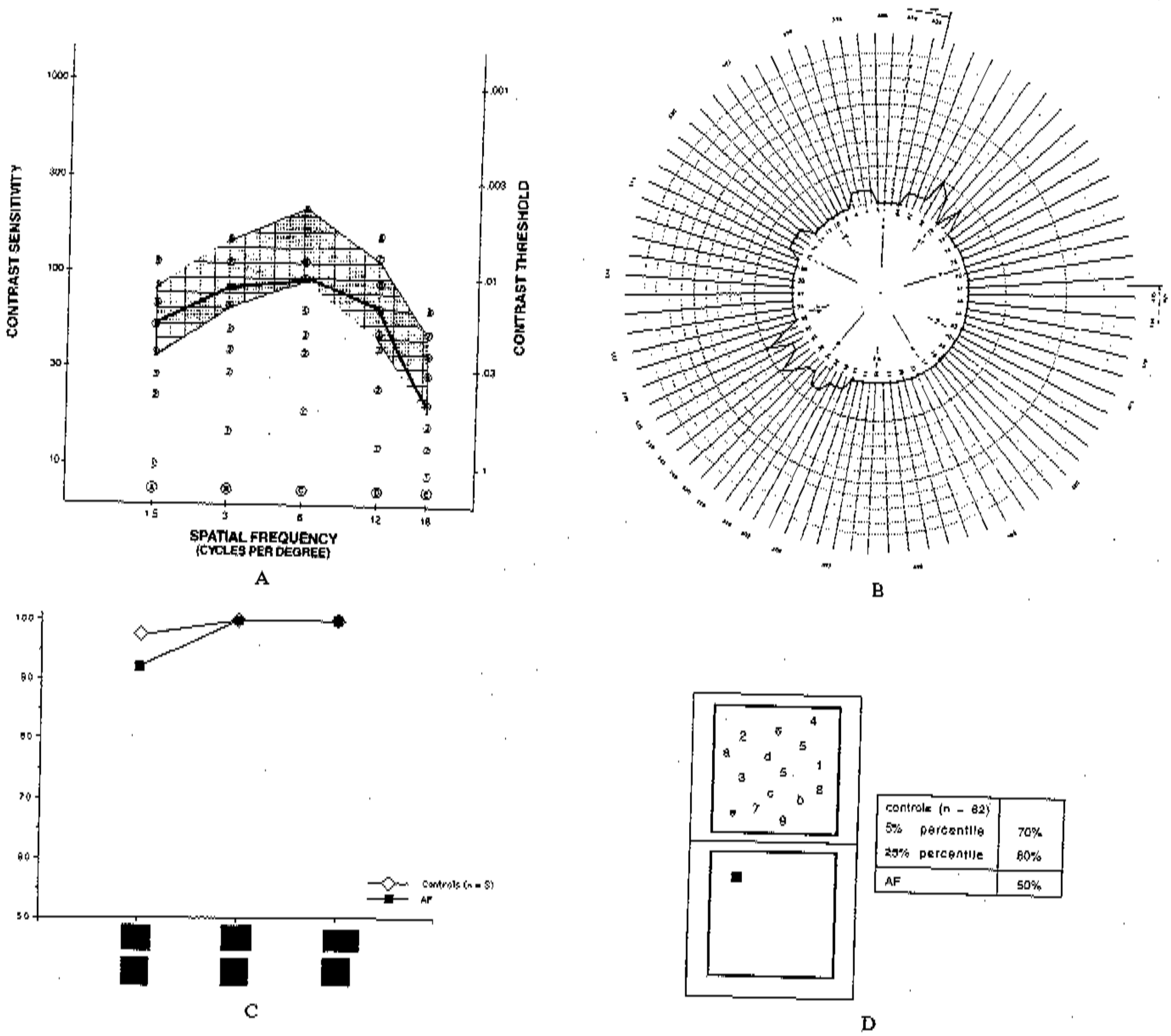


Fig. 3. Results on non-motion visual tasks. A: Results on the Vistech Contrast Sensitivity Test: The figure is a plot of foveal contrast sensitivity for pattern detection as a function of grating spatial frequency. The spatial frequency varied between 1.5–18 cycles/deg. The hatched area indicates the corresponding mean control data as presented by Vistech Inc. (Ginsburg, 1983), and the continuous line shows A.F.'s results which fall within the normal range. B: Results of the Farnsworth-Munsell 100 Hue Test. Errors are plotted radially and positions on the circumference represent points along a continuum of hue. A.F. scores were normal. C: Shape Discrimination Task. Percent correct performance of five normal controls and A.F. on shape discrimination for three ratios plotted as decreasing difficulty. D: The spatial localization task. An example of the display is presented on the left side. The target is the black square in the bottom box. The observer's task is to choose the corresponding spatial location from among the numbers and letters in the top box. On the right are presented the percentage of correct responses in the spatial localization task of A.F. as compared with those of the worst 5% and 25% responses of a large group of normal controls.

indicated that he had no strabismus or any sign of oculomotor imbalance.

Neuropsychological evaluation

The neuropsychological examinations addressed the patient's perceptual, cognitive, and motor abilities.

Evaluation with the Wechsler Adult Intelligence Scale-Revised (WAIS-R) showed an average verbal IQ of 104 and a severely depressed performance IQ of 68. His perceptual skills were severely impaired as most strikingly shown on a task of visual recognition of objects photographed from noncanonical angles. This task, adapted from the test originally devised by Warrington and Taylor (1973), is very sensitive to right posterior parietal lobe damage. A.F. could recognize all these objects when presented as photographs taken from the usual, canonical view. When presented with pictures of noncanonical views of a bucket or a basket, for example, A.F. described them as "*a* *slice taken from the trunk of a tree,"* or a *glove* he described as an *"overshoe or a boot"* and a *frying pan* as a *"magnifying glass."*

A.F. had satisfactory performance on tasks that addressed the visual semantic level such as recognition of pictures of objects and actions, functions of objects, and matching textures to silhouettes of objects (Vaina, 1987). Spontaneous speech, repetition, auditory comprehension, and prosody were intact. Naming of definitions, tactile naming, and naming of environmental sound were also normal.

Discussion of nonmotion tasks

A.F. scored in the normal range on tasks of shape, color, and contrast discrimination. These results strongly suggest that the form and color pathways were not involved in the deficits manifested at the sensory processing level.

He was unable to bisect lines in the middle, appreciate lengths of lines, and scored at a severely impaired level on formal and informal tasks of spatial localization under visual control. Global and local depth perception from binocular cues was severely impaired. His drawings were poor and spatially disorganized (Fig. 2) and he had severe difficulties recognizing objects presented in noncanonical views.

The constellation of the above visual deficits is suggestive of the syndrome of *visual disorientation* described by Gordon Holmes (Holmes, 1918; Holmes & Horax, 1919) in several detailed and careful studies of World War I veterans wounded symmetrically in both posterior parietal regions. In the cases in which the anatomical site of the lesion was known, the lesions involved the upper fibers of the optic radiations beneath the supramarginal and angular gyri. These patients presented severe defects in visual orientation, judgement of size and relative distance, visual localization of objects, and manual reaching. One of these patients (Holmes, 1918), similar to A.F., also lost stereoscopic vision and was unable to judge depth.

Visual-motion perception

Both the patterns of visual loss and the combined CT and MRI findings indicate damage to the posterior parietal cortex in A.F. As such, it might be expected that cortical area MT as well as other motion-processing areas would be undercut. Consequently, we hypothesize that A.F. should show losses in motion perception. To delineate the severity of such possible losses, a battery of psychophysical tasks was designed to pose a graded series of challenges to A.F.'s motion system. These tasks can be divided into three broad categories.

First, we followed the lead of Julesz (Julesz, 1971; Braddick, 1974) to delineate the short-range process by seeing whether A.F. could see form or boundary from differences in motion in neighboring spatial regions. In three tests of increasing difficulty, we ask whether perceptual segregation can occur between regions having motion-*vs.*-no-motion (experiment 1a), regions having differences in velocity direction (experiment 1b), and differences only in velocity magnitude (experiment 1c).

Second, we evaluated the ability of A.F. to integrate motion information over a larger retinal area by evaluating his ability to precisely encode the average speed of a cluster of dots moving in random directions (experiment 2a), and his ability to detect coherent motion defined by a small number of dots in noise (experiment 2b).

Third, we addressed the issue as how A.F.'s visual system might employ motion information to perform higher-order visual tasks, in particular, whether he had the ability to see a rotating three-dimensional cylinder from a two-dimensional dot display (experiment 3a) and whether he would be able to identify different patterns of biological motion simply from the trajectories of moving points (experiment 3b).

