Eye movements and their role in visual and cognitive processes
E. Kowler, Editor

CHAPTER 3

The role of eye movement in the detection of contrast and spatial detail

Robert M. Steinman and John Z. Levinson

Department of Psychology, University of Maryland, College Park, MD 20742, U.S.A.

1. Introduction

1.1. Scope, goal and plan of this chapter

The role of eye movement in a variety of basic visual information processes, such as the detection of contrast, the detection of fine details and the discrimination of the positions of such details relative to each other, has a long history in visual science. Eye movements have been treated as beneficial, as detrimental, and as both beneficial and detrimental, depending on the nature of the visual stimulus. Eye movements have also been treated as irrelevant for understanding basic visual processing, on and off, during the last 125 years. This chapter traces out the reasons for such varying interest in the relationship between eye movement and basic visual information processing throughout this entire period. We adopted a long-term historical perspective because we feel that both, now highly specialized subjects, oculomotor control and visual information processing, as well as current understanding of the relationship between these specialized subjects, would benefit from a relatively detailed examination of prior attempts to work out the relationship between what the eyes do and what the human being can see. Reasons for ignoring the potential significance of eye movements in the past are as worthy of review as reasons for emphasizing their role. We base this claim on our belief that now that we have just entered a period in which eye movements are coming into prominence again, we are inclined to believe, based on the history of this area, that interest will last a decade or two to be followed by a period of neglect, once again, for seemingly plausible reasons. We hope that the reader will not interpret this last remark as simply the cynicism of older investigators. He will, we believe, come, as we have, to view the work in this area as exceptionally cyclical with peaks for and against an important role for eye movement, reflecting technical or theoretical developments in one or the other specialty. Other problem areas in visual science show analogous perturbations but we feel that they are more prominent here perhaps because those of us who specialize in vision or in oculomotor control tend to return to safer, more familiar ground once difficulties are encountered at the interface of these two relatively technical specialties. Those few who have tried to work on problems at this interface, bringing only expertise in one but not the other specialty area, have not, in our opinion, advanced far beyond their largely uninformed, initial assumptions about how eye movements interact with visual processing. It is for this reason that we undertook this chapter together: one of us (JZL) is a veteran visionary, the other (RMS) a veteran oculomotorist. We have tried, by working together on research at the interface of vision and oculomotor control, as well as by collaborating in writing this chapter, to provide some new information and also to provide the reader with a view of prior and present work that does not do violence to either of our specialties. If we have succeeded, we will have provided the reader with some insight into where this problem area has been, where it is, where it might profitably go and the problems likely to be encountered along the way.

Tracing out in some detail the history of what has become an interdisciplinary problem area over a period of more than a century required considerable selectivity in our choice of examples. We have neglected scores of interesting papers and failed to cite directly many very worthy investigations, investigators and their ideas. Our selection was dictated by knowing the end of the story we wanted to tell before we began and we chose our examples accordingly. The chapter reads as though it was historical, but it is history seen from a parochial point of view. Specifically, our 'history' sets the stage for our current work, in which it is proposed that eye movements play a very prominent role in basic visual information processing. This treatment is self-serving, but we tried to write the chapter in such a way as to encourage the reader to believe that we have treated antecedent developments fairly, albeit idiosyncratically. We tried to accomplish this in two ways: By referring the interested (or suspicious) reader to major contemporaneous reviews of the material under discussion and by adopting the practice of using quotations rather more extensively than is current practice. We adopted this practice not only because it allows the reader to judge for himself whether we are being fair but also because it avoids our being credited with the ideas of others – an outcome almost unavoidable in the currently popular writing style in which paraphrase is preferred to quotation.

We begin the chapter with a brief section describing so-called 'types' of eye movement and a brief treatment of the role each type might serve in visual information processing. Prevailing ideas about different types of eye movement and their implications for defining specific oculomotor 'subsystems' are derived from observations and classifications originally made by Raymond Dodge, a psychologist who began the objective, systematic study of human eye movement early in the present century (Dodge, 1901, 1903). Currently, it is popular to describe the oculomotor system as if it is composed of a number (at least 5) of independent 'subsystems', each of which has different velocity-amplitude characteristics, is controlled by different stimulus characteristics and serves a somewhat different function. This approach has influenced a great deal of oculomotor research since its introduction by Dodge. Currently, the value of this approach has begun to be questioned (Collewijn, 1989; Erkelens et al., 1989b) and it is likely to be replaced in future years by a simplification into only two subsystems, that is, a fast, jump-like saccadic subsystem and a smooth eye movement subsystem, whose action includes what is now described as the output of the smooth pursuit, slow control, optokinetic nystagmus, vestibulo-ocular response and vergence 'subsystems'. We have adopted current usage in our introductory sections and throughout the chapter because it makes it easier for the reader to evaluate the classical literature on our topic and also allows us to use quotations which would not be possible if the oculomotor material were treated in less conventional language. We expect that a good deal of the oculomotor terminology we have used will seem quaint in a few years. The ideas should, however, remain clear despite the fact that such distinctions as vergence, smooth pursuit, OKN, VOR and slow control may fall into disuse. The introductory sections, just below, also serve to define a number of terms and important concepts used throughout the chapter.

1.2. Eye movement characteristics and putative functions

The human eye rotates for a number of quite different reasons. Specifically, it may rotate in a rapid jump-like manner ('saccade') so as: (1) to bring the retinal image of an attended, but eccentric, stationary object to fall at the foveal center. Once the attended object is at the foveal center, the object is said to be 'fixated' or, less commonly, 'foveated'. Attention and fixation may remain coincident once fixation is established, or they may, once again, go their separate ways, establishing a complex set of asynchronous cognitive processes and visuomotor acts (see Kowler, Ch. 1 of this volume, and Viviani, Ch. 8 of this volume, for discussion of this complex interaction). The eye may also rotate smoothly so as: (2) to reduce the retinal image velocity ('slip') of an attended object which is moving in the visual field. When this occurs, the eye is said to be 'smoothly pursuing (tracking)' the attended object. Both saccades and smooth pursuits are usually observed in the oculomotor pattern when moving objects are tracked. This mixed style of tracking reflects a subject's desire to follow or to lead a moving object by combining 'catch up' or 'get ahead' saccades with smooth, retinal image slip-reducing, eye rotations – a strategy that can minimize tracking error if it is used cleverly. This mixed tracking strategy seems to be used by most, if not all, subjects, but it is not a hard-wired oculomotor characteristic. Simple instruction is sufficient to modify the manner in which saccades and smooth pursuits are mixed when a moving object is tracked (Puckett and Steinman, 1969).

The eye may also rotate so as: (3) to compensate for movements of the head and body, allowing the line of sight to remain near a fixated stationary object as the orientation of the fixator changes in space. These eye rotations are called 'compensatory' eye movements. They include (a) saccades which re-establish fixation of an attended stationary object displaced from the foveal center by head or body movement, (b) smooth pursuit of an attended stationary object when its image slips on the retina because of deficiencies inherent in the operation of (c) the vestibulo-ocular subsystem, which uses signals from the semicircular canals to produce smooth rotations of the eye opposite in direction to rotations of the head. These vestibularly activated smooth compensatory eye rotations serve to help maintain objectively stationary visual details relatively stationary on the retina as the head or body rotates. Such compensatory eye movements are frequently referred to as the 'VOR', an acronym derived from vestibulo-ocular reflex (or response).

Finally, (4) the eye rotates because it is not fused to the skull. This arrangement guarantees that the eye will rotate with respect to the skull despite the operation of the position- and velocity-sensitive biological 'control systems' that have evolved to compensate for movement of the head and body or for movement of objects in the visual field (Walls, 1962). Real biological control systems are certain to fall somewhat shy of perfection when they operate. These imperfections always produce some irreducible variability of eye position and velocity (oculomotor 'noise'). There is also, what we now consider to be a noisy, high-frequency small-amplitude eye movement that can be seen while a subject attempts to maintain fixation with his head immobilized on a biting-board. These tiny eye movements are called 'physiological nystagmus' or 'high-frequency tremor'. These tremors, which can only be observed with the most sensitive recording instruments, have been ascribed to incomplete tetany of the extrinsic oculomotor muscles during maintained fixation (e.g., Cornsweet, 1956). Discussion of attempts to establish the functional significance of the various kinds of eye movement, described above, constitutes a large part of the content of this chapter.

1.3. Eye movements and the fovea

Eye rotations, regardless of their origin and purpose, have considerable significance for basic visual processing. A good deal of this significance arises from the fact that the human retina is exceedingly heterogeneous in its functional properties. A foveate animal, such as ourselves, must be able to orient its eye with respect to the direction of objects in space and to maintain the eye's orientation exceedingly well because our best detail vision depends on specialized tissues (very densely packed cone receptors) found only at the center of the floor of the fovea – a region occupying less than 0.02% of the retinal surface area (Polyak, 1941). The presence of such a tiny highly specialized fovea guarantees that a very high degree of oculomotor skill will evolve so that a foveate animal, for example, such disparate creatures as a chameleon or a human being, is able to make use of its specialized foveal cells to discriminate fine details in the visual environment (see Walls, 1962, and Steinman, 1975, for elaboration of this point).

1.4. Eye movements as sources of neural transients

There are additional reasons for taking eye rotations seriously when considering basic visual processing. Eye rotations might be beneficial for the detection of contrast or spatial detail by producing changing or transient retinal stimulation. Changing or transient stimulation might help, or might even be required, to prevent photochemical depletion, neural fatigue or adaptation. Transient stimulation might actually be required to generate such basic physiological events as afferent bioelectrical potentials or lateral inhibitory processes in the retina or higher visual centers. Alternatively, eye rotations might be detrimental to basic visual processing because they could create, or allow, excessive retinal image motion. Such retinal image motion could smear the 'proximal stimulus' (the distribution of light on the retina where the transduction of physical to biophysical messages takes place). However, commands to generate eye rotations might initiate efferent or associative neural processes which serve to elevate detection thresholds, and thereby reduce the visual significance of the smear or retinal blur caused by high-velocity eye movements (see Matin, 1974; Volkmann, 1986; or Sperling, Ch. 7 in this volume, for reviews of research on 'saccadic suppression' – the name given to threshold elevations observed before, during and following a saccade).

All of these potentially beneficial and detrimental effects have been described in the visual psychophysical literature. All call attention to the potential importance of eye movement for visual processing. All have been studied and discussed, with varying enthusiasm, for more than a century. Important trends in the development of ideas about the relationship between eye movement and basic visual processing will be traced out next. Current thinking about this relationship, as in other areas of visual science, sometimes fails to take antecedent mistakes and progress into account. Our review includes this old material, partially for completeness but also because it will be shown later that many very old ideas bear directly on recent and current work.

Ideas about the relationship of eye movement to basic visual processing had well-established roots by the time visual science entered the final quarter of the 19th century. Like most areas in visual science in this period, very clear alternative positions were available. Alternatives, proposed clearly and authoritatively by H. von Helmholtz (1866), were traditionally contested and/or elaborated by E. Hering (1899, 1920). Their antagonistic relationship can be found in the subject matter of this chapter as well as in many other problem areas. Consider first Helmholtz's treatment of 'visual acuity'.