General methods and procedures

All stimuli (except in experiment 3b) were generated and presented, and responses collected and analyzed, using a Macintosh IIcx computer with an extended 8-bit video card. The stimuli were presented in the center of the Macintosh standard RGB monitor with a resolution of 640 × 480-pixel frequency of the vertical retrace interrupt of 66.7 Hz and P4 white phosphor.

Random dots were used to minimize familiar position cues and to isolate motion mechanisms (Nakayama & Tyler, 1981). In all of the experiments, each picture element (pixel) subtended 1.8 × 1.8 arc min at the viewing distance of 65 cm. The background in the display was black and the random dots were painted white. In experiments 1a–c, we used dense random-dot pattern and each pixel in the raster dot had a 50% probability of being light or dark. Viewing time in all of the experiments was 2 s for each trial.

A.F. and control subjects were first familiarized with the display and task through examples and feedback. Experimental sessions followed, during which response feedback was not provided. Each trial was initiated by the subject by pressing a specially designated key. The room illumination was maintained at a low photopic level, and the observer was instructed to restrict fixation to the center of the screen.

The control group consisted of normal volunteers without any known ophthalmological, neurological, or psychiatric disorder and matched by age with A.F. A substantial fraction of the group consisted of the spouses of other patients who have been tested in the laboratory. They were unexperienced as subjects, this being their first participation in a psychophysical experiment. All of the subjects had corrected to normal vision.

Low-level motion: Form and boundary perception from motion differences

Because our preliminary results indicated that A.F. had intact two-dimensional form discrimination, we used the form-from-motion task introduced by Julesz (1971) to evaluate early motion processing. Psychophysical studies of motion segregation have demonstrated that the human visual system is capable of detecting motion discontinuities, when the background is static (Julesz, 1971) or when the figure and the background move at different velocities (Baker & Braddick, 1982). The extraction of the objects' boundaries appears to be mediated by a *short-range process* (Braddick, 1974), a visual mechanism that matches up corresponding local pattern elements of the same luminance polarity in successive time frames and operates over short time and spatial intervals.

Experiment 1a: Two-dimensional form from motion in a static background

In this experiment, the sensation of two moving textured planar surfaces is elicited by two patches of contiguous random dots uniformly displaced from one frame to the next in translational motion across the random-dot display which formed the stationary background (Fig. 4A). The forms were defined solely by the relationship of the displacement between the moving patches and the static surround. The moving patches had one of the following shapes (illustrated in Fig. 6A): square, circle, triangle, cross, or oblong (oriented horizontally or vertically). The square, circle, and the cross had roughly the same area, and the oblong was half of the square area.

The display subtended 10 × 10-deg and two moving forms, each covering approximately 2.2 × 2.2-deg of visual angle, were shifted by roughly 3.5 arc min between successive frames, thus the resulting speed of the motion was 0.8 deg/s. This displacement gave a vivid impression of two forms translating in opposite directions across the static random-dot background. The interstimulus interval between two consecutive frames was in all cases about 70 ms.

The subject's task is to make a simple two alternative, forced-choice judgement concerning whether the two shapes moving on the screen were the same or different. There were 40 trials, and chance was at 50% correct responses. The interval between trials was 1 s. Figure 4B shows the performance of A.F. on comparison to normal control observations. Note that his performance is excellent and is statistically indistinguishable from the six age-matched controls. This lack of any deficit seen for A.F. is not surprising, given the results of earlier reports from this laboratory. In these studies (Vaina, 1989; Vaina et al. 1988, 1989), it has been shown that perceived shape of a moving figure against a stationary background does not distinguish motion deficient from normal subjects. The above reinforces our views that A.F. does not have a significant form-recognition problem in random-dot displays, that he can pick out the correct figure from a display which would have no structure were it to remain stationary.

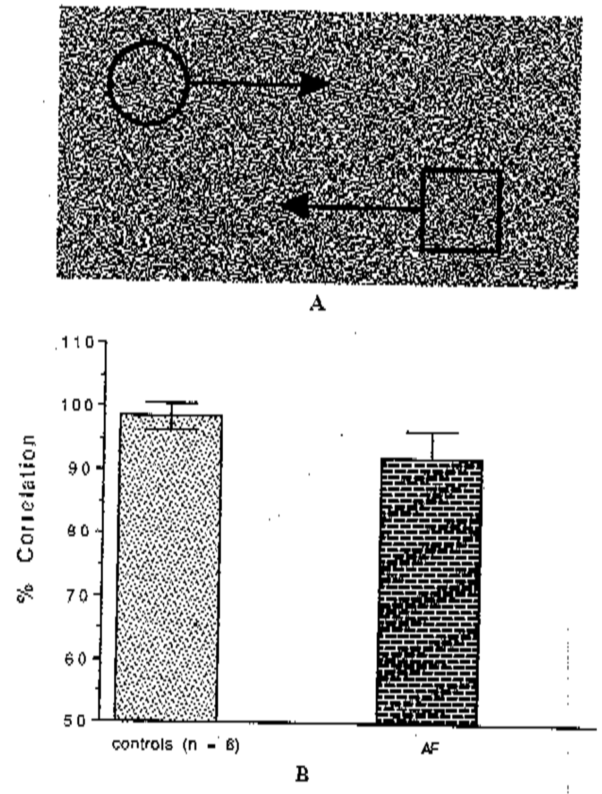


Fig. 4. Two-dimensional form from motion in a static background. A: The display consisted of a fine grained (pixel size = 1.72 arc min) random-dot pattern resulting from 50% black and 50% white dots. Two patches of contiguous dots were coherently translated across the static background and with equal speeds. The observer's task was to determine whether the two translating two-dimensional shapes were the same or different. B: Percent correct responses as a function of correct matches of the moving shapes. The graph shows the mean percent of correct responses and error bars for A.F. and mean percent of correct responses and standard deviations for the normal control group.

Experiment 1b: Boundary localization

Conceivably more difficult than the two-dimensional form from motion in static background case is the situation where both regions are moving but the directions of motions are different. Such a case is examined in the next experiment, an evaluation of boundary detection where the regions are moving at the same velocity magnitude but differ in direction only.

The stimuli (similar to those used by Hildreth, 1983) were dense dynamic random-dot patterns subtending 12 × 8 deg of visual angle. In any given experimental trial, a sequence of 50 frames was constructed in such a way that there was a discontinuity in the velocity field along a vertical line (Fig. 5A). Four angular differences (18.4, 37.1, 90, and 180 deg) were used. Along the left side of the boundary was a 1.4 × 1.4-deg notch whose distance from the point of fixation varied along the y axis from trial to trial but remained within 2-deg visual angle above or below a black fixation mark. The vertical boundary and the notch were entirely defined by the difference in velocity to the left and right of the boundary and were not visible in any static frame.

The subject was instructed to maintain fixation at the center of the display indicated by the black fixation mark present in the first frame of the sequence. On the subject's signal, the additional frames appeared one after another, in rapid succession. The experimental session consisted of 20 trials for each condition tested. This was a two-alternative, forced-choice (2-AFC) task in which the subject was required to determine whether the notch was above or below the black fixation mark.

Figure 5B shows that the normal subjects performed the task essentially without error for all direction differences tested. A.F.'s scores were also in the perfect or near-perfect range except for the smallest direction difference of 18 deg, indicating that angular direction differences of 37 deg or more were sufficient to perform an accurate boundary localization. As such these results indicate that delineation of image regions by relative motion remains essentially intact except from the smallest difference in directions.

Experiment 1c: Two-dimensional form from differences of velocity magnitudes alone

Perhaps the most difficult motion task is that which requires a comparison of velocity magnitude without any difference in direction. Nakayama (1985), for example, has summarized recent work indicating that velocity magnitude is encoded with 3–4 times less precision than velocity direction. For this reason, we