2. Helmholtz's treatment of visual acuity

2.1. The retinal mosaic and intensity discrimination

As early as 1866 Helmholtz pointed out that the finite size of the light-absorbing retinal elements, which were already believed to be the rods and the cones on the basis of psychophysical experiments (Brindley, 1960, p. 151), could place a theoretical limitation on the ability to discriminate details in an extremely fine pattern. Imagine the fovea to be stimulated by two mathematical point-images. In this *hypothetical* case it is apparent that the ability to discriminate two discrete points requires their separation to be at least as large as the diameter of one of the visually sensitive retinal elements. If the separation were smaller, only a single element would be stimulated and discrimination would be impossible. There would be no physiological basis for the perception of the gap between two mathematical point-images if the gap were smaller than a retinal element. In Helmholtz's words: "The light which falls upon a single sensitive element can produce only a single light sensation, within which it is impossible to distinguish whether individual parts of the element are strongly illluminated, others weakly illuminated" (p. 215). In other words if we could test with infinitely small points or their extension as infinitely thin lines, retinal element size would place a limit on acuity. This limit, as Helmholtz knew well, applied only to hypothetically thin stimuli and that once acuity is measured with real points or real lines, the ability to discriminate detail could be appreciably better than this 'theoretical' limit imposed by the retinal mosaic. Real stimuli always have appreciable extent in the proximal stimulus, guaranteeing a gradient of light stimulation which will fall on a number of elements, particularly near the central fovea where these elements are densely packed. According to Helmholtz, when real stimuli are used to measure visual acuity, acuity is limited by the capacity to discriminate intensity, not by the size of the retinal elements.

2.2. Elaboration of Helmholtz's ideas

Helmholtz's emphasis on intensity discrimination in limiting visual acuity was picked up and elaborated by Hartridge (1922) and Hecht (1927) more than 60 years later. Intensity discrimination theories of visual acuity subsequently came to be known as 'static' theories (Falk, 1956). They were called static because they ignored the potential importance of eye movements. The proximal stimulus was treated as though it were stationary. This treatment implies one or more of the following assumptions: (1) visual information is sampled only within relatively brief intervals during which the test stimulus cannot move far or fast, (2) oculomotor compensation for movement of the stimulus or the body of the observer is virtually perfect, producing effectively stationary test stimuli in the presence of motion, or (3) oculomotor noise is too small to have visual consequences. Much more will be said about each of these possibilities later. Static theories can be contrasted with 'dynamic' theories of visual acuity which are based on Hering's ideas (1899). Dynamic theories emphasize the role of eye movement. Ultimately, they develop to the point where eye movement is treated as a *necessary* condition for visual acuity. Hering's ideas and their development will be described later after elaborations of Helmholtz's approach have been presented.

Many authors since 1866, during the early years of the current century as well as more recently, have been under the mistaken impression that Helmholtz proposed an anatomical limit for real, as well as for hypothetical, acuity targets. For example, Hartridge (1922) implied that he may have made this mistake by saying that "One assumption has been made in the past, namely that the finite diameter of the foveal cones sets a maximal limit to the resolving power of the eye, and various calculations of the performance of the eye under different circumstances have been made on this basis" (p. 52). Hartridge does not attribute this idea specifically to Helmholtz, nor to anyone else for that matter, but it seems likely that its basis is Helmholtz's treatment of hypothetical test stimuli. Le Grand (1967, p. 106) made a clear attribution of this idea to Helmholtz more recently. Hartridge went on to elaborate Helmholtz's treatment of real acuity stimuli by calculating the illumination differences on adjacent foveal cones produced by particular test stimuli and relating the calculated intensity differences to the likelihood that such differences would be sufficient to be perceptually discriminable. His calculations of retinal light distributions include aberrations and other errors expected in a biological optical system such as the living human eye. In short, Hartridge's often-cited paper is an elaboration of Helmholtz's intensity discrimination approach to the limit of visual acuity. Byram (1944) subsequently pointed out that Hartridge's calculations could be in error by as much as a factor of two because Hartridge used Rayleigh's equations intended for a square aperture on image formation in the human eye, whose aperture is round. Hecht (1928) will make the same mistake with more modest consequences (15%) a few years later. A clear understanding of Helmholtz's (1866) priority for his treatment of the acuity limit-intensity discrimination question has not, however, been completely lost from the secondary vision literature. It reappears from time to time where the persistent misinterpretation of Helmholtz is duly noted and criticized (e.g. Wilcox and Purdy, 1933, or Walls, 1943). The retinal mosaic-acuity limit idea is often, but not always, falsely attributed to Helmholtz. It kept reappearing periodically for more than a century probably because it could be made to serve a useful didactic purpose (e.g. Senders, 1948; Riggs, 1965; or Le Grand, 1967). More recently, textbook treatments of spatial vision tend to ignore the retinal mosaic-acuity limit idea when they discuss the problem in terms of contrast sensitivity and the 'spatial modulation transfer function' – an approach brought from optical engineering by Schade (1956). The retinal mosaic and the functional significance of the fineness and symmetry of its cytoarchitecture have been the subject of renewed interest very recently (see Yellott, 1983; Hirsch and Hylton, 1984; Williams, 1985; for representative research).

3. Hering's approach to spatial vision

3.1. The direction, as well as the number, of perceived details

Consider next Hering's approach to the problem of visual acuity. His approach came to lean heavily on the role of eye movement in the hands of subsequent theorists. Consider first the traditional definition of visual acuity commonly found in the literature of Hering's day. This definition, 'visual acuity is the capacity to perceive fine details' (e.g. Hofmann, 1920), was usually 'operationalized' (studied empirically) by measuring the smallest separation, in visual angle, at which two neighboring objects could be resolved in the sense that two objects, rather than one object, were perceived. Secondary source treatments of visual acuity, starting with Helmholtz's first edition (1866) of his *Treatise on Physiological Optics*, traditionally include accepted values of visual acuity which had been measured up until that time with a variety of different test stimuli. Fixed stars, parallel lines of the same width, parallel lines with wider and narrower intervals, white squares separated by a black grating, spider webs and rod gratings were some of those mentioned in Helmholtz's original compilation. Tables of acuity values with such different test stimuli permeate the vision literature on acuity. Hartridge (1922), for example, continued this tradition and added new test results for such stimuli as absorption bands, and black and bright brass wires to the list of tasks. By the time Senders (1948) published her review article on visual acuity, stereoscopic acuity and vernier acuity had become traditional members of the list.

These were important additions. They highlight a distinction made by Hering almost 50 years earlier. Hering (1899) distinguished two kinds of detail vision. In this he was influenced by a report of Wülfing (1892), who measured what we now call the vernier displacement threshold of the offset contained within a dark bar ('vernier acuity') and found it to be only a few seconds of arc – very much better than anyone's ability to perceive two objects when they are separated by a gap of only a few seconds of arc. Hering was already familiar with Volkmann's (1863) demonstration that differences in the width of bars as small as 7 seconds of visual angle could be discriminated when the two bars were seen at the same time. Such delicacy of detail vision, about an order of magnitude better than the ability to discriminate gaps, led Hering to propose distinguishing resolving power (*Auflosungsvermogen*), the capacity measured in the traditional visual acuity task, from the space sense (*Raumsinn*), the capacity to distinguish differences in spatial position as measured in vernier, stereo and size estimation tasks. Interest in Hering's *Raumsinn* was rekindled recently by Westheimer (1981), who has been studying what he calls 'hyperacuity', i.e., the capacity to perceive position differences much smaller than the smallest retinal element. Hering's *Raumsinn* is a quite remarkable capacity, not only because of its delicacy, but also because it has been known since the turn of the century to be insensitive to blurring, intensity and appreciable motion. Hering (1920) also anticipated another currently important characteristic of visual acuity when he emphasized the importance of contrast in the perception of detail. Hering, as well as Helmholtz, was well aware that imperfections in the dioptrics of the normal human eye degraded the quality of the proximal stimulus and proposed that such degradations could be overcome by contrast effects produced by antagonistic neural processes – a quite modern view. He said: "The retinal image of every contour line, even when seen with perfect accommodation, is blurred. But our 'inner eye' has the power to create in the psychological visual field (providing that the intensity of the imperfect line-image is sufficient) a sharp boundary between two contiguous colors – thereby producing sharply-contoured elements in the retinal image. Our visual system owes this power to interaction of visual areas. The retinal image is always blurred; but, like the photographer who retouches a defective print, this interaction corrects the picture of external objects" (p. 154).

3.2. Weymouth's development of Hering's ideas

Hering's influence on dynamic theories of visual acuity can be seen first, and most explicitly, in the work of Averill and Weymouth (1925). These authors began by pointing out that thresholds [vernier and stereo] "far below the known size of the retinal elements" have been measured but "no satisfactory explanation has been offered of the way in which the delicacy of a perception might exceed the retinal grain" (p. 147).* They go on to propose that the idea of the position of a straight line is a complex percept based on at least three factors; namely, mutual effects of adjacent retinal elements, the averaging of successive stimulus patterns on the retinal mosaic caused by the continual occurrence of small eye movements, and the combining of similar stimulus patterns from each of the eyes. They propose that the visual system employs a statistical method, which uses signals from a large number of retinal elements, to calculate a percept based on what they call the 'retinal mean local sign'.

3.3. An experimental test of Hering's ideas

Averill and Weymouth (1925) acknowledged that their proposals were anticipated by Hering (1899) – a fact that Weymouth, Andersen and Averill (1923) had overlooked in an earlier brief report. They go on to point out that Hering, despite his clear, but neglected priority, gave no empirical support for his ideas – a situation they set out to remedy. Averill and Weymouth (1925) describe an ingenious simulation experiment in which they attempted to reproduce the stimulating conditions they believed would occur during a vernier task in a living eye. First, they made a replica of a portion of the fovea centralis (magnified 350x) by drilling many fine holes in a thin sheet of aluminum. The placement and density of these holes were carefully contrived to imitate the receptor surface as it was believed to be at the time. Light from an optically distant source passed through this perforated diaphragm and fell on a frosted glass screen which was viewed by a 'reagent' (observer). The projected spots of light represent the visually sensitive elements of the fovea. Objects placed between the perforated diaphragm and the screen will cast a shadow on the screen, cutting off some of the 'retinal elements' from the light. This situation is intended to simulate a proximal stimulus, where some retinal elements are stimulated while others are not. Opaque, inverted V-shaped objects were used to cast the object-shadows. These objects were attached to a motor-driven eccentric cam which was used to simulate the effect of eye movements by oscillating the shadow used to simulate the proximal stimulus. The inverted V-shaped objects were of two types. One could have an offset on one of its edges at a variable position along a contour. The other type had a uniform contour. The reagent's task was to report the presence and position of the offset, if any, and to give a confidence rating about his report. He

---

* In this, Averill and Weymouth, like many others, seem to be unaware of Helmholtz's original treatment of the relationship of visual acuity to the capacity to discriminate intensity differences rather than to the size of retinal elements, see section 2.1.

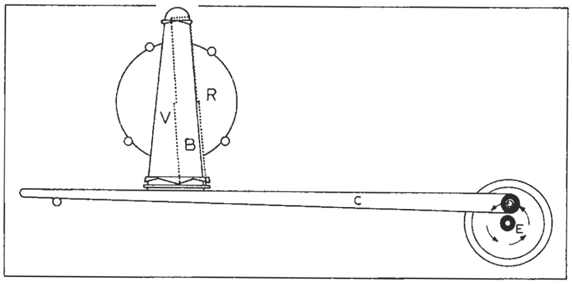

Fig. 1. Diagram of apparatus. R, replica of fovea (Fritsch) on an aluminum disc, cones being represented by minute perforations (see Fig. 2 for details). V, inverted V-shaped shield used to produce the image-shadow. B, brass rod with offset (dotted) held in such a position that its broken edge projects just beyond the margin of the shield. C, wooden cross-bar to which the shield is attached. The cross-bar and shield move in an elliptical path whose horizontal diameter is 8 mm (5 cone diameters) and vertical diameter about one-third as great. E, motor-driven eccentric which produces oscillation of the cross-bar and shield in an elliptical path. (From Averill and Weymouth, 1925)

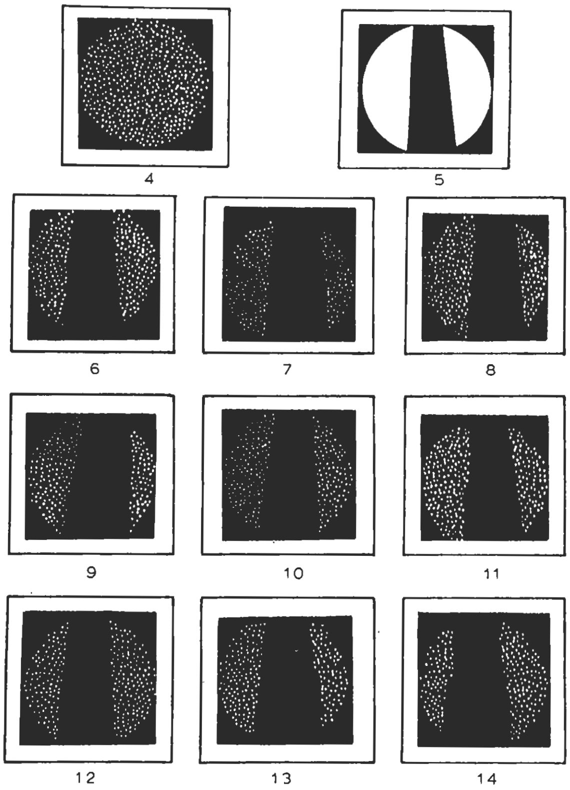

Fig. 2. (4) Diagram of retinal mosaic after Fritsch. Note the irregular arrangement of the cones and the great variation in inter-conal distances. This diagram is a replica of the perforated aluminum disc. (5) Appearance of the image shadow with displacement along the left margin and with the wider portion above the offset. (6–14) Representations of the retinal field as observed by the subject (reagent), who was required to judge the presence of an offset and its location when an offset was present. For example, in (14) there was a relatively large offset on the left which was wider in the lower part of the retinal field. (From Averill and Weymouth, 1925)

did this when the shadow of the V-shaped object was seen while it was stationary and also while it was moving so as to simulate fixational eye movements as they understood them to be at the time (see Figs. 1 and 2 for illustration of their technique).

They found: displacement thresholds far smaller than the ‘retinal elements’ (the projected spots of light); thresholds were lower when ‘eye movements’ (oscillations of the V-shaped object) were present than when the V-shaped object was stationary; thresholds were lower when longer objects were viewed (this allowed more retinal elements to contribute to the average); and short exposures with both eyes were better than short exposures with a single eye, demonstrating the role of binocular summation in ‘the delicacy of perception of offsets’. Averill and Weymouth (1925) concluded by claiming to have shown that relative motion between the retina and an image greatly increases the delicacy of visual perception.

The ‘dynamic’ theory of visual acuity was launched empirically by this experiment, which made the approach seem plausible despite somewhat indirect tests. Further progress with the dynamic approach had to wait more than 20 years to gain prominence after Selig Hecht’s static theory had fallen out of favor and a physiological variant of dynamic theory was proposed by Talbot and Marshall (1941) and Marshall and Talbot (1942). Hecht’s commanding position in the development of visual science and his static theory of visual acuity will be presented next.

4. Hecht's static photochemical theory

4.1. The man

Selig Hecht (1892–1947) was "one of the most vivid scientific figures of his time; a pioneer in the development of general physiology in this country [USA]; and for more than two decades leader in his chosen field, the physiology of vision" (Wald, 1948). Hecht's approach to visual problems was physicochemical. The origin of this approach is traced by Wald (1948) in his obituary of Hecht, where Hecht is described as having been profoundly influenced by three factors, namely, the birth of the science of photochemistry, which occurred near the turn of the century, Jacques Loeb's treatment of animal phototropism, with its source in ordinary physicochemical processes, and the publication of Arrhenius's book, 'Quantitative Laws in Biological Chemistry', in 1915 while Hecht was a graduate student at Harvard. Hecht favored biochemical, over neurophysiological, explanations of visual processes whenever possible, applying peripheral, biochemical explanations to potentially complex central visual processes (e.g. color), as well as to potentially simple, retinal visual processes (e.g., dark adaptation). Present-day visual biophysicists, undoubtedly wiser for having had such an influential, but only fleetingly successful, intellectual antecedent, tend to be less ambitious. Hecht's theory of visual processing was very broad in scope as well as parsimonious with respect to the number of required explanatory principles. For example, Hecht explained absolute sensitivity to light, the discrimination of brightnesses and the detection and discrimination of spatial details by reference to the concentration of photopigments in the rod and cone receptors and to the distribution of the thresholds of these receptors. Only the first of these problems – the absolute threshold for light (Hecht et al., 1942) – is still held to be amenable to photochemical explanations, but, even here, at the very beginning of visual processing, neural reorganization associated with dark adaptation has added a significant neurophysiological embellishment (Barlow et al., 1957). Hecht's approach was taken very seriously in his day – the simplicity and wide-reaching scope of the approach undoubtedly adding to its influence. Hecht employed classical Fechnerian psychophysical techniques with experienced human observers (his collaborators and himself) in most of his work. Ascidians, clams, bees and flies also served as his subjects.

4.2. Hecht's theory

Hecht's (1927, 1928) theory of visual acuity is an elaboration of Hartridge's (1922) approach, which, in turn, is derived, without attribution, from Helmholtz's idea that visual acuity is limited by the ability to discriminate differences in light intensity. Hartridge advanced this idea by calculating the light distribution in the retinal image of an optical device, such as the human eye, when it is presented with a variety of conventional acuity targets and comparing such light distributions to acuity thresholds measured with these targets. Hartridge concluded that the limit of acuity was determined by the ability to discriminate intensity. His calculated limit was about 5–10%. Hecht, adopting a similar approach, concluded that the limit of intensity discrimination and, therefore, the limit of visual acuity was about an order of magnitude better.*

Hecht's elaboration of the Helmholtz-Hartridge idea went far beyond demonstrating that the exquisite delicacy of spatial vision could be limited by a similar sensitivity to intensity differences. He developed a general theory, encompassing the full range of variation of acuity with a variety of differences in stimulating conditions, related these variations to the functional density of the retina complete with its two kinds of receptors, and explained all of these phenomena at the level of the

---

* See Hecht and Mintz (1939) for the 'minimum visible' acuity limit, that is, the ability to make out the presence of a dark bar against a moderately intense background, where the threshold was found to be about one half second of arc – a value that was calculated to correspond to an intensity difference across adjacent receptors of about one half of one percent with this type of display.

chemistry of the photopigments. Hecht's grand scheme could only seem to succeed by being selective with respect to the data it was willing to consider and by being indifferent to or, perhaps, naive about the desirability of statistically testing the goodness of fit of data to theory. A discussion of Hecht's theory of acuity and traditional criticism of the theory will be presented next (see Senders, 1948, for a more complete and particularly fine review of Hecht's theory published near the end of its useful scientific life).

Hecht's theory begins by recognizing that the human eye is not a perfect optical device and the presence of chromatic aberration, spherical aberration and diffraction guarantee that the distribution of a point of light will be spread when it is imaged on the retinal surface. The retinal image of an acuity target is considerably more diffuse than the calculated geometrical image of the same acuity target. He goes on to assert that the ability to discriminate details reduces to the ability to discriminate differences in intensity in this diffuse retinal image. Next, Hecht calls attention to the established fact that intensity discrimination improves with the intensity of the stimulating light and proposes that if acuity is actually an intensity discrimination, acuity should improve in the same manner. In other words, the functions for the intensity difference threshold ($\Delta I/I$)/intensity function and the visual acuity (1/minimum angle resolved)/intensity functions should be similar in form. Hecht supported this point in his 1928 and later papers by showing that the log of visual acuity plotted against the log of stimulating intensity is similar in form, that is, sigmoidal, to plots of $\Delta I/I$ against the log of the stimulating intensity. Hecht pointed out that Helmholtz had previously suggested that the poor visual acuity measured at low intensities of background stimulation (the shallow lower limb of this sigmoidal function) might be related to the coarseness of intensity discrimination under these conditions. Hecht goes on to consider why this might occur. He credits König with realizing that this lower limb might reflect the operation of rods in the periphery at low levels of illumination, but not before pointing out that Broca supposed that the connection between retinal elements and nerve fibers might not be fixed and that the number of retinal elements which communicate with each nerve fiber might vary with the intensity of illumination – an idea rather ahead of its time. Hecht claims credit in his 1928 paper for realizing that the lower limb in the acuity-intensity function is the rod limb while the rest of the function describes activity of the cones. He encourages the reader to confirm this simply by looking at an acuity target in low illumination, where the reader will note that fixation is eccentric, causing the target to fall on the peripheral retina where the highly light-sensitive rod receptors are located. Fig. 3 (top) reproduces the acuity-intensity functions Hecht used to develop his theory.

Hecht develops his ideas by pointing out that: "the fineness of detail which a surface can register depends on the number of receiving elements present in a unit area of the surface. In other words, its resolving power varies in inverse proportion to the average distance between the centers of the sensitive elements. This is very evident in such a case as the photographic plate. The retina is a surface of this kind since it is composed of discrete rods and cones which function as individual units or groups of units. The way in which visual acuity varies with illumination indicates the way in which the resolving power of the retina varies. A low visual acuity means that the average distance between the retinal elements is large; whereas a high visual acuity means that the distance is relatively small. To account for the large variation of visual acuity with illumination, one must suppose that the number of sensitive units per unit area of retina can and does vary nearly a hundred-fold. But the number of rods and cones in the retina is fixed anatomically. Therefore it is necessary to assume that the number of these elements is variable functionally" (pp. 259–260). Hecht next assumes that the thresholds of these sensitive retinal elements are not the same; rather, they vary in the "manner of populations" (p. 260).

Hecht's last assumption, that is, that receptor thresholds are normally distributed is illustrated in

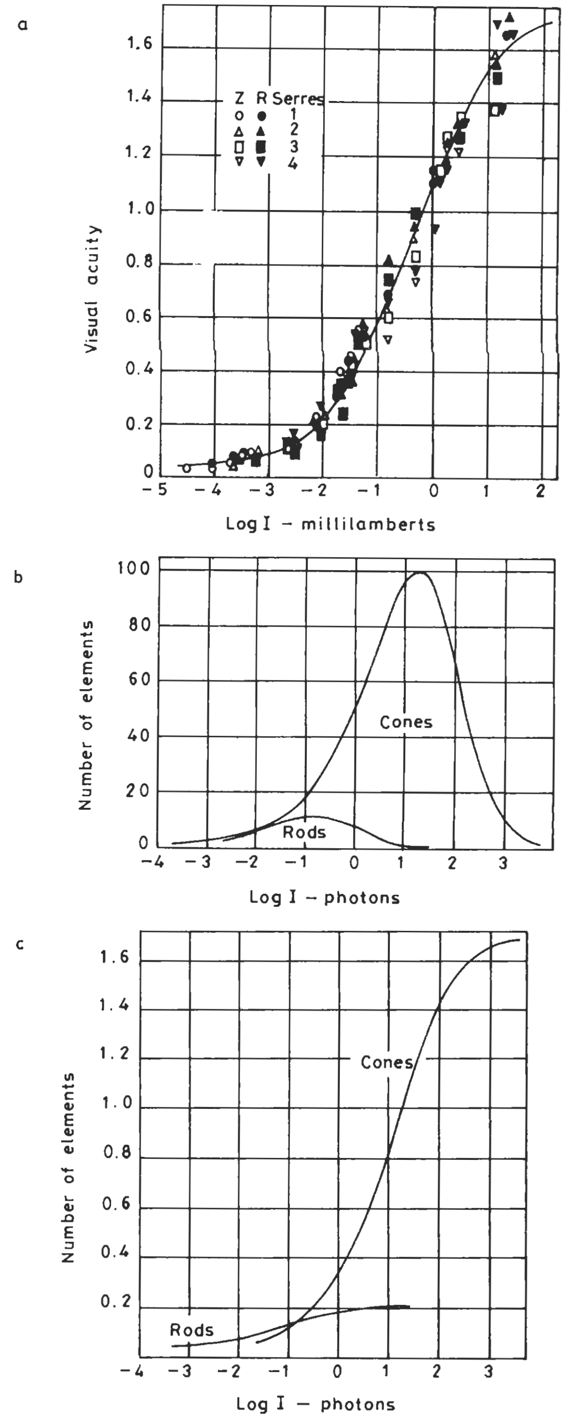

Fig. 3. (a) Relationship between visual acuity and illumination. Data of Roelefs and Zeeman re-plotted by Hecht. (b) Distribution of thresholds of rods and cones. (c) Statistical distribution of the sensitivity of rods and cones. (From Hecht, 1928)

Fig. 3 (middle), which reproduces his assumed threshold distributions. The 'functional density' of the retina can be estimated by summing the rod and cone threshold distributions for each potential illuminating intensity plotted in Fig. 3. Functional density (the integrated curve) is illustrated in Fig. 3 (bottom). **N.B.** Hecht emphasized that he is assuming that a level of stimulating intensity will be reached where the functional density asymptotes (saturates), that is, the number of cones stimulated to threshold no longer increases as intensity increases. This assumption clearly implies that visual acuity will cease to improve at these high levels of stimulation. This counter-factual prediction will, along with other problems, eventually make trouble for Hecht's approach. Hecht next goes on to test his theory quantitatively by using Reeve's data on the effect of stimulating intensity on pupillary diameter to correct König's acuity-intensity data. This correction is necessary once one wishes to estimate the actual retinal intensity associated with the König's measurements of visual acuity because König did not control for pupillary diameter. Hecht makes similar corrections for Roelofs and Zeeman's data and also examines the acuity data of two completely color-blind subjects (described by König and Uhthoff) whose eyes should only show the rod limb of the acuity function. In all cases there is excellent agreement between the calculated and observed relationships between illumination and visual acuity (see Hecht's, 1928, Tables I and II).