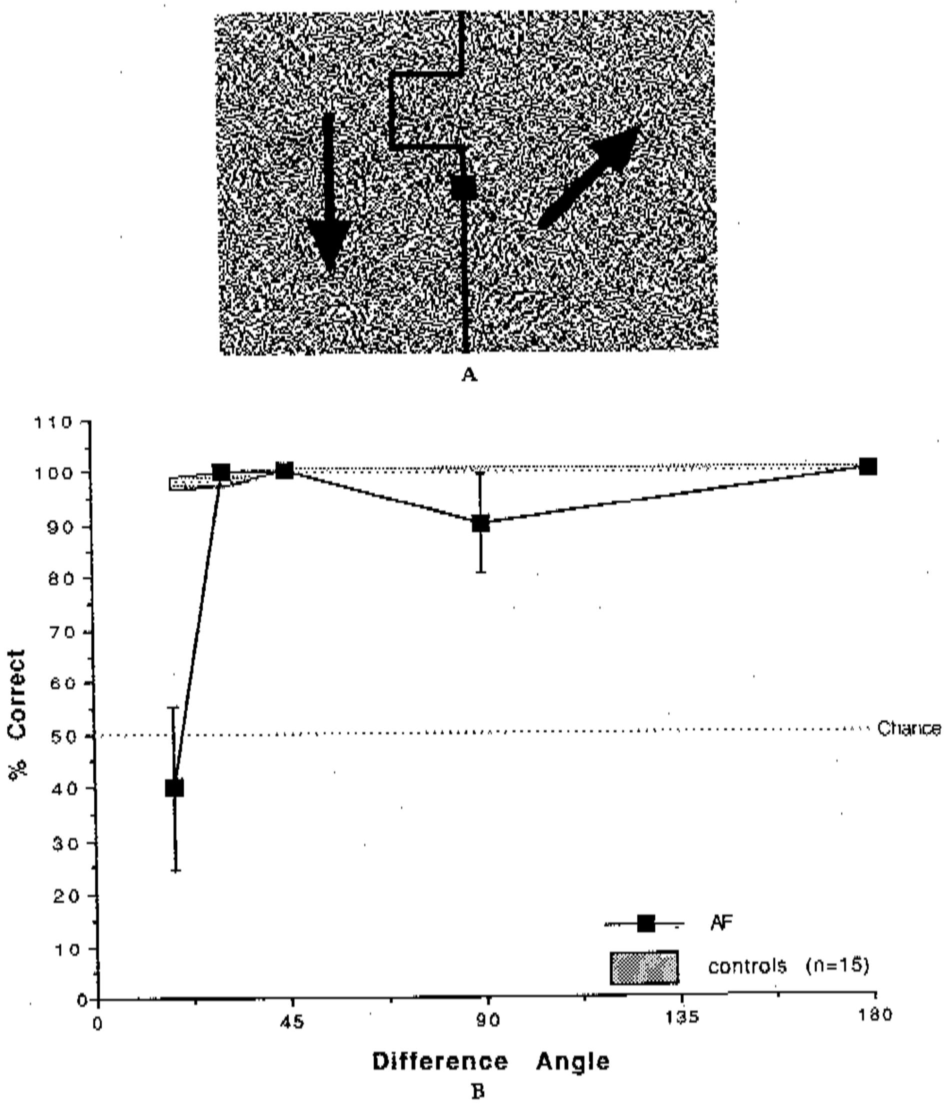


Fig. 5. Boundary localization. A: The random dot pattern characteristics are the same as in Fig. 4A. The differential movement of the two adjacent areas of the pattern reveals a boundary with a rectangular notch defined solely by angle of differences in the directions of motions of the two areas. The position of the notch and of the invisible dividing line between the two areas in relative motion varies randomly within 1 deg on the x axis. The heavy square indicates the fixation mark. B: Percent correct responses for correctly identifying the location of the notch (above or below the fixation mark) for the following angular differences between the two areas: 18.4, 37.1, 90, and 180 deg. The black squares and the error bars indicate the percentage of correct responses of A.F. for these angular differences. The hatched area indicates the percentage of correct responses of the control group ± 1 S.D. obtained under identical conditions.

employed a task where observers were required to recognize two-dimensional (2D) shape from velocity magnitude differences alone.

An example of the display is portrayed in Fig. 6A. In each trial, the figure subtended an area of roughly 2 × 2-deg and it was randomly positioned within a 6 × 3.75-deg rectangular aperture of dense random-dot pattern. Both the figure and the background were displaced in the same direction but with different speeds. In 60% of the trials the background speed (displacement of 3.5 arc min between two frames) was slower than the figure, and in 40% of the trials the background was faster (displacement of 8 arc min between frames) than the figure. The speed ratios between the figure and the background, using as a reference the speed of the background, had the following values: 3/2, 2, 3, 2/3, and 1/2. In all trials, the interframe interval was roughly 70 ms and the maximum displacement between two frames did not exceed 8 arc min.

Observers were asked to make a six-alternative forced-choice judgement indicating which of the six figures displayed at the bottom of the screen constituted the correct match for the moving shape. There were 60 trials and each one of the six shapes was the correct answer 10 times. Chance level was considered at 17% correct responses.

The results for this experiment are shown in Fig. 6B. It should be clear that the normal control group had near-perfect performance, approximately 95% correct. This stands in con-

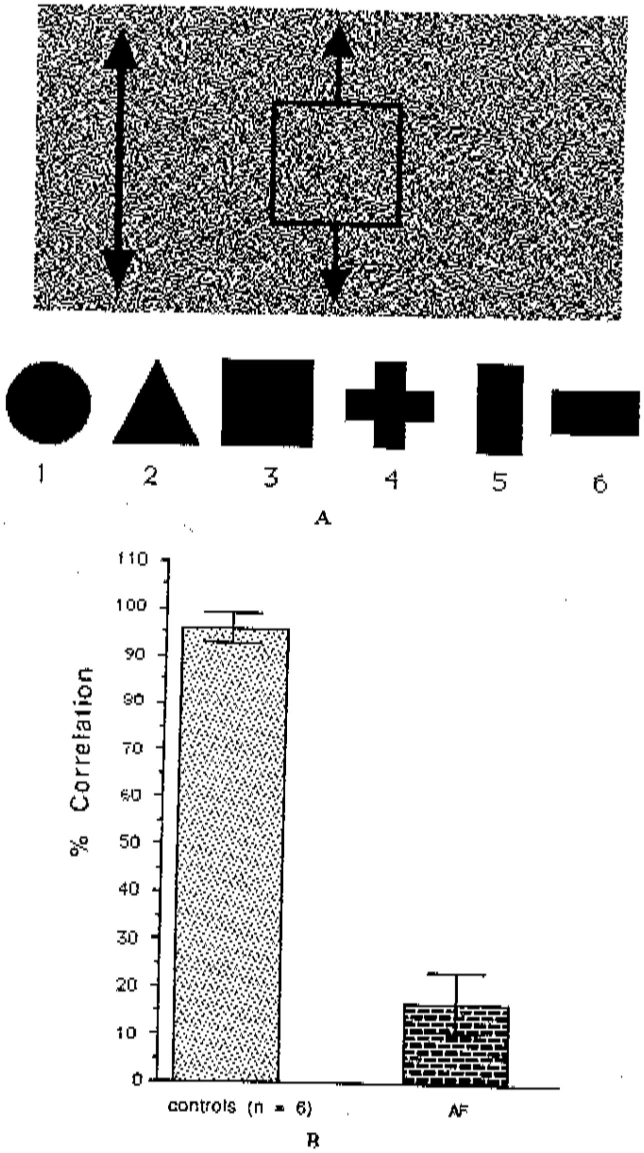


Fig. 6. Two-dimensional form from differences of velocity magnitudes alone. A: Illustration of the stimulus. The random dot pattern characteristics are the same as in Fig. 4A. The background scrolls vertically and a central patch of contiguous dots subtending 2 × 2 deg moves in the same direction as the background with different speed. The shape of the central patch is randomly assigned to one of the six shapes illustrated on the bottom of the display. B: The bar graphs plot percentage correct responses of the total number of trials for A.F. and the normal controls together with the standard deviation for the normal group and the standard error of A.F.

trast to A.F.'s very poor score of 17% correct responses, which is at chance.

Encoding average speed and coherent motion in dynamic displays

Experiments conducted so far have employed stimuli in which local motion was highly coherent. In the next two experiments to be described, the stimuli were very different, consisting of rapidly fluctuating dynamic dot displays. Here the observer is confronted with motions that are locally incoherent and where the task is to judge the average speed of the overall cluster of moving dots or to detect motion of a small number of coherently moving dots within this dynamic field.

Experiment 2a: Local speed discrimination

Experiment 1c indicated that A.F. was profoundly deficient in seeing two-dimensional form from differences in velocity magnitude. To examine A.F.'s possible "speed" deficit more directly using a very different technique, we measured the ability to judge the average speed of a cluster of random dots.

The stimuli consisted of two sparse dynamic random dot cinematograms each comprising 20 computer-generated dots. The cinematograms were displayed in two rectangular apertures arranged one above the other, each subtending an area of 4 × 2.5 deg, thus giving a dot density of 2 dots/deg^2 (Fig. 7A). The distance between the centers of the apertures was 2.75 deg. In any single trial, each dot took an independent, two-dimensional random walk of constant step size defined by the speed. The direction in which any dot moved was independent of its previous direction and also of the displacements of the other dots. The speeds of the dots, defined as a function of the distance a dot was displaced between successive frames, was uniform within a box and was assigned independently for each box. A base speed of 4.95 deg/s was always compared to five other speeds, giving five speed ratios of 1.1, 1.47, 2.2, 3.6, and 5.5. The assignment of the highest speed to the top or bottom aperture was pseudorandomly selected.

Each frame was on for 66 ms with no interframe interval. A "wrap-around" scheme was used, in which dots displaced beyond the boundary of the aperture in the next frame reappeared on the opposite side.

Subjects were asked to determine which of the two apertures contained the faster moving dots. A two-alternative forced-choice (2-AFC) procedure was used for measuring the subject's ability to detect difference in speed.

In comparison to the age-matched control group, who were performing almost perfectly for 2:1 speed differences, A.F. had a severe deficit on this speed discrimination task. He failed to discriminate reliably speed ratios smaller than 5.5 and was performing at near chance when normal observers were performing better than 90% correct.

Experiment 2b: Motion coherence

Some of the most direct indications of MT's critical role in motion perception come from recent work in nonhuman primates. In particular, Newsome and Paré (1988) have found that selective lesions of MT reduced motion sensitivity. Newsome et al. (1989) have found that the activity in single MT cells correlates directly with the monkey's trial-to-trial performance in a motion-discrimination task and finally Newsome et al. (1989) showed that microstimulation of tiny regions of MT can bias a monkey's motion perception in a manner predictable from the response characteristics of cells adjacent to the stimulating electrode.

To begin the process of making a direct link to this important work, we compare A.F.'s performance to that of normal controls in a task essentially identical to that used in the monkey studies.

The stimuli were dynamic random-dot cinematograms with a correlated motion signal of variable strength embedded in a

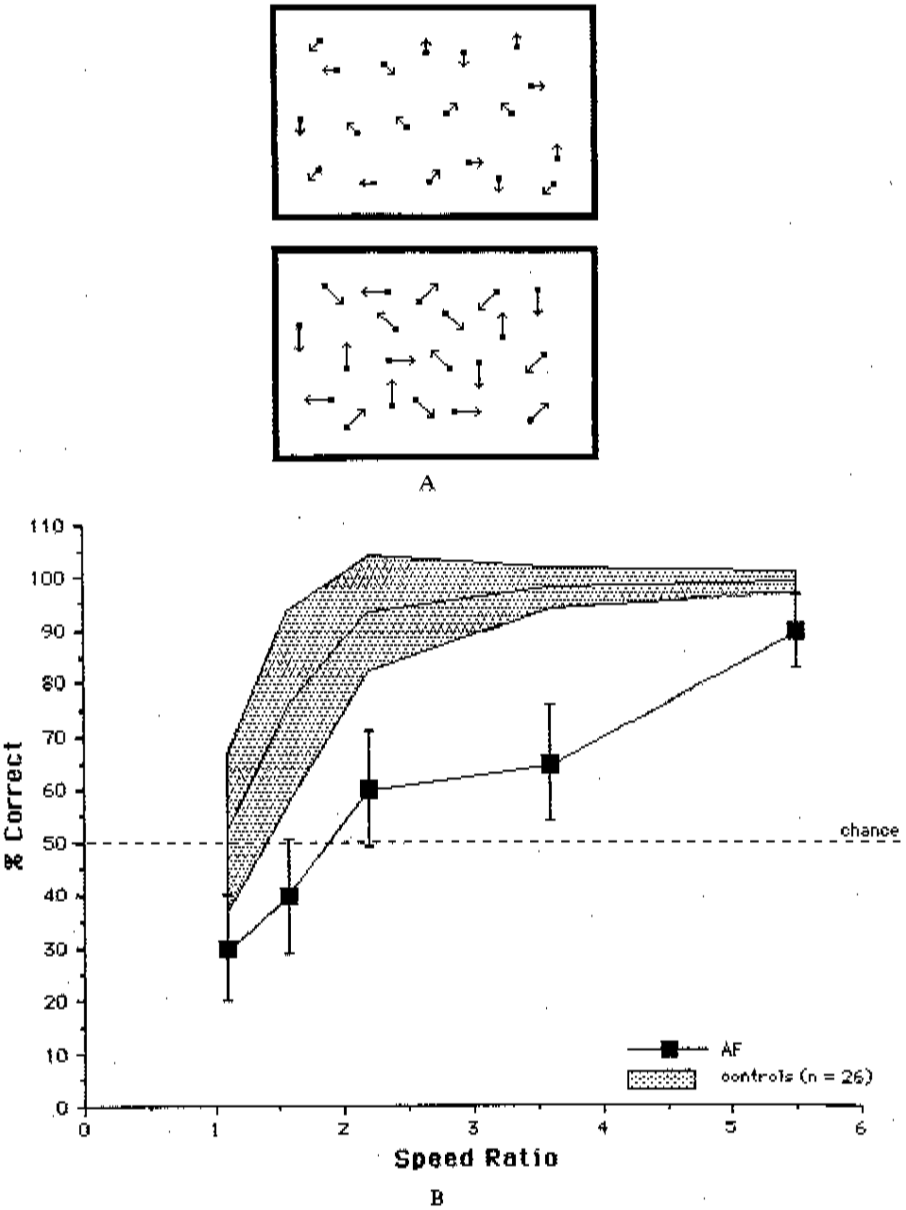


Fig. 7. Local speed discrimination. A: A schematic representation of the stimuli employed. The display consisted of two sparse random dot fields each displayed in a rectangular aperture subtending 4 × 2.5 deg. Within each aperture the dots, plotted here as vectors, move in random distributed directions with constant speed. The observer's task is to decide which of the two apertures move faster. B: Results for A.F. and normal controls. The graph plots the percentage of correct answers as function of the speed ratios between the two apertures. The data for the normal control group are presented as a shaded area representing mean ± 1 S.D. A.F.'s data present the mean and the standard error.

motion noise. The strength of the motion signal, that is the percentage of the correlated dots moving in the same direction, varied from 0–100%. For intensity less than 100%, the dots carrying the motion signal were embedded in a field of dots moving in random directions. At 0% there was no motion signal (Fig. 8A, left), and at 100% the correlation (Fig. 8A, right) corresponded to a motion luminance-based random-dot pattern. In most of the trials, the subjects viewed a stimulus that was intermediate between the zero-correlation state and the 100% correlation. Figure 8A (middle) shows a stimulus with 50% correlation in which 50% of the dots are always plotted at random locations within the aperture while the other 50% of the dots are plotted in correlated motion; they continue to move in the same direction during the stimulus display.

The aim of this task was to determine the threshold-of-motion intensity for which a subject could reliably discriminate the direction of motion, up–down, left, or right.

The stimulus consisted of a dynamic random-dot display presented in a square 10 × 10-deg aperture situated at 2 deg left or right of a white fixation mark. The subject is instructed to maintain fixation on the fixation mark for the entire testing period. The dynamic random-dot field consisted of 15 dots, each subtending 1.8′ × 1.8′, thus giving a dot density of 0.3 dots/deg^2.