4.2.1. The unit retinal area

An appreciation of this relationship and Hecht's treatment of it requires brief discussion of his treatment of what he calls the 'unit retinal area'. His unit retinal area is "a minimal retinal area which contains the equipment for recording the various properties such as intensity perception, color vision, visual acuity, and the like, usually ascribed to the retina as a whole" (p. 275). Hecht bases his estimate of the unit retinal area on König's brightness discrimination data, in which König found 572 discriminable steps over the complete range of intensities to which the eye can respond: in other words, 572 just noticeable differences (JNDs), assuming, of course, that one JND (a mental unit) is equal to the difference threshold (DL), measured in physical units, at each level of stimulating intensity.[*] Hecht, on the basis of his work on intensity discrimination, asserted that 30 of these 572 steps are based on rod function and the remaining 542 are based on cone function. This number, 542, then, represents the number of cones that must be contained in the unit area of the rod-free part of the fovea, that is, the minimal foveal unit area is at least this large. According to the acuity-intensity discrimination theory, this minimal unit retinal area, which is based on intensity discrimination data, should be able to mediate all visual acuities as well because intensity discrimination provides the basis of visual acuity. Hecht supports this analysis by pointing out that the unit retinal area, based on König's acuity data, would be 0.04 mm^2, which, according to Helmholtz, would contain 540 cones (cone density in the fovea was believed to be 13500 per mm^2 at that time) 'the same number as that derived from intensity discrimination' (Hecht, 1928, p. 276).

The next step was to show that intensity discrimination, which could explain visual acuity, was "a necessary consequence of the photochemical system" action (p. 269) that Hecht had developed to explain dark adaptation. The general form of this photochemical system can be summarized as follows:

$$S \underset{\text{dark}}{\overset{\text{light}}{\rightleftarrows}} P+A$$

where S is a sensitive substance, P and A are decomposition products and also precursors of S.

The sensitivity of each retinal receptor depends on the concentration of decomposition products necessary to generate an action potential in the nervous element monitoring the state of the receptor. The total number of active elements is a linear function of the concentration of decomposition products. It follows that the number of active elements can be described by the following equation:

$KI = x^2/(a—x)$, where a is the initial concentration of photosensitive substance S, x is the concentration of the decomposition products (A and P), I is intensity and K is a constant.

The final form of Hecht's quantitative treatment of visual acuity in his photochemical system was presented by Hecht and Mintz in 1939. The following treatment of the acuity-intensity relationship is based on the equations presented in that paper.

$\Delta I/I = c[1+1/(KI)^{1/2}]^2$, where c is the minimal value of $\Delta I/I$ at the highest value of I, and K is the reciprocal of the intensity at which $\Delta I/I$ is 4 times the minimal value. This represents the photochemical equation as applied to intensity discrimination data, but the minimum angle resolved (the acuity threshold) can be viewed as a function of $\Delta I/I$ and therefore: $a = b' \Delta I/I$, where a is the visual angle and b' is a constant.

Then: $a = b[1+1/(KI)^{1/2}]^2$ where $b=b'c$. The b fixes the curve on the ordinate and K fixes it on the abscissa.

4.2.2. Summary of Hecht's theory

Briefly, Hecht's theory of visual acuity can be summarized as follows: the density of active receptor elements in the retinal mosaic sets the limit on visual acuity; the larger the number of active elements per unit retinal area, the better the acuity. The number of receptor elements is, obviously, fixed by the anatomy of the retina, so the variation in acuity observed, when acuity is tested on different background illluminations, must be produced by variations in the number of anatomical elements that are functionally rather than anatomically present, under a given set of illuminating conditions. Individual retinal elements respond differentially to the same amount of photochemical decomposition compounds. A few have very low thresholds, a few very high thresholds, while most thresholds cluster near the middle. In other words, thresholds are normally distributed. The intensity

---

[*] This equality of JND and DL is a common assumption, going back to Fechner (1860), but it is worth noting that it is only an assumption – the JND and the DL are quite different entities – the former is mental; the latter, physical.

of the stimulating light determines the amount of photochemical decomposition. This relationship between stimulating intensity and the concentration of decomposition products predicts both intensity discrimination and the discrimination of visual angles in the acuity task. In other words, visual acuity is based on intensity discrimination.

4.3. Criticism of Hecht's theory

Hecht's theory was criticized on three grounds, namely, some of the facts used as basis of the theory were not correct, the theory was insufficient to account for some well-established facts and, finally, the theory was not the only or the best interpretation of the facts. In short, when all was said and done, there really was not a great deal of lasting merit in the acuity theory other than its simplicity and breadth. We will next give examples of some of the problems which brought it down.

First, there were problems arising from facts known in Hecht's day. Hecht makes much of the 'excellent' quantitative agreement between the number of cones in the 'minimal unit foveal area' and the number of discrete steps (JNDs) of intensity discrimination across the range of lights to which human vision responds. This quantitative agreement is the heart of his intensity discrimination theory of visual acuity. It rests on Helmholtz's (1896) estimate of 13500 cones per square millimeter in the fovea. In 1941 Polyak provided a better, more modern estimate of 55000 cones per square millimeter. This density is for the foveal floor. The cone density in the central bouquet of the foveal floor, whose diameter is about 20 minutes of arc, is still greater. The foveal bouquet would be the likely preferred fixation position when the absolutely best acuity is desired, providing, of course, that the intensity of the target background is sufficient to stimulate the densely packed cones found in this retinal region. So, within 2 years of Hecht and Mintz's (1939) definitive quantitative confirmation of the intensity discrimination–acuity theory, the minimal retinal unit area in the central fovea (the 0.04 mm^2) claimed by Hecht (1928), on the authority of Helmholtz, to contain 540 cones was actually shown to have more than 2000 cones. Recall that 540 cones was within 2 cones of the number needed to account for König's number of JNDs of intensity discrimination (namely, 542) so as to bring intensity discrimination into agreement with the fineness of grain of the retinal receptor mosaic required to resolve details of acuity targets. This, of course, is only a quantitative disagreement, but a factor of more than four was considered by some sufficient to raise serious doubts about the soundness of the underlying idea that visual acuity is based exclusively on intensity discrimination.

There were additional problems. The number of JNDs of intensity depend on the conditions under which the difference thresholds (DLs) are measured and also on the statistical convention used to calculate the DLs used to estimate the JNDs. The particular conditions König used, the size of his test field and/or his exposure duration, for example, affected the size of his DLs. DLs measured with fields of different sizes or durations would give different values. These facts imply that the splendid quantitative agreement emphasized by Hecht is actually a coincidence based on König's particular choice of conditions rather than on a general relationship between visual acuity and intensity discrimination. Other problems were clearly evident before the sudden untimely death of Hecht in 1947. Byram's (1944) criticism of Hartridge's (1922) use of Rayleigh's equations for a square aperture on the round pupil of the eye applies to Hecht's calculations of retinal light distributions of acuity targets as well as to Hartridge's. Byram calculates Hecht's error with the minimum visible dark bar used by Hecht and Mintz (1939) to be smaller, however, only about 15% – a modest difficulty when compared to the factor of four error in estimating the number of cones in the minimal foveal unit area. The similarity of the form of the $\Delta I/I$ vs. I and VA vs. I functions, so important to Hecht's formulation, as well as the exact retinal light distribution of his acuity targets, can also be questioned. The acuity function seems to be stable or to continue to improve even at the highest intensities measured, whereas intensity discrimination does deteriorate at high levels of intensity as Hecht believed (see Walls, 1943, p. 493, for a review of these facts in the context of Hecht's theory).

There is also reason to wonder about the adequacy of Hecht's photochemical theory to explain facts well-established when the theory was first proposed. The decline of interest in Hecht's theory in the 1940s derives, in part, from his making a poor guess about what he could choose to ignore in the interests of developing his theory of acuity within the limitations of the simple biochemical model he preferred. Hecht chose to ignore the role of eye movements in visual acuity – a problem that had already received serious attention when he began working on his general biochemical theory of visual processing. It continued to interest other investigators during the period from 1921 to 1942 while Hecht was extending his ideas to encompass visual acuity as well as adaptation, intensity discrimination, color and the absolute threshold. Weymouth and his co-workers had published their dynamic theory of visual acuity, based on Hering's idea of the mean local retinal sign, at about the same time that Hartridge was publishing his acuity–intensity discrimination paper that Hecht would build upon within a few years. Hering had emphasized the importance of eye movements and the mean local sign to explain how a straight line can appear straight despite the fact that it is imaged on a receptor mosaic believed at the time to contain cones arranged in a rather haphazard spatial configuration. Fig. 4, taken from Walls (1943), ilustrates Averill and Weymouth's development of Hering's idea as applied to the perceived straightness of edges.

Hecht ignored this problem of how the sharpness and straightness of a straight line can be perceived when its proximal stimulus would be expected to contain offsets and jagged edges. This was a good thing to ignore. This 'problem', taken seriously by visual scientists in the first half of the present century, is analogous to the problem of how the proximal stimulus, known to be an inverted image of the distal world on the retina by the beginning of the 17th century, is perceived as upright. It is hard at this point in time to understand why perceived straightness continued to worry people long after the phenomenology of the upright ceased to be troublesome. It is not obvious why anyone would expect the physiological correlate of a perceived straight edge to be a straight edge in the receptor mosaic any more than that the physiological correlate of an upright world need be an upright image on the retina. Hecht's failure to worry about this problem does not seem to discredit his work despite Walls's (1943) contrary claim. There were, however, other, more serious questions about the sufficiency of

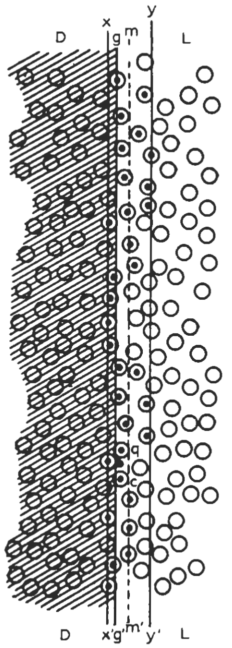

Fig. 4. Illustrates Weymouth's theory of vernier and stereoacuity. The diagram shows retinal conditions at the margin of a stimulated area. D–D is in darkness, L–L is illuminated. The geometrical margin of the image is g–g'. The cones are shown as circles. Cones a, b and c (near the bottom of region g–g') have local signs whose 'center of gravity' is amidst them, tending to pull b to the right. This action among all of the cones cut by g–g' smooths the percept of the contour despite the raggedness of the line of cones concerned. Furthermore, normal nystagmus shifts g–g' back and forth between the extreme positions x–x' and y–y', so that m–m' represents the center of gravity of all the points stimulated, and is the 'local sign' of the percept. The localization of this percept is independent of such factors as the size of one cone. (From Walls, 1943)

Hecht's approach that were worth considering. For example, Averill and Weymouth (1925) had shown that visual acuity depends on the length of the test stimulus up to some critical value of about 30 minutes of arc. It is difficult to imagine how a photochemical theory such as Hecht's can deal with such a fact. Hecht used long stimuli in his acuity experiments (e.g., wires subtending about 9–14 degrees of arc in Hecht and Mintz, 1939), but avoiding the effects of target length in this way does not solve the problem for a photochemical theory introduced by the fact discovered by Averill and Weymouth and used to support their dynamic theory.