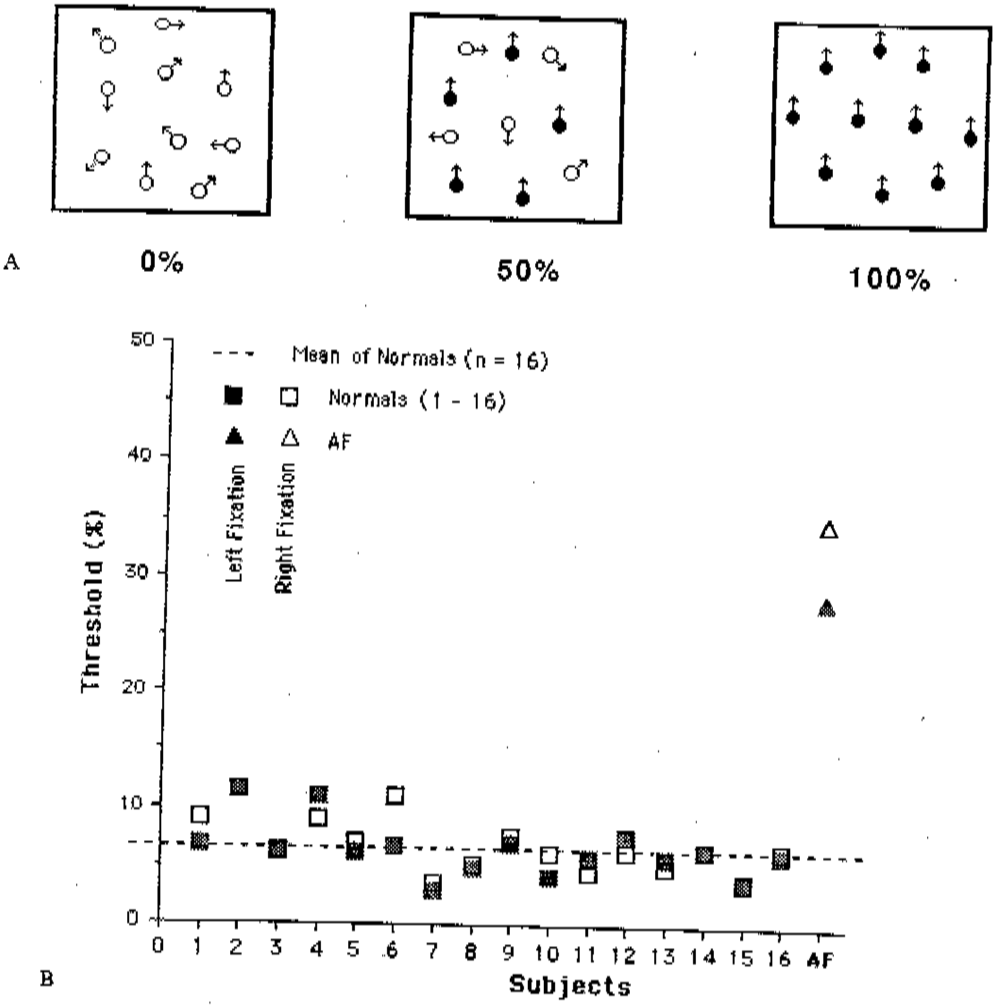


Fig. 8. Motion coherence. A: A schematic representation of the dynamic random dot stimuli employed. Each stimulus presented dots in rapid succession and each dot survived for a brief period of time before being replaced. In the 0% correlation case, on the left, the replacement dots were plotted at random locations within the aperture so that the dot field appeared as twinkling visual noise without motion in any single direction. In the 100% correlation case, on the right, each dot was replaced by a partner with a constant offset in space and time, so that the dot field appeared to move in a single coherent direction. Intermediate conditions, such as that for 50% correlation displayed in the middle panel, were generated as follows. There was a finite probability (here 0.5) that the dots would be replaced by the constant offset in space in the next frame. Otherwise it was plotted elsewhere in the display on a random basis and provided dynamic masking noise. The stimulus is presented 2 deg to the left or to the right of a fixation mark. A four alternative forced-choice procedure was used to determine the threshold correlation for which observers could successfully discriminate the direction of motion. B: The graph plots the threshold correlation for normal controls and A.F. was determined by a staircase procedure as the mean of the last nine reversals. The open symbols represent thresholds for the right fixation mark, and the black-filled symbols represent thresholds for the left fixation mark. The dotted line represented the mean of percent correlation required by the 16 normal controls. A.F.'s threshold for detecting the direction of motion was nine and 12 standard deviations higher than that of the normal controls, indicating that he was severely impaired on this task.

The algorithm by which the dots were generated was similar to that of Newsome and Paré (1988). Each dot in the aperture has an equal probability to be paired with a correlated dot in the subsequent frame and thus to contribute to the total motion signal. The correlated partners in turn had the same probability to be succeeded by another correlated partner. Thus, for example, if the probability was chosen to be 0.1, then the probability that a dot would continue on the same path for three consecutive frames was 0.001. Thus, it is unlikely that the perceiver could follow a single dot or a local cluster of dots. Rather, the impression of coherent movement in some dominant direction had to be derived from a global computation that integrated the local motion measurements. The size of the step of the dots was held constant at 9 min and thus, similarly to the previous experiments, the value for dx is in the range of the values reported to fall within the spatial limits that characterize the short-range motion mechanism at the same retinal eccentricity (Baker & Braddick, 1984).

A conventional "wrap-around" scheme was used, in which dots displaced beyond the aperture reappeared on the opposite side of the aperture. The speed of motion, defined as a function of the distance a dot is displaced between successive frames, was 3 deg/s.

Motion threshold was obtained in a four-alternative, forced-choice judgement (4-AFC) concerning whether the perceived direction of motion was predominantly up, down, left, or right. The stimuli were displayed by an interactive staircase procedure driven by the subject's responses and controlled by computer. The staircase procedure went from easy to hard, and after two consecutive correct answers the percent of correlated dots was reduced. The percentage of correlated dots was increased after one incorrect answer. The step size of the staircase was one-eighth of a log unit of percentage of motion coherence. The test terminated after 15 reversals (minima or maxima). The motion threshold was computed from the average of the last nine reversals.

A.F. was significantly impaired on the motion coherence task (Fig. 8b). His direction-discrimination threshold was 28.4% for left fixation and 35.2% for right fixation. The mean of the motion coherence threshold for the normal controls ($n = 16$) was 6.5% for the left fixation and 6.9% for the right fixation.[‡] Thus, in both the right and left visual-field stimulus presentation, A.F.'s threshold was significantly higher than that of the normal controls (9 and 12 respectively, standard deviations above the scores of the normal controls computed by a Z score; $P < 0.0001$).

Thus, it is plausible that A.F.'s deficit on the local speed discrimination and motion coherence tasks demonstrates a deficit in a higher motion system which integrates the output of the short-range motion processes, such as local velocity measurements or simply, a deficit of computing local velocity.[§]

Discussion of experiments 2a and 2b: A deficit of motion integration?

A.F.'s elevated thresholds on the motion coherence task are a strong indication of his inability to compute global motion fields. His poor performance on this task might be underlaid by an inability to extract motion signal and thus resulting in failure to resolve the ambiguity present in the motion flowfield. It is generally assumed that the ambiguity-solving process is fed by the direct outputs of simple and local motion measurements such as detection of local velocity. We have seen that A.F.'s performance was extremely impaired on the local speed discrimination task and that he was not able to use speed information to extract form from relative motion.

Thus, it is plausible that the system that integrates the measurements obtained by the local motion detectors of the short-range motion process was impaired in A.F. Moreover, the motion coherence task implies the use and integration of local motion measurements which is a processing stage in the analysis of visual motion that is beyond the plain detection of motion.

In summary, A.F.'s performance on the above five motion experiments showed severe deficits in the following areas: (1) inability to use speed differences to extract two-dimensional form from motion; (2) inability to extract magnitude of motion from local velocity measurements and compute global speed fields; and (3) inability to extract common global velocity from motion noise. All of the experiments discussed above used stimuli that fell in the range of the short-range motion process, thus suggesting that A.F. had a deficit in either sensing local motion or integrating the output of the local motion system, particularly in the presence of noise.

Higher-order motion

The experiment 3a and 3b examined the ability of an observer to recover three-dimensional (3-D) information from the relative motions of elements in the changing two-dimensional image. The experiment 3a is the generic structure from motion task (SFM) in which the stimulus consisted of the parallel projection of points covering the transparent surface of a rotating cylinder. The human visual system has a remarkable ability to recover the three-dimensional structure only from the two-dimensional projection of these dots.

Our previous studies of the SFM task (Vaina, 1988*b*; Vaina et al., 1988; Vaina, 1989) showed that patients with lesions to the right occipital-parietal area were unable to perceive three-dimensional structure-from-motion, while normal controls and patients with right occipital-temporal lesions did not have difficulties with this task. The generic display consisted of computer-generated dots painted at random location on the orthographic projection of a transparent rotating imaginary cylinder. The subjects were told that the display represented a simple, moving object and they were asked to describe in words what they perceived in the display. Typical answers of the subjects with right occipital-parietal lesions were that the display portrayed "birds high in the sky, snow blown by the wind, or ants crawling on the ground;" in no circumstance did these subjects have the impression of a rotating cylinder. Their responses contrasted with those of the subjects with lesions to the right occipital-temporal area who described perceiving "a cylinder, or rotating lamp shade, revolving door, a rotating garbage can, or a child's rattle." In all cases, their responses conveyed that this subject group perceived a rotating three-dimensional cylindrical object.

When presented with the 2-D pattern of moving dots, A.F. without any hesitation reported seeing a rotating cylinder. We were extremely surprised that A.F. had no difficulties with this stimulus, since on all of the other motion tasks his performance was consistent with the performance of the group with right occipital-parietal lesions. Furthermore, there was no indication that A.F. had any tendency to make guesses in any of the many perceptual demonstrations that were given to him. Yet, it could still be argued that in this particular case he "guessed" the correct answer without actually perceiving the cylinder as such. So to get a more detailed and quantitative assessment of A.F.'s ability to perceive 3-D structure from motion, we needed a test

---

[‡] The thresholds of motion coherence obtained by the human observers in the Newsome and Paré's task were roughly 2%. This may be accounted for by slight differences in the display characteristics. Dot density in their task was roughly 5 times higher than in our display (1.7 dots/deg^2 compared with 0.3 dot/deg^2). In addition, their subjects were experienced with psychophysical testing and in addition they also received extensive training on the task. Our data (unpublished observation) also suggest the dependency of the threshold of training and psychophysical experience of the observer.

[§] The motion coherence task reported here was presented separately in the right and left visual field, while in the other motion tasks reported in this study the subject freely inspected the display (the stimulus exposure time was 2 s). Thus, one may think that if free inspection were permitted A.F. might have scored better on the motion coherence task. We have experimental evidence against this possibility since, initially, A.F. was actually given a version of this task which was adapted from that of Williams et al. (1986). In this task, the dynamic random dot field was displayed in the center of the screen without any specific instruction referring to eye movements. The subject's task was to discriminate the direction of motion in a field of dynamic random dots whose fraction of coherently moving dots varied. A.F. could discriminate the direction of motion only when 70% of dots or more moved coherently, whereas normal subjects can perform the discrimination when only 6% of dots are moving coherently.

with graded levels of difficulty, either varying the percentage structure of random dots or varying the lifetime of the individual points in the display (see Siegel & Andersen, 1988).

Experiment 3a: Rotating cylinder

Two dynamic random-dot cinematograms were presented simultaneously, one containing the cylinder, the other a pattern of scrambled velocities.

The display (Fig. 9A) consisted of two dynamic random-dot cinematograms each subtending an area of 3×3 deg^2. Each cinematogram was composed of 128 dots with an average point density of 14 dots/deg^2. At the end of its lifetime, the point disappeared and was replotted at a new random location on the screen (within the boundary of each display) and it began a new trajectory. One of the two cinematograms portrayed a revolving hollow cylinder with a vertical axis. The angular velocity of the cylinder was 30 deg/s. The other cinematogram contained an unstructured stimulus generated by spatially shuffling the velocity vectors present in the structured display and thereby destroying the local spatial relationship between velocity vectors. Both the structured and the unstructured displays were generated by 50 frames each displayed for 33 ms, which were repeated for the total duration of the display. The maximum distance travelled by a dot between two consecutive frames was 4.3 arc min. Before describing the moving displays in further detail, it is important to note that one needs to rule out the possible contamination issue of texture gradients. This arises because the orthographic projection of dots "painted" uniformly on a cylinder will have a nonuniform density distribution with many more dots occurring at the edges of the perceived figure. Thus, subjects might correctly identify the cylinder on the basis of texture density alone. In order to remove the density cue, the random positions of dots were first generated in screen coordinates and then back-projected orthographically onto the surface of the cylinder. This procedure is repeated for every point when it gets replotted at the end of its lifetime; there is always approximately equal density at all locations on the cylinder at all times. Even if a small texture gradient remains, the gradient is shallow, particularly since it has been found that sensitivity to density modulation in random dots is rather poor (van Meeteren & Barlow, 1981). However, to make sure that this method removed this possible texture cue a temporally scrambled version of the display was constructed. Here the average density of the dots is identical to the experimental conditions but the coherent motion presumably required for the structure-from-motion is absent. Normal subjects were unable to distinguish the cylinder in the temporally scrambled display, but in the moving case they could correctly distinguish the correct from the random cylinder with greater than chance performance. As such we argue that the subjects' responses in the moving case are not contaminated by density cues.

The dynamic motion experiment consists of two versions. In version A, the point lifetime was held constant at 400 ms and the percentage of structure varied as follows: 100, 90, 80, 70, 65, 42, 27, and 11. The spatial positions (left *vs.* right) of the structured field and the unstructured field were randomly assigned. There were 160 trials equally distributed among the eight percentages of structure in the cylinder. In version B, the fraction structure was kept constant at 100% and the point lifetime of the dots in both the structured and unstructured display was varied as follows: 400, 200, and 100 ms. There were 20 trials for each point lifetime.

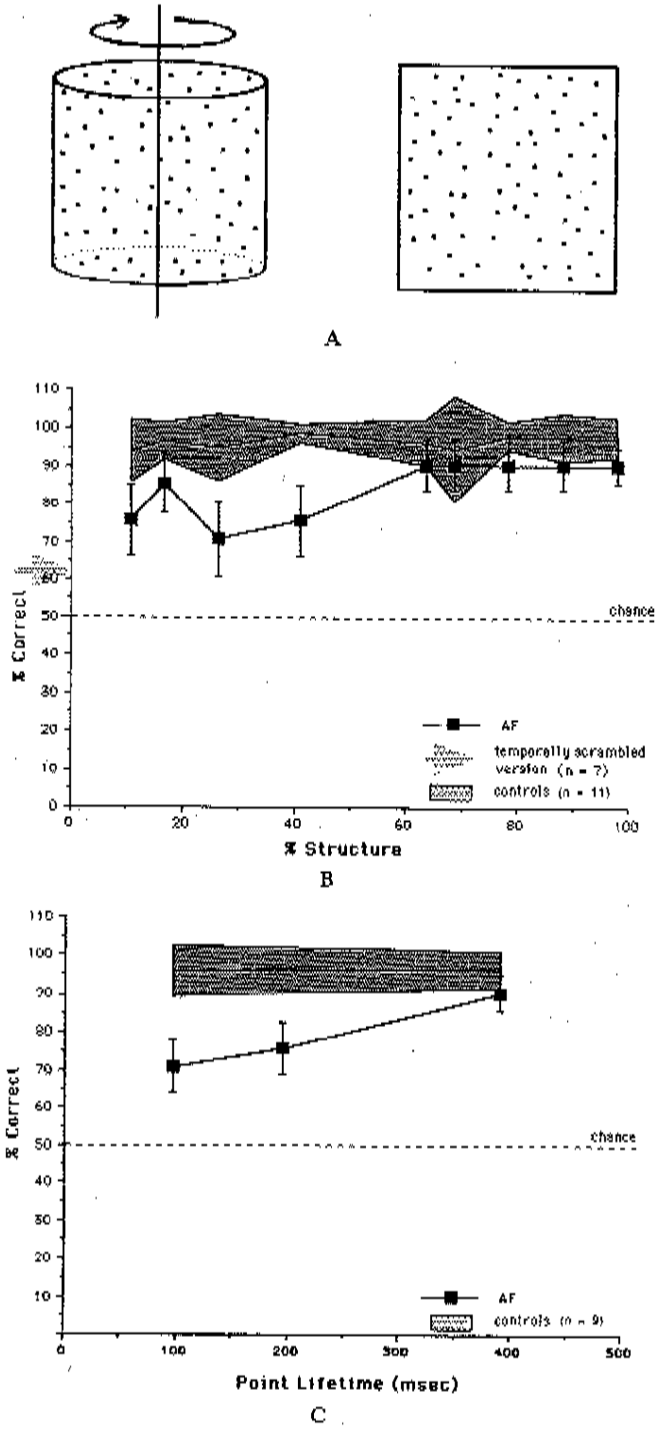


Fig. 9. Forced-choice identification of the rotating cylinder. A: A schematic representation of the two dynamic random dot fields employed. On the left is portrayed the structured stimulus generated by the orthographic projection of a square-shaped random dot velocity field onto a transparent cylinder which is rotated. On the right is portrayed a velocity field representing the unstructured stimulus. The dots survival time and the percentage of structure in the structured stimulus were independently varied in the display. B: Comparison of A.F. with normal controls where percentage of structure was varied in the cylinder (point lifetime was constant 400 ms). The shaded arrow to the left of the *y* axis represents the results from the normal subjects with the task modified such that the frames in the cylinder were temporally scrambled. C: The performance of the normal controls and A.F. on the experimental condition in which the structure of the cylinder was a constant 100% and the point lifetime varied.

In Fig. 9B, the data for A.F. (black square) is plotted along with the data obtained from 11 normal observers. In condition (A) when the percent structure was varied, A.F.'s performance was clearly above chance for all fractions of structure in the cylinder, and his results for the percent structure of 60% or more were virtually indistinguishable from those in the low normal range. Yet it should also be noted that when the percent structure of the cylinder was 42% or less, A.F.'s performance deteriorated. It did remain noticeably above chance, however, even for the lowest percent structure tested. The difference between his performance on the high amount of structure (65% and above) and low degree of structure (42% and below) was statistically significant ($\chi^2 = 7.41$, $P < 0.007$).

When the point lifetime was varied (version B), A.F.'s performance remained high (Fig. 9C) for the longest "lifetime," but deteriorated for shorter "lifetimes," a trend substantiated by Cochran-Mantel-Haenzel statistics ($F = 4.7$, $P < 0.031$). The general differences among his performances on this task at different point lifetimes was close to significant ($\chi^2 = 5.106$, $P < 0.08$).

Thus, despite the very poor performance on tasks designed to test the early motion mechanisms, A.F.'s performance on a 3-D structure-from-motion task is always above chance. In the experimental conditions in which there was no noise added (all of the dots in the stimulus were projected on the surface of the imaginary cylinder) and the integration time was relatively long (400 ms), A.F.'s performance was within the range of that obtained from the normal control group (90%). His ability to discriminate the cylinder deteriorated when a large amount of noise was introduced (more than 60% noise), which is consistent with his poor performance on the motion coherence task. His performance also deteriorated when the integration time available was less than 400 ms, while this did not affect the scores of the normal control group.