Even more important to Hecht's theory was its inability to handle the effects of adaptation on visual acuity. Retinal adaptation has a clear photochemical basis and the failure of Hecht's theory to deal with the influence of this variable did a great deal to undermine the theory. The critical experiments on adaptation were done by Craik (1938, 1939). According to Hecht's theory, visual acuity depends on the intensity of the stimulating light and intensity differences in the retinal light distribution produced by the acuity target. Acuity will be best when the intensity of the stimulating light is sufficient to exceed the threshold of even the least sensitive receptors, providing the receptors are ready to catch quanta because their bleached photopigments have been allowed sufficient time to regenerate completely. So, the best acuity will be obtained when an intense stimulus is delivered to a dark-adapted eye. This is not the case. Craik found that acuity is best for a given level of test intensity when the eye is adapted to the test intensity before acuity is measured. Acuity suffers markedly when it is tested in a dark-adapted eye with even moderately intense backgrounds.

Hecht's receptor 'recruitment' proposal, the term he used to describe the increases in functional density attributed to the normal distribution of receptor thresholds, was also criticized on physiological grounds as soon as modern single unit recordings from visual neurons became available. There were also experiments by Senders (1949) and Nachmias (1958) showing that acuity depends on the subjective brightness, rather than the intensity, of the background field – another result not amenable to a simple photochemical explanation.

There is another class of criticism that is particularly germane to our treatment of Hecht's theory. Remember, we undertook a rather detailed examination of Hecht's photochemical theory for two rather different reasons. The first was because his theory was the most developed of the static theories of acuity, that is, theories that ignore eye movements. An appreciation of the most complete static theory sets the stage for an appreciation of theories that make eye movement a necessary requirement for acuity. The second reason we chose to go into such depth was because Hecht's quantitative theory was very influential in its day and criticism of this theory contributed to the development of an awareness of modeling issues in the contemporary visual science community that is often lacking in the much younger contemporary oculomotor community. An appreciation of such issues needs to be developed in the oculomotor community if we are to approach the level of theoretical sophistication that was apparent in our parent discipline 40 years ago. Specifically, how should one go about testing a model? Consider Senders treatment of this question in 1948.

"Hecht says that the distribution of sensibilities in the manner of populations is fundamental to his theory of visual acuity. The basis of this distribution is the photostationary state equation. Hecht feels justified in arriving at this conclusion because the photostationary state equation curves 'fit' the obtained data. This question of curve-fitting deserves extensive consideration, but only a few words will be devoted to it here. What is the most acceptable criterion of goodness of fit? In general, there are three main classes of criteria: (a) inspection, (b) statistical (e.g. least squares) and (c) parametric analysis. Each type of criterion is suitable to some types of data, and in a practical sense may be inapplicable to others.

"Essentially the same curve may be obtained from totally different equations. Crozier, for example, has pointed out (1937) the complete formal identity of the log logistic and the photostationary state equation as used by Hecht. In many cases, the curves predicted by the photostationary state equation and those predicted by the normal probability integral are so similar as to be indistinguishable by any visual criterion, and a statistical criterion is sometimes also inadequate. In such a case, the most suitable way to distinguish between the curves is by an analysis of the parameters of the function... When Hecht says that his curves fit his data, he means that a good visual inspection fit is obtained. In the absence of further analysis, this cannot be taken to mean that the photostationary state equation describes the data better than any other." (pp. 491–492)

Senders goes on to show that Crozier, her main doctoral advisor at Harvard and an undergraduate and graduate school classmate of Hecht's at CCNY and Harvard, had shown (1937) that Hecht's treatment of visual acuity as a special case of intensity discrimination was often arbitrary and not the best interpretation of the functional relationships observed.

5. Marshall and Talbot's new dynamic theory

5.1. Hartline's influence

Soon after Hecht's static theory of visual acuity had been confirmed in its final quantitative form by Hecht and Mintz (1939), Talbot and Marshall (1941) and Marshall and Talbot (1942) introduced a new dynamic theory that would dominate research on visual acuity for more than 20 years before it was replaced by the application of Fourier optics to the problem of spatial vision. The Marshall-Talbot dynamic theory placed emphasis on three things, namely, the cortical magnification factor, averaging in the visual cortex of signal transients generated in the fovea, and generation of these transients by tremor of the fixating eyeball. The appearance of this theory in the early 1940s was timely in large part because Hartline (1938, 1940) had just demonstrated phasic, as well as tonic, discharges from the ganglion cells of the frog retina – discharges to changes in stimulation, as well as discharges that began when a light came on and continued as long as the light was present, albeit at a diminished frequency. The latter kind of response, which was the only kind of response that had been observed in the compound eye of the horseshoe crab, was not tuned specifically to transients and would not, therefore, have led theories in the direction of a new, physiologically grounded, dynamic theory of visual acuity. Alternatively, the new transient neurons discovered by Hartline in the frog, which has a simple eye with a retina somewhat like the human being's, could and did provide the impetus for the development of such a theory. Neurons that signalled when a light came on and/or when the light went off were clearly capable of being driven by lights and shadows that were moved in and out of their receptive fields by eye movements. Hartline (1938) introduced the term 'receptive field' of a neuron in the same paper in which he reported the discharges in response to changing stimulation – terminology and observations at the core of modern visual neurophysiology.

The Marshall-Talbot theory was reviewed rather sceptically by Walls (1943) soon after its appearance. It was reviewed somewhat more sympathetically by Senders (1948), whose main criticisms were directed largely to Hecht's theory (see above), and subsequently by Falk (1956), whose review of visual acuity is largely devoted to criticisms of the Marshall-Talbot theory near the end of its useful life. These reviews, as well as the source papers, provided the material used in the following discussion and the reader is directed to them for additional details.

5.2. Walls's scepticism about the Marshall-Talbot theory

Walls (1943) introduced his criticism of the Marshall-Talbot dynamic theory by way of his discussion of the earlier publications of Weymouth and his associates (see sections 3.2 and 3.3), who had based their dynamic theory on Hering's idea of averaging 'retinal local signs'. (Walls's approach to this material was very clear and we will base our treatment of the Marshall-Talbot theory on his paper.) Hering introduced the idea of retinal local signs as processes which could compensate perceptually for irregularities assumed to be present in the spatial distribution of the receptor elements making up the retinal mosaic. The average of such retinal local signs can provide the physiological correlate of a straight-looking edge from the jagged edge of stimulation assumed to be present in the proximal stimulus. The concept of average local retinal signs can also provide a mechanism for abstracting the very tiny offsets in a proximal stimulus that are produced by a vernier acuity target (see Walls's illustration of Weymouth's theory reproduced in Fig. 4).

Weymouth used the assumed irregularities in the retinal mosaic coupled with the involuntary, continuous oscillations of the eyeball (physiological nystagmus) to facilitate the resolution of such details as straight edges and vernier offsets. This was accomplished by assuming, as Hering did, that each retinal receptor element had a unique directional local sign, whose activity would produce the percept of a point of light in a unique direction in visual space. When two such adjacent elements were stimulated, the perceived direction would be the spatial directional average of the local signs of each of the elements. Jagged edges in a proximal stimulus caused by irregularities of the retinal mosaic would be averaged out in this way. Differences in such averages could also provide reliable indications of offsets in vernier acuity targets. In short, characteristics of the receptor grain do not, in themselves, limit the perception of spatial details once average directional local signs, rather than discrete unrelated activity of individual receptor elements, provides the physiological correlate of the perception of the sharpness or position of spatial details. The longer the contours over which the local signs could be averaged, the larger is the number of elements contributing local signs to the average; the larger the number of elements contributing to the average, the sharper the edge. A larger number of elements allows greater reliability of the averages, increasing the likelihood of detecting any offset that might be present in the edge. These benefits derive from ordinary statistical considerations – averages are more likely to be accurate and differences in these averages can be estimated more reliably when sample size is increased. Confirmation of this theory can be had by showing that acuity improves as length of contour increases and also that acuity improves as exposure duration increases because increasing exposure duration allows more time for the tremor of the eyeball to contribute samples to the calculation of the average local signs. The former prediction was confirmed by Weymouth and his coworkers (see section 3.3), but their confirmation of the effects of exposure duration on acuity were more ambiguous theoretically as well as less clear-cut experimentally. The theoretical significance of exposure duration for dynamic theories of visual acuity would remain uncertain for many years until exposure duration could be varied with acuity targets stabilized on the retinal surface. This procedure permits unconfounding the contribution of the number of independent elements contributing to a particular sample with the contribution of the effects of the same sampling elements acting for longer periods of time (see Riggs et al., 1953, discussed in section 6.2.1, for the answer to this question).

The dynamic theory of Marshall and Talbot was much more elaborate than the dynamic theory of Weymouth and his coworkers. Part of the interest it engendered probably came from its complexity. If the visual system and the brain are complex, theories of their function should also be complex. Not everyone was prepared to adopt this view. Consider their theory as described by Walls (1943).

Marshall and Talbot came to the problem of visual acuity from their work on the physiology of the visual system and "erected a ponderous machinery with which they can explain almost everything that happens in vision. It is based upon supposed anatomical relationships and to an even greater extent upon such neurophysiological phenomena as recovery cycles, multiplication of paths, reciprocal synaptic overlap, facilitation, channeling, funneling, and peaking. These matters are most difficult for all but a chosen few – not including the writer – to comprehend fully. One tangible experimental fact is that two minutes of arc of the Rhesus monkey fovea (= ca. 9 μ) project to one linear millimeter of the visual cortex. Talbot and Marshall believe that the effective ratio of cortical ganglion cells to foveal cones is at least 600:1, which could account for vernier and stereo-thresholds very much smaller than those actually observed – whose smallness is already so frightening" (Walls, 1943, p. 502).

Walls often showed himself to be a man of considerable prescience and good taste in his numerous publications (e.g., he was not terribly keen about Hecht's theory either). The best Walls could manage to say about the Marshall-Talbot theory is that it is essentially the Weymouth version of "Hering's local signs pulled up to the *cortical* visual cells (where it probably always belonged) in order to account for the precision of spatial localization" (Walls, 1943, p. 502).

5.3. Assumptions of the Marshall-Talbot theory

If Hecht's theory of acuity could be criticized because it seems to be overly simplified, the Marshall-Talbot theory cannot be criticized for the same reason. To illustrate, seven mechanisms are assumed to be necessary to explain performance in the various kinds of acuity task, ranging from the minimum visible dark bar, through gaps and gratings and back to the stereo and vernier task. These 'mechanisms' include (1) diffraction by the pupil, which produces a statistical spatial distribution of light from every point in an acuity target; (2) physiological nystagmus, which applies the graded light distribution produced by diffraction at the edges of the pupil to the receptors; physiological nystagmus, itself, follows a statistical distribution; (3) reciprocal overlap between neural pathways, which provides a mechanism for increasing a gradient of excitation and producing peaks in the gradient of excitation because of temporal summation; (4) the neural recovery cycle, which can amplify or depress the level of excitation, depending on when in the recovery cycle the stimulus is presented; (5) multiplication of cells in the visual pathway in the sense that each retinal receptor projects to a distribution of cortical cells rather than to a single cell (it does not always do this by means of the same pathways; in other words, the specific population of cortical cells is a random variable); (6) threshold mechanisms in all of the cells in the pathway which can pass more or less of the information received; and, finally (7) an operating range of neural activity that covers about two log units of intensity, depending on the level of adaptation. The level of adaptation may be determined by photochemical processes. The total number of impulses arriving at the cortex is a function of the number of receptors active at that level and the number of impulses delivered by each receptor to the next level. The number of these impulses is determined both by the intensity of the stimulus and by the way these impulses are modified by other events in the visual pathway. These impulses may be facilitated, inhibited, peaked, etc. This menu was proposed as sufficient to develop a complete theory of visual acuity. Fortunately, this theory was built on estimates of the fineness of the receptor grain and the size and frequency of physiological nystagmus, which provided the transient stimulation-generating samples for the cortical averaging process. We say 'fortunately' because it will prove possible to remove the underpinnings of the scheme without getting into the nitty gritty of its other assumed mechanisms.

5.4. An example of how the theory worked

We next describe how this theory worked before getting into the experimental difficulties it encountered once visual scientists were able to give up their war-related research and return to basic science. Some essential features of the theory are illustrated diagrammatically in Fig. 5, which reproduces a figure from Falk (1956).

Fig. 5 shows the proximal stimulus of a bipartite field made up of a light and a dark half. The curve is the intensity distribution of the border between the light and dark halves. The slope of the distribution

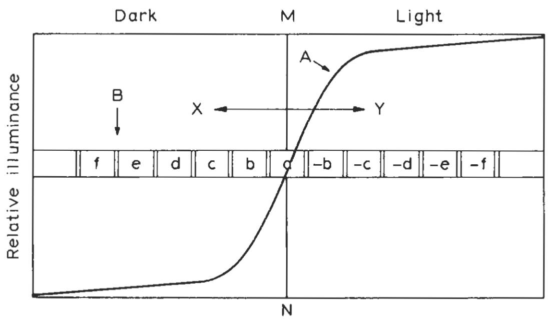

Fig. 5. The distribution of illuminance on the retina across the geometrical boundary (MN) separating Light and Dark halves of a field. See the text for details. (From Falk, 1956)

is caused by diffraction and aberrations of the edge in this proximal stimulus. The distal stimulus would be an intensity step, containing an abrupt transition from the light to the dark side of the stimulus (the MN border in Falk's figure). B is the foveal cone distribution in which cones have a center-to-center distance of 1.5 μm (a retinal distance corresponding to about 19 seconds of visual angle). The distance XY represents the average amplitude of physiological nystagmus as it was believed to be by Marshall and Talbot, that is, about 2 minutes of arc, which would contain about 6 cones. XY marks the limits of tremor for cone *a* whose stimulation by the blurred border of MN is represented by the light distribution curve A. As physiological nystagmus vibrates the retinal light distribution back and forth across cone *a*, cone *a* will receive the greatest rate of change in illumination because it lies on the steepest part of curve A. Cone *b* receives a lower rate of change and would, therefore, generate a lower rate of neural firing. The relationship between firing rate and rate of change of illumination proposed by Marshall and Talbot had been observed in the eccentric cell of the lateral eye of the horseshoe crab by Hartline and Graham (1932). These authors had also demonstrated that firing rate depended on the amount of illumination. It was these two factors, the combination of differential rate of change of illumination and the amount of illumination across the cones, that create a gradient of firing rates, which is peaked in the center of the row in the example illustrated in Fig. 5. Cone size is important in this theory, unlike the Hering-Weymouth theory, which averages local signs in order to produce the physiological correlate of an edge. Marshall and Talbot were quite specific about the importance of the relationship of cone size to the retinal light distribution, going so far as to point out that: "Smaller receptors would be useless, because though traversing the optical gradient oftener, they would gather proportionally less brightness differential. The limiting retinal factor in acuity seems to be the relation of receptor width to the highest optical *gradient* in a moving pattern, rather than the average static differential illumination on one cone, compared with its neighbors." Multiplication of pathways and reciprocal overlap on the way to the primary visual area in the cortex produce additional sharpening of the peak produced in this way in the retina, as does the 'neural recovery cycle', which favors the successful centripetal travel of the signal peaks as compared with signals arising from other receptors.

5.5. Criticism of the Marshall-Talbot theory

5.5.1. Byram's criticism

Criticisms of the dynamic theory of Marshall and Talbot were as diverse as the backgrounds of the critics. Consider for example, Byram (1944), who approached the problem of acuity from the viewpoint of a physical scientist, and developed a Hartridge-Hecht variety of static theory in which 'energy rate discrimination', rather than 'illumination discrimination', which he correctly attributed to Helmholtz, provided the basis for the discrimination of details in the retinal light distribution produced by an acuity target. In Byram's words, Marshall and Talbot "use a neuro-physiological approach in which most of the emphasis is on the post-retinal system. According to this point of view, neural mechanisms play an important role and visual acuity and contour recognition are achieved to a large extent in the visual cortex and associated systems. They point out that the grain of the cortex is in effect much finer than the grain of the retina and attribute certain accomplishments of the visual system, such as the perception of vernier offset, to both the fineness of the cortical grain and the operation of certain neural mechanisms. This would mean that the cortical image and the test object would have in common certain energy differentials which would be lacking in the retinal and photochemical images. In view of the physical concept of entropy, it is hard to see how the energy in the cortical image could be in a more highly organized state, with respect to the test object, than the energy in the retinal image. This would be equivalent to obtaining a sharp enlargement from a blurred photographic negative by manipulating the focusing mechanism of the enlarger so as to obtain an 'equal and opposite blur" (pp. 736–737).

5.5.2. Empirical tests of the theory

Entropy can proceed slowly in science, as well as in the Cosmos, and the Marshall-Talbot theory was still under serious consideration, albeit with continuing scepticism, 12 years later, shortly after one of its most basic assumptions, namely, an explicit quantitative functional relationship between cone density at the foveal center and spatial and temporal characteristics of physiological nystagmus, had been discredited in a number of difficult, carefully executed, experiments (Ratliff and Riggs, 1950; Riggs and Ratliff, 1951; Ratliff, 1952; Barlow, 1952; Ditchburn and Ginsborg, 1953). Marshall and Talbot had used the measurements of physiological nystagmus provided by Adler and Fliegelman (1934) to construct their theory. Adler and Fliegelman had used an optical lever with a plane mirror mounted at its fulcrum on the eye to measure fixational eye movements, apparently without appreciating the angular amplification factor of two inherent in this kind of optical arrangement. They reported that the average amplitude of physiological nystagmus was more than 2 minutes of arc – a value that would permit the proximal stimulus produced by an acuity target to flutter over more than 6 foveal cones once cone-to-cone separation in the foveal bouquet was known to be less than 1/4 minute of arc. Polyak (1941) had published such values for the primate retina precisely when such measures of cone density were needed by Marshall and Talbot to support their dynamic theory. Polyak's views of the potential for reciprocal overlap in the afferent pathways arising in the central fovea, however, did not offer direct support of this aspect of their theory (Polyak, p. 431). Adler and Fliegelman had also reported that the frequency of physiological nystagmus was high, ranging from 50 to 100 Hz. This range of frequencies seemed capable of synchronizing with the neural recovery cycle.

Ratliff and Riggs (1950) corrected Adler and Fliegelman's data for their neglect of the optical lever's amplification factor. This reduced the amplitude of physiological nystagmus, actually observed by Adler and Fliegelman, to about one minute of arc. Physiological nystagmus with an amplitude of only one minute of arc would sweep an acuity target image over about 3 foveal cones on average – a rather small sample of receptor elements from which to calculate reliable transient peaks signalling illumination differences – the essential requirement for all later processing proposed in the Marshall-Talbot theory. Fixational eye movement characteristics present additional difficulties for the theory once allowance is made for drifts of the eye and fixational microsaccades, both of which would move critical features of the proximal stimulus onto new receptors – shifts that would broaden the peak of the average transient generated by the moving stimulation. In electrical jargon, DC shifts in average fixation position, produced slowly by drifts or abruptly by microsaccades, would blur the peaks generated by physiological nystagmus, the relatively fast AC component in the fixational eye movement pattern.

Even more troublesome, however, were Ratliff and Riggs's (1950) measurements of physiological nystagmus, which they reported to have a median amplitude of only 17.5 seconds of arc and a frequency ranging from 30 to 70 Hz. The distribution of amplitudes was such that nystagmic movements as large as one minute of arc were rare. The somewhat lower frequency observed by Ratliff and Riggs was not an insurmountable problem because Marshall and Talbot had not worked out the quantitative details required to synchronize physiological nystagmus and the neural recovery cycle. The difference of almost a factor of 7, however, in the average amplitude of physiological nystagmus was critical. Seventeen and a half seconds is about the diameter of a foveal cone in the central bouquet where acuity is best. Physiological nystagmus, the high-frequency fixational eye movement so essential to the dynamic theory of acuity, was actually found to be so small that it seemed unlikely to have any functional significance whatsoever. It would allow the intensity distribution of an acuity target to sweep back and forth over only a single cone – precisely the situation described by Helmholtz almost a century earlier to be incapable of providing the nervous system with a signal indicating a difference in the stimulation in one or another part of the proximal stimulus. Barlow (1952) and Ditchburn and Ginsborg (1953) confirmed Ratliff and Rigg's values, Barlow agreeing with these authors that such results cast doubt on the Marshall-Talbot theory. Ditchburn and Ginsborg, however, thought that although such values did not support the theory, they were, nevertheless, compatible with the theory. This seemingly odd interpretation probably stems from Ditchburn's early committment to an important role for fixational eye movements in visual processing – a preference that would persist and become increasingly difficult to maintain in later years (see Ditchburn, 1980; and Kowler and Steinman, 1980; for an exchange of letters on this point couched in terms of the functional significance of fixational microsaccades – the miniature fixational eye movements that Ditchburn chose to emphasize when he abandoned the functional significance of physiological nystagmus).

5.6. Motor theories of perception after Marshall-Talbot

By the middle of the 1950s, the Marshall-Talbot dynamic theory, which made small high-frequency tremor a necessary condition for visual acuity, was fatally compromised by new measurements of fixational eye movements. These measurements required the development of very sensitive, accurate eye movement measuring instrumentation. This instrumentation was soon elaborated so as to stabilize the retinal image of visual targets – a development that ushered in an era in which a lot of effort was expended in the study of the role of transient stimulation in initiating and maintaining the visibility of simple and complex visual displays. Most of this work was done within a relatively loose theoretical framework. Lessons learned from the fate of the Hecht static and Marshall-Talbot dynamic theories of visual acuity and their derivatives, for example, Jones and Higgins (1948), had apparently tempered enthusiasm for attempting elaborate general theories of basic visual processes as well as compromising physiological nystagmus as an important contributor to basic visual processing. The treatments of more complex visual processes were also influenced by these new measurements of fixational eye movements. For example, Osgood's (1952, 1953) use of what he called the 'statistical theory' to explain 'figural aftereffects' (Köhler and Wallach, 1944) died along with its explanatory base. Osgood's 'statistical theory' was based on the Marshall-Talbot treatment of physiological nystagmus. By the mid 1950s, researchers interested in the role of eye movement in basic visual and more complex perceptual processes, such as the perception of form or size, began to confine interest to the functional significance of saccades and drifts. Only work on basic visual processes will be reviewed in this chapter. Hebb's motor theory of form perception (1949) and Festinger's (1971) 'efferent readiness' theory of visual extent will not be covered (see Steinman, 1976, or Murphy et al., 1974, for a review of this work, and Steinman, 1986b, for comments on more recent efforts; Skavenski (Ch. 5 of this volume) reviews the role of eye movement in the perception of direction and Wallach (Ch. 6 of this volume) reviews the role of eye movement in the perception of motion and shape).

6. Stabilized image research

Experiments with stabilized images were very prominent in the study of spatial vision from their introduction by Ratliff and Riggs in 1950 until the mid-1960s when the new direction of visual science became the application of Fourier analysis to optical problems ('Fourier Optics'). Schade (1956) was the harbinger of this approach for spatial vision, but earlier instances of its use have been cited (cf. Le Grand, 1967). The application of Fourier Optics to temporal factors in vision was pioneered by de Lange (1952, 1954, 1958), followed shortly by Levinson (1960). The application of Fourier Optics to spatial vision – the topic of this chapter – began with studies of image formation by the human eye (Westheimer, 1960), moved inwards to consider the relative importance of neural, as well as optical, factors in limiting spatial resolution (Campbell et al., 1966), and then extended the approach by making psychophysical measurements of the 'spatial modulation transfer function' (the Spatial MTF; currently called the Contrast Sensitivity Function (CSF)) of the human visual system by Campbell and Robson (1964). Thinking of the visual system as a spatial frequency analyser gave rise to the 'channel hypothesis' introduced by Blakemore and Campbell (1969). This hypothesis treats the visual system as though it contained a limited number of band-pass filters tuned to different portions of the spatial frequency distribution. This hypothesis seemed far-fetched to many, perhaps most, visual scientists active at the time. In retrospect such wide-spread scepticism seems unwarranted. The channel hypothesis, despite its novelty, only assumed that the physiological processes underlying spatial vision in the visual system are analogous to the kinds of processes long believed to serve color vision and qualitative dimensions in other sensory systems. To illustrate, a limited number of independent channels, assumed to be only three in both the Young-Helmholtz and Hering theories, had accumulated a good deal of empirical support in the century since the introduction of a 'channel hypothesis' in color vision by Helmholtz, who extended J. Müller's (1826) Doctrine of Specific Nerve Energies to 'hue' – the qualitative dimension in the visual modality. The channel hypothesis also brought theories of spatial vision into line with theories of auditory pitch discrimination. Helmholtz's (1863) auditory 'place' theory assumed a large number of independent neural channels, each tuned to a different temporal frequency of air pressure change, to explain the human being's capacity to discriminate several thousand pitches.

Once viewed in this broad historical perspective, the channel hypothesis, which flowed quite naturally from the attempt to use Fourier Optics in the study of spatial vision, fits in with a long tradition of 'labelled line' sensory theories going back at least to J. Müller. The channel hypothesis will be described in greater detail later after research from the initial period of image stabilization has been reviewed. At this point it is sufficient to note that the shift to Fourier Optics from traditional ways of studying visual acuity forced a change in thinking about spatial vision. In traditional approaches it had been appropriate to pit static theory against dynamic theory – static theories emphasized spatial factors and largely ignored temporal factors; dynamic theories included both but placed greatest emphasis upon the temporal changes caused by physiological nystagmus. This dichotomy ceased to be aesthetically pleasing by the late 1960s because both spatial and temporal properties of visual stimulation were intimately bound together when the new theoretical paradigm was adopted. The relative importance and interactions of these factors, however, in the search for the physiological underpinnings of spatial vision constitute an important area of current concern.

6.1. Early stabilized image research

Most of the specific interests and techniques of the scientists working with stabilized images during the almost 20 years that intervened between the first accurate measurements of the fixational eye movement pattern, which did the Marshall-Talbot theory in, and the beginning of interest in the channel hypothesis can be traced to three sources; Riggs in the USA, Ditchburn in England and Yarbus in the Soviet Union. Developments during the first 10 years were summarized in an unusually perceptive review by Fiorentini (1961) and reviewed again by Heckenmueller (1965) towards the end of this period of greatest initial interest. The material which follows is based, in part, on these reviews. Both reviews provide descriptions of the stabilizing techniques as well as summaries of the visual effects produced by stabilization.

6.1.1. Dependent variables in early research

The two main dependent variables in stabilization research were: (1) the 'disappearance time fraction', i.e. the proportion of time a stabilized image was visible during an experimental run, usually 30 or more seconds in length; and (2) unrestricted phenomenological reports of the appearance of critical features during stabilization. Both have serious methodological limitations. The first cannot be used conveniently to test quantitative models of visual function for reasons described below. The second can be even more troublesome. Phenomenological reports are readily influenced by expectations of the subjects. Such reports can produce reliable scientific information only when done by unusually discerning observers, otherwise providing masses of what, in the long run, amounts to useless speculation. Examples of the latter can be found in the stabilized image literature, where phenomenological reports, widely cited as supporting Hebb's and Gestalt theories of perceptual grouping, were subsequently shown to be more parsimoniously interpreted as response biases inherent in the linguistic constraints on open-ended verbal reports about ambiguous visual stimuli (see Steinman, 1976, for references and comments on this topic, which is beyond the scope of this chapter).

These dependent variables were used frequently in early stabilization research because of severe limitations on the length of experimental sessions. This restriction made it awkward to use traditional psychophysical methodology (see, for example, Woodworth, 1938, Ch. 17). Fechner's Method of Constant Stimuli was, and remains, the generally preferred traditional psychophysical method. This method bases its threshold estimate on hundreds of *independent* observations made in experimental sessions, typically lasting 2 or 3 hours. Such lengthy sessions are usually repeated a number of times to verify the reliability of the estimated threshold. Thresholds estimated in this way have calculable probable errors based on normal curve statistics. A large number of studies have demonstrated the appropriateness of normal curve statistics for the treatment of such independent psychophysical observations. More efficient 'interactive' procedures with known probability interpretations of their threshold estimates have come into use during the last 20 years, e.g. double random staircases, but neither the techniques nor appropriate statistical treatment had been worked out when stabilization research started (see Penner, 1978, for an introduction to the newer methods). The psychophysical procedures with unambiguous statistical treatment, available during the 1950s and early '60s, required a number of long experimental sessions. Contact lenses, fitted so as to be suitable for stabilization, could not be worn for more than 40–60 minutes and normally are not inserted more than once in 24 hours. Yarbus-type suckers placed even greater restrictions on wearing time. Such methodology does not readily lend itself to 'proper' psychophysical threshold measurement. It was these restrictions, rather than indifference to methodology, that encouraged the use of the disappearance time fraction and unrestricted phenomenological reports in most of the early stabilization research. 'Thresholds' are often reported in the old stabilization literature but they are usually based on the length of time a critical feature was visible, for example, the threshold was "the contrast at which the targets would be seen during 50% of the [30 second] viewing period" (Krauskopf, 1957*). The relationship of this kind of threshold measure to thresholds derived from large sets of independent tests with calculable probable errors has not, to our knowledge, been worked out. Ignorance of this relationship makes it risky to use the 'threshold' data from early stabilization experiments to test quantitative visual theories that were based on thresholds measured in traditional ways.

6.2. *Important early results with stabilized images*

The old stabilization literature contains some important generalizable facts about the relationship between eye movement and the perception of contrast and spatial detail. We will describe a few papers from this period in some detail. These papers were chosen because they introduce important ideas and have clear implications for the relationship of eye movements to spatial vision. The reader is directed to the reviews, cited above, for discussion of additional papers in what became a relatively voluminous literature in the period between 1950 and 1965 despite the difficult research methodology.

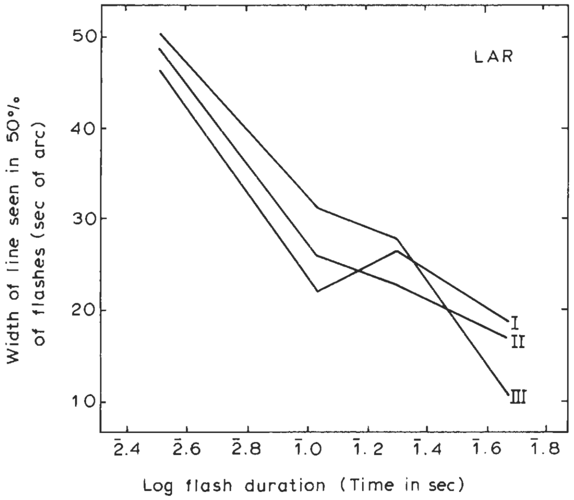

Fig. 6. Width of a line seen during 50% of flashes as a function of flash duration under viewing conditions I (stabilized), II (normal) and III (exaggerated motion, i.e., twice normal). The subject was L. Riggs. (From Riggs et al., 1953)

6.2.1. *Effect of exposure duration on stabilized vision*

Riggs et al. (1953) reported results of a short flash threshold experiment, as well as results from a 50% disappearance time experiment, in which the width of a dark line against a lighted background was varied under normal, stabilized and exaggerated movement conditions. The main results of this seminal experiment have stood the test of time. They also contain an observation, which will be replicated by other investigators, whose significance will not be appreciated for 27 years.

Two 'minimum visible' dark bar visual acuity experiments were reported. The dark bars were images, projected on a magnesium-oxide-coated screen, of wires, oriented vertically, whose visual angles ranged from about 6 to 93 seconds. In the first experiment the disappearance time fraction was measured during one-minute viewing periods. Fractions were measured under three conditions, namely, (I) a horizontally stabilized (compensated) wire seen against a homogeneous, moderately intense (5 ft-L) background, (II) normal, that is, unstabilized viewing of the same stimuli, and (III) twice-normal image motion while viewing the same targets. The results, in the words of the authors, were as follows:

"In Condition I, the 'compensated' condition, the black line target was clearly seen when it first appeared. The subject was surprised by the fact that the line was always at the center of the field regardless of eye movements. Soon, however, the line began to fade out. Finally, it disappeared altogether, so that the projected image seemed to consist only of a bright circular field. Occasionally, the bright field also disappeared; in these intervals the subject saw only the stationary annulus [an unstabilized surround used to keep the stabilized central test area from drifting beyond the operating range of the contact lens optical lever stabilizing apparatus]. A fine black line usually disappeared during the first few seconds of viewing, and failed to reappear later. Heavier lines took longer to disappear and often reappeared from time to time during one minute of

---

* Krauskopf was aware of the fact that thresholds inferred from disappearance-time fractions were not necessarily the same as thresholds measured in traditional psychophysical experiments and provided the rationale he used for drawing this inference. The reader should consult his paper for a description of the assumptions underlying his inferred relationship.

steady fixation.

"In Condition II, the 'normal' condition, the fading of the image did occur for the fine lines, but the lines reappeared sporadically. Heavier lines seldom disappeared.

"In the 'exaggerated' condition, Condition III, there was scarcely any disappearance of even the finest lines. The impression was that the target was 'locked in place' so that steady fixation was effortless, automatic" (p. 498).

In their second experiment flashed targets were used to measure 50% frequency of seeing thresholds. This was a genuine threshold experiment, rarely attempted in this early period. It suffered from limitations imposed by wearing tight-fitting contact lenses. Only 432 judgments for one of the subjects (the senior author) were obtained in two experimental sessions in which there were 18 experimental conditions (6 line widths and 3 viewing conditions) each of which was replicated 24 times – a modest number for estimating a threshold from a psychometric function based on Fechner's Method of Constant Stimuli (100 or more replications for each of the 18 conditions would probably have been obtained if the contact lens had not restricted the length and number of experimental sessions). Four flash durations were used, namely 34, 110, 213 and 472 ms (5 to 7 durations would probably have been used if a contact lens had not been worn).

The results of this flash experiment, summarized in Fig. 6, were described as showing "no striking differences for short flashes among the three experimental conditions. Consistently, however, the 'compensated' image of condition I yielded the best seeing for the shortest flashes. The 'exaggerated' condition, III, begins to excel at exposure durations beyond 0.2 sec. It is of interest to note that in all cases the intermediate Condition II [normal viewing] yields results which lie between those of Conditions I and III" (p. 500).