Experiment 3b: Recognition of biological action sequences

A very different yet compelling "structure-from-motion" demonstration was introduced by Johansson (1973). He called it "biological motion" and defined it as the pattern of movement generated from the evolving pattern of dots placed at the joints of a human actor. By placing patch light sources at the points of articulation of a moving human actor, Johansson demonstrated that the motion of the lights alone was sufficient to enable observers to perceive unequivocally the activity of the human actor. Such biological motion patterns have been shown to be sufficient for the perception of specific actions, the perception of gender (Kozlowski & Cutting, 1977), and the perception of sign language (Poizner et al., 1981). What is important about these types of experiments is that the only information available in the display is provided from the dynamic source alone. The pattern of the organization of the dots, and the pattern of their spatial modification over time, is the only clue for recognition.

In our experiment, the stimulus was presented on a videotape using scenes from the original Johansson movie. In the display, one sees only the pattern of lights attached to the joints of the human actor during the performance of some prototypical actions: *walking*, *stair-climbing*, *riding a bicycle*, *push-up*, *two men walking*, *shaking hands*, and *hugging*. When static, such a pattern gives no clue as to the "cause" of the lights, i.e. the identity of the gap of dots or the activities concerned. However, just from the motion of the pattern of lights, one gains the vivid impression of a human being performing these specific actions.

In the present experiment, the above prototypical actions were used. The examiner did not give any cues about the nature of the display; the subject was told only that he will see a bunch of dots moving in a certain direction and that his task was to first report the direction of motion and second, to describe what it looked like.

We were very surprised to find that A.F. presented no difficulties on this biological motion task. In all of the trials, he quickly recognized that the display portrayed a human actor carrying out a series of specific activities. The first three actions portrayed a man walking in different directions and A.F. identified the human walking and the correct direction of the movements. He also identified quickly and without hesitation all of the other actions—a man doing push-ups, climbing stairs, and riding a bicycle (although the wheels of the bicycle were not in view). The perception of bicycle ride emerged from the specific pattern of leg movements. In all of the trials presented, the motion overall is nonrigid, but there are, however, local rigidity constraints among the pattern of dots.

General discussion

In this paper, we have examined contrast sensitivity, stereopsis, color and form discrimination, visual spatial localization, and several stages of motion processing in a patient with bilateral lesions involving the posterior parietal-temporal areas as identified by MRI techniques. Neurological, neuro-ophthalmological, and neuropsychological studies indicated a mild form of the visual disorientation syndrome (Holmes, 1918), thus functionally confirming the posterior parietal involvement in the lesion. The anatomical extent of A.F.'s lesions, although large, are in the position to also involve the subcortical circuitry of the suggested human homologue of MT bilaterally. The conjecture that MT is involved in A.F.'s lesion is consistent with human positron emission tomography (PET) studies (Miezin et al., 1987; Allman, 1988; Zeki, 1990) where low-contrast moving stimuli were used to isolate "magnocellular" stream activity. The most significant responses to these stimuli were centered in the striate cortex and in the fundus of the parietotemporo-occipital fossa (PTOF) (Polyak, 1957, Fig. 271) which is located laterally in the occipital-parietal cortex at the junction of the Broadmann's areas 19 and 37. Figure 1G suggests that A.F.'s lesions undercut the neural circuitry corresponding to the area PTOF. As such, we predicted that the patient would have motion deficits. The basis for the prediction were recent results on behavioral tasks obtained by Newsome and Paré in monkeys with ibotenic acid lesions which selectively damaged the area MT. Newsome and Paré found that for the reliable discrimination of motion direction, postoperatively these MT lesioned monkeys required a significantly increased percentage of coherently moving dots in a dynamic random dot field. Preoperatively, the threshold of direction discrimination was equal to that of humans tested on the same psychophysical task. Correspondingly, in our study, in order to discriminate the direction of motion A.F. required a significantly higher percentage of correlated dots than the normal controls. In addition, A.F.'s speed discrimination was also greatly impaired.

Yet, A.F.'s performance on higher-order motion tasks was essentially normal when there was no noise added. He spontaneously identified the rotating cylinder in two experimental conditions: when the point lifetime was infinite and when the point lifetime was 400 ms. His performance on a forced choice selection of the "best cylinder" in a task of graded difficulty was also very good. When a large amount of noise was added to the image at the cost of loss of coherence, his performances degraded but still remained above chance. Furthermore, he correctly identified all of the human activities portrayed in a Johansson biological motion task and the direction of motion.

These results are puzzling, particularly given the commonly held assumption that higher-order structures-from-motion computations are derived from earlier, metrical measurements of local motion. For example, the theoretical models mostly rely on spatial-temporal derivatives of the velocity field (Longuet-Higgins & Prazdny, 1980; Lawton, 1983; Koenderinck & van Doorn, 1986). Such models use algorithms that perform a reverse geometrical projection of a sequence of 2-D images to a 3-D structure. Common to these approaches is the requirement that the system has available graded measurements that provide metrical information as an input. In addition, experimental findings suggest that damage to earlier extrastriate motion mechanisms leads to deficits on structure from motion tasks (Siegel & Anderson, 1986, 1988).

How can A.F., whose lesion extends bilaterally into the extrastriate areas corresponding to the presumed human homologue of MT and who has associated deficits of early motion mechanisms, perform normally on the two types of structure from motion tasks?

Three classes of explanations come to mind.

First, it is possible that structure from motion tasks do not require the precision of motion signals implied by many current models. Evidence for this view comes from several other sources. Ramachandran et al. (1987) has found that the exact velocities of a surface are not as critical as the outline of field of random dots. Nakayama (1985) has indicated that the nature of early motion hardware precludes the accurate registration of spatio-temporal derivatives of the velocity field. Loomis and Eby (1989) have shown that the recovery of depth is not compatible with 2-D and 3-D reverse geometric optic procedure.

Second, an alternative perspective is that our structure-from-motion tasks do not require any motion processing at all. Other forms of local measurements such as relative position (Ullman, 1979, 1984), shape deformation (Koenderinck & van Doorn, 1986), or texture gradients might provide sufficient information to obtain the structure.

A third possibility is that A.F.'s performance in the forced choice task is mediated by motion but not necessarily by rotating 3-D structure from motion. As one increases percent structure, for example, local motion coherence increases. Thus, this increased coherence alone may be sufficient to make the forced choice decision. A.F.'s reduced performance in comparison to normals might reflect this reduced ability to see coherent motion.

In terms of cues involved with the change of the pattern over time, one candidate is Ullman's incremental rigidity scheme that requires the comparison of positional information in successive frames or Koenderinck's shape deformation which relies on changing local orientation in different views. In addition, very approximate measurements of motion in combination with the cues listed above may be useful. Whatever the specific information utilized, very precise motion measurements do not seem to be involved.

A.F.'s excellent performance on the Johansson biological motion task is perhaps more startling. We cannot think of any obvious strictly non-motion cues that would have allowed A.F. to perform so well. Nor can we point to any lower-level motion explanations such as perceived coherence (described above) since A.F. had not received any prior exposure to the moving Johansson figures, and could not have given the correct answers except by actual "recognition" of the human actions. Either the required motion signals for the perception of action are extremely imprecise and/or it could imply the existence of a separate motion pathway, specialized for the perception of biological action. Supporting the latter view are anatomical and physiological findings. Perrett et al. (1985) have demonstrated that in awake-behaving rhesus monkeys, cells in the anterior portions of the superior temporal sulcus (STS) were selectively sensitive to biological motion. Recent anatomical studies (Colby et al., 1988) suggest that a separate pathway, independent of MT, could provide motion information to this site, raising the possibility of an "action recognition" system in the temporal lobe receiving a separate motion input.

Conclusion

This detailed study of the visual abilities of a patient, A.F., strongly indicate that the psychophysical study of motion deficits in humans correlated with the specific anatomy of the lesion delineated by MRI studies, and the description of the pattern of eye movements provides a promising avenue for understanding the neural substrate of the visual-motion pathway in humans. Also such studies are a rich source for thinking about the mechanisms necessary for visual-motion perception and, together with the psychophysics of the normal visual system, provide direct means for validating experimentally the theories of visual motion processing. The patient, A.F., presented in this study is a particularly interesting challenge for current theoretical models of motion processing which would predict that his deficits on computing global motion fields in an ambiguous field of dots, his failure to extract the magnitude of a velocity vector and use it as a basis for grouping, and his inability to use speed information to extract form must strongly interfere with his ability to perceive 3-D "structure from motion." A.F.'s excellent performance on the 3-D structure-from-motion test at least suggests that none of the many existing elegant computational solutions for this problem provide a complete and unique solution.