The conclusions drawn from these experiments are as follows:

"1. Vision is impaired under conditions such that the retinal image of an object remains essentially motionless with respect to the retina. During prolonged viewing under these conditions single-line test objects gradually disappear from view. The rate of disappearance is related [inversely] to the angular width of the line.

"2. Normal involuntary eye movements prevent the disappearance of test objects during long periods of observation. Exaggerated movements of the retinal image [twice as fast as normal movements] are even more effective in preventing the disappearance of images.

"3. In the case of short exposures (less than 0.10 sec) of test objects, the above relations appear to be reversed. Vision is poorer under conditions of normal or exaggerated motion than under conditions of reduced motion of the retinal image" (p. 501).

These conclusions were based on the following considerations. The disappearance of the stabilized target after prolonged viewing is "consistent with the theory that under uniform stimulation conditions each photoreceptor may attain a stationary state in which a minimum number of impulses are initiated in the retina". With shorter exposure durations the situation becomes more complex. When exposure durations are set at or below 10 ms, drifts are too slow to be significant and physiological nystagmus allows the eye to move through only about 5 seconds of visual angle. This discussion rested on a prior fixational eye movement experiment in which Riggs and Armington (1952) had shown that the eye was virtually stationary during 10-ms intervals and only moved about 25 seconds of visual angle during 100-ms intervals (these values would soon be confirmed by Riggs et al., 1954, as is shown in Fig. 7).

With flashes shorter than the exposure interval known as the 'critical duration' (an interval of about 100 ms in which the intensity of the stimulating light and its exposure duration are reciprocally related) the light–time product was held to be the primary determinant of the ability to see a detail. Acuity improved as the light–time product increased with durations shorter than the critical duration because 'the differential responses of stimulated and unstimulated retinal elements [in the proximal stimulus of an acuity target] might be

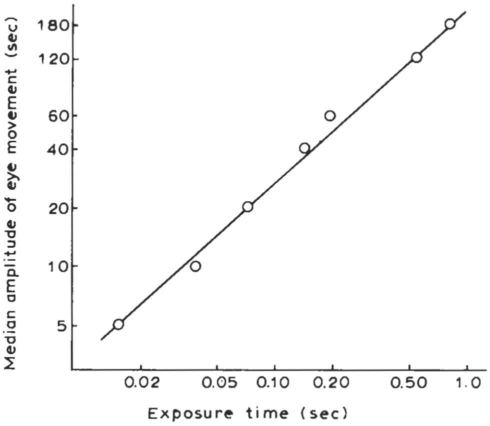

Fig. 7. The median extent of motion of the retinal image as a function of exposure time. (From Riggs et al., 1954)

expected to increase with exposure duration'. This explanation can be sufficient, providing the stimulus remains on the same retinal elements. Cone diameter in the central fovea is only about 18 seconds of visual angle, which means that target details would remain on the same receptors during about two-thirds of the critical duration. Stabilization would be expected to be beneficial for acuity once the exposure duration is sufficient to allow the unstabilized image to begin to move onto new receptors – a result consistent with Ratliff (1952), who reported that eye movements could be detrimental (stabilization helped) when a grating target was exposed for 75 ms. With exposures longer than the critical duration (100–500 ms), there is no longer reciprocity of intensity and time, and the eye can move through a couple of minutes of arc. Here, stabilization is only beneficial up to about 200 ms. Beyond 200 ms *exaggerated* motion of the retinal image leads to the best acuity. When exposure duration is longer than a half second, "eye movements, and even more those of double the amount in Condition III, clearly serve to maintain prolonged seeing. Hartline's experiment on the frog retina lends support to the idea that motion of the retinal image may serve to trigger 'on' and 'off' responses of individual retinal units... A brief, but inadequate summary of these points might be to the effect that eye movements are bad for acuity but good for overcoming the loss of vision due to uniform stimulation of the retinal receptors" (p. 501).

Riggs et al.'s (1953) interpretation of their results was described in some detail because it calls attention to two questions that continue to be of interest. First, why does the visual system perform better with exaggerated retinal image motion than with normal retinal image motion? Most, if not all, visual scientists like to believe that evolutionary pressures have tuned physicochemical and physiological processes rather exactly to psychophysical function. Why are normal fixational drift eye movements too slow by a factor of two to be consistent with this popular teleological belief?

Second, Hartline (1938, 1940) reported that only 50% of his frog units were phasic 'on-off'; the units Riggs et al. (1953) credited with maintaining vision with normal, unstabilized input. Prolonged failure to stimulate these units was held to cause the target to disappear. Note, however, that 20% of Hartline's remaining units were tonic 'on' units, similar to the units invariably observed in the horseshoe crab. They burst shortly after stimulus onset, pause and then continue to fire at a reduced rate as long as the stimulating light is present (the other 30% were called 'off' units, bursting when the light went off). Why were Riggs and his coworkers not puzzled by the failure of these tonic units to sustain vision when the image was stabilized? A possibility, not discussed by Riggs and his coworkers, that will come to be of interest subsequently is that the complete disappearance of a stabilized image depends critically on having low-contrast and/or low-intensity test stimuli. Both were modest in their experiment. Such conditions might have rendered the sustained 'on' unit responses subliminal, permitting complete disappearance of the stabilized portion of their display. The question of whether high-luminance, high contrast, foveally centered stabilized targets ever disappear completely will be encountered periodically in subsequent research. This question is still unsettled. We will discuss it further after we have considered the first publication in which the effect of motion, controlled by the experimenter, was imposed on a stabilized acuity target image – currently, a much-touted technique for studying spatial vision.

6.2.2. *Motion imposed on stabilized images*

Krauskopf (1957) was the first to publish detailed results of this kind of research (he credits Cornsweet and Riggs with a prior report at an Eastern Psychological Association Meeting in 1954). The logic behind this and subsequent work is the desire to control, and thereby simplify, the waveform, frequency and amplitude of retinal image motion while concurrent psychophysical measurements of spatial vision are made. If this can be done, the effect of retinal image motion on spatial vision can be understood. Once such knowledge is at hand, it may then become possible to understand the functional significance of the various motions observed in the natural fixational eye movement pattern.

Krauskopf used the Riggs et al. (1953) contact lens optical lever technique to stabilize his targets. The stabilized targets could be oscillated in a controlled manner by mounting a mirror in the stabilizing optical path on a galvanometer which could be driven by a signal generator. Targets consisted of a bright bar of variable width (10 sec arc, 1, 4 or 8 min arc of visual angle) presented at the center of a 1 degree diameter circular background field of 20 ft-L. This background field was enclosed in a 10 ft-L unstabilized annulus, which prevented the line of sight from drifting outside the range of the optical stabilizing system and also served as the fixation stimulus – the subject being required to maintain his line of sight at its center during tests. The 'contrast' of the test bar (the ratio of its brightness [photometric intensity] to the brightness [photometric intensity] of the background field) was the dependent variable used to estimate 'the 50% contrast threshold'. This 'threshold' was based on the percentage of time a particular target was reported as visible during a 30-s test interval (such threshold estimates require assumptions about the relationship of 'disappearance-time fractions' to 'frequency of seeing' measures; see above). The frequency of the imposed sinusoidal oscillations was varied (1, 2, 5, 10, 20 or 50 Hz) as was the peak-to-peak amplitude of these oscillations, which were 30 sec arc, 1, 2 or 4 min arc of visual angle.

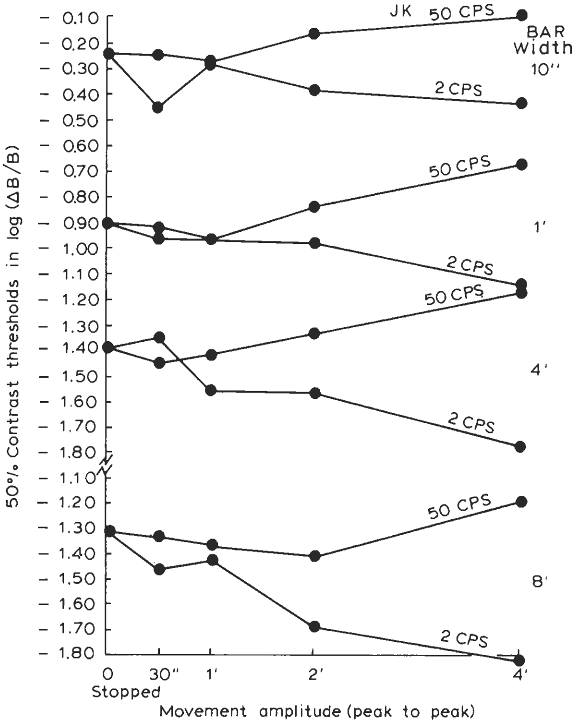

Fig. 8. Fifty percent contrast thresholds as a function of amplitude of vibration of 2 and 50 Hz for four bar widths, ranging between 10″ and 8′ of arc. There were four determinations for each datum point. (From Krauskopf, 1957)

Krauskopf found that "the effect of low-frequency motion [1 and 2 Hz and, to a lesser extent 5 Hz] was to improve seeing while higher frequency [10 Hz and above] had a generally detrimental effect. The curves suggest quite strongly, however, that motions at amplitudes below 1 minute of arc had little effect" (p. 743). He goes on to conclude that "The results of the present experiments as well as those of Cornsweet and Riggs (1954) suggest that the disappearance of stopped [stabilized] images during prolonged viewing is the result of the removal of the low-frequency components of normal retinal image motion. The present results suggest that oscillations at frequencies below 10 cps [Hz] may be constructive if they are of sufficient amplitude. The critical amplitude appears to be in the neighborhood of 1 min of arc (peak-to-peak). High frequency motion on the other hand appears to have detrimental effects. Again it seemed in the present experiments that the amplitude of these motions had to exceed 1 min of arc to have a demonstrable effect on the contrast thresholds" (p. 744). He goes on to point out that "it is dangerous to generalize from experiments with externally controlled sinusoidal motion to the normal fixation case" but ventures, nevertheless, to suggest that physiological nystagmus, whose amplitude is mainly below 1 min of arc, is not likely to be visually effective – a conclusion he treats as consistent with the earlier brief E.P.A. report of Cornsweet and Riggs, who imposed variable-amplitude 30, 50 and 70 Hz sinusoidal oscillations on a stabilized target. Krauskopf did not attempt to extend his results showing that low-frequency imposed motions were visually beneficial to the role of low-frequency fixational eye movements because he adopted the then-current view that "the low-frequency components are much more irregular [than physiological nystagmus]", and ended his elegant paper by leaving open the possibility that these low-frequency movements, along with fixational microsaccades, "may well be beneficial to maintenance of vision": both assumptions will continue to be the subject of some controversy many years later (see Ditchburn, 1980; Kowler and Steinman, 1980).

6.2.3. Exposure duration revisited with a variety of target types

Keesey (1960) examined the effect of stabilization and exposure duration (7 durations, ranging from 20 to 1280 ms) with three types of dark target presented on a lighted background (a single dark bar, two dark bars, one above the other in a vernier arrangement or multiple equally-spaced dark bars, in effect, a square-wave grating). The single dark bar was produced by a thin wire stretched across an aperture filled with collimated light. The grating and vernier targets were photographic negatives which could be mounted in the same aperture. The single bar would have very high contrast; the other targets somewhat less, depending on the photographic film and processing technique used to reproduce the 'opaque' portions of the test stimuli. The intensity of the 58 min arc diameter circular test field and its matched unstabilized annulus was not specified, but it was likely to be at least as high (20 ft-L) as in the Krauskopf (1957) report, described above, which used rather similar optical arrangements, that is, the test field was presented in Maxwellian view (a lens formed an image of the test field at the entrance pupil of the eye).

There were important methodological advances in this doctoral thesis (L. Riggs was chief advisor) in addition to the use of different kinds of acuity target presented with relatively high contrast on backgrounds of relatively high light intensity. Namely, traditional acuity