Acknowledgments

This research was supported by NEI Grants EY07861-R-01 and EY 05408-5 and by the Boston University Seed Grants for Biomedical Research #7712-5-ENG-909. We thank the staff of the Department of Neuropsychology and Robert Haile, M.D. from the Young Stroke Unit of the New England Rehabilitation Hospital for referring this patient to us and for making available to us the clinical results relevant to this study. We thank the Vistech Consultants Inc. for kindly providing the Vision Contrast Test System. We also thank Richard Andersen and his laboratory members for discussions about the 3-D structure from motion experiment. Informed consent was obtained from the patient and healthy volunteers used in this work in accordance with the Declaration of Helsinki.

References

Adelson, E.H. & Movshon, J.A. (1985). Phenomenal coherence of moving visual patterns. *Nature* **300**, 523–525.

Allman, J.M. (1988). The search for area MT in the human brain. In

Proceedings of the European Brain and Behavioral Society Workshop on Segregation of Form and Motion, Tubingen.

Andersen, R.A. & Siegel, R.M. (1987). Effects of ibotenic acid lesions in area MT on motion perception threshold in the macaque monkey. *Investigative Ophthalmology and Visual Sciences* (Suppl.) **28,** 197.

Baker, C.L. & Braddick, O.J. (1982). The basis of area and dot number effects in random dot motion perception. *Vision Research* **22,** 1253–1260.

Baker, C.L. & Braddick, O.J. (1984). Eccentricity-dependent scaling of the limits for short-range apparent motion perception. *Vision Research* 25, 803–812.

Braddick, O.J. (1974). A short-range process in apparent motion. *Vision Research* **14,** 519–527.

Colby, C.L., Gattass, R., Olson, C.R. & Gross, C.G. (1988). Topographic organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. *Journal of Comparative Neurology* **238,** 1257–1299.

Efron, R. (1968). What is perception? In *Boston Studies in the Philosophy of Science, Vol. 4,* ed. Cohen, R.S. & Wartofsky, M.W., pp.137–173. Dordrecht: D. Reidel.

Ginsburg, A.P. (1968). A new contrast-sensitivity vision test chart. *American Journal of Optometry and Physiological Optics* **6196,** 403–407.

Ginsburg, A.P. (1983). *Vision Contrast Test System.* Vistech Consultants Inc., Ohio.

Hess, R.F., Baker, C.L. & Zihl, J. (1989). The motion-blind patient: low-level spatial and temporal filters. *The Journal of Neuroscience* 9, 1628–1640.

Hildreth, E.C. (1983). *The Measurement of Visual Motion.* Cambridge, Massachusetts: MIT-Press.

Holmes, G. (1918). Disturbances of visual orientation. *British Journal of Ophthalmology* **2,** 449–468.

Holmes, G. & Horax, G. (1919). Disturbances of spatial orientation and visual attention, with loss of stereoscopic vision. *Archives of Neurology and Psychiatry* **1,** 385–407.

Howard, H.J. (1919). A test for judgement of distance. *American Journal of Ophthalmology* **2,** 656–675.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics* **14,** 201–211.

Julesz, B. (1971). *Foundation of Cyclopean Perception.* Chicago, Illinois: University of Chicago Press.

Koenderinck, J.J. & van Doorn, A.J. (1986). Depth and shape from differential perspective in the presence of bending deformations. *Journal of Optical Society of America A* **3,** 242–249.

Kozlowski, L.T. & Cutting, J.E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception and Psychophysics* **21,** 571–580.

Lawton, D.T. (1983). Processing translational motion sequences. *Computer Vision, Graphics, and Image Processings* **22,** 116–144.

Longuet-Higgins, H.C. & Prazdny, K. (1980). The interpretation of moving retinal images. *Proceedings of the Royal Society B* (London) **208,** 385–397.

Loomis, J.M. & Eby, D.M. (1989). Relative motion parallax and the perception of structure from motion. In *Proceedings of IEEE Workshop on Visual Motion,* Irvine, California, pp. 204–211.

MacQuarrie, T.W. (1953). *MacQuarrie's Test for Mechanical Ability.* Monterey, California: California Test Bureau.

Maunsell, J.H.R & Van Essen, D.C. (1983*a*). Functional properties of neurons in the middle temporal visual area (MT) of the macaque monkey, I: Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology* **49,** 1127–1147.

Maunsell, J.H.R. & Van Essen, D.C. (1983*b*). Functional properties of neurons in the middle temporal visual area (MT) of the macaque monkey, II: Binocular interaction and sensitivity to binocular disparity. *Journal of Neurophysiology* **49,** 1148–1167.

Miezin, F.M., Fox, P.T., Raichle, M.F., Allman, L.M. (1987). Localized responses to low-contrast moving random dot patterns in human visual cortex monitored with positron emission tomography. *Society of Neuroscience Abstracts* **13,** 634.

Nakayama, K. (1985). Biological motion processing: a review. *Vision Research* **256,** 625–660.

Nakayama, K. & Tyler, C.W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Research* **21,** 427–433.

Newsome, W.T., Wurtz, R.H., Dursteler, M.R. & Mikami, A. (1985). Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *Journal of Neuroscience* **5,** 825–840.

Newsome, W.T. & Paré, E.B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience* **8,** 2201–2211.

Newsome, W.T., Britten, K.H. & Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. *Nature* **341,** 52–54.

Perrett, D., Chitty, A., Mistlin, A. & Harries, M. (1985). Visual cells sensitive to biological motion. *Behavioral Brain Research* **16,** 153–170.

Poizner, H., Bellugi, U. & Lutes-Driscol, V. (1981). Perception of American Sign Language in dynamic point-light displays. *Journal of Experimental Psychology: Human Perception and Performance* **7,** 430–440.

Polyak, S. (1957). *The Vertebrate Visual System.* Chicago, Illinois: University of Chicago Press.

Ramachandran, V.S., Cobb, S. & Rogers-Ramachandran, D. (1987). Recovering 3-D structure from motion: some new constraints. *Society of Neuroscience Abstracts* **13,** 630.

Siegel, R.M. & Anderson, R.A. (1986). Motion perceptual deficits following ibotenic acid lesions of the middle temporal area (MT) in the behaving rhesus monkey. *Society of Neuroscience Abstracts* **2,** 1183.

Siegel, R.M. & Anderson, R.A. (1988). Perception of three-dimensional structure from motion in monkey and man. *Nature* **331,** 259–261.

Ullman, S. (1979). *The Interpretation of Visual Motion.* Cambridge, Massachusettes: MIT Press.

Ullman, S. (1984). Maximizing rigidity: the incremental recovery of 3-D structure from rigid and rubbery motion. *Perception* **13,** 255–274.

Ungerleider, L.G. & Desimone, R. (1986). Cortical connection of the area MT in the macaque. *Journal of Comparative Neurology* **247,** 190–222.

Vaina, L.M. (1987). Visual texture for recognition. In *Matters of Intelligence,* ed. Vaina, L.M., pp. 89–115. Dordrecht, Holland: D. Reidel Press.

Vaina, L.M. (1988*a*). Deficits of motion analysis in right occipito-parietal lesions in humans (Abstract). In *Proceedings of the European Brain and Behavioral Science Workshop on "Visual Processing of Form and Motion,"* Tubingen.

Vaina, L.M. (1988*b*). Effects of right parietal lobe lesions on visual motion analysis in humans. *Journal of the Optical Society of America* (Suppl.) **29,** 434.

Vaina, L.M., LeMay, M., Naili, S., Amarillio, P., Bienfang, D. & Montgomery, C. (1988). Deficits of visual motion analysis after posterior right hemisphere lesions. *Society of Neuroscience Abstracts* **14,** 458.

Vaina, L.M. (1989). Selective deficits of visual motion interpretation in patients with right occipito-parietal lesions. *Biological Cybernetics* **61,** 1–13.

Vaina, L.M., LeMay, M., Choi, A., Kemper, T. & Bienfang, D. (1989). Visual motion analysis with impaired speed perception: psychophysical and anatomical studies in humans. *Society of Neuroscience Abstracts* **15,** 1256.

van Meeteren, A. & Barlow, H.B. (1981). The statistical efficiency for detecting sinusoidal modulation at average dot density in random figures. *Vision Research* **21,** 765–778.

Warrington, E.K. & Taylor, A. (1973). The contribution of the right parietal lobe to object recognition. *Cortex* **9,** 152–164.

Warrington, E.K. & James, M. (1988). Visual apperceptive agnosia: a clinico-anatomical study of three cases. *Cortex* **24,** 13–32.

Williams, D., Phillips, G. & Sekuler, R. (1986). Hysteresis in the perception of motion direction as evidence for neural cooperativity. *Nature* **324,** 253–255.

Zeki, S.M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of rhesus monkey. *Journal of Physiology* **236,** 547–573.

Zeki, S.M. (1990). The form vision of achromatopsic patients. *The Brain,* LV Cold Spring Harbor Symposium on Quantitative Biology, May 30–June 6, p. 16.

Zihl, J., von Cramon, D. & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain* **106,** 313–340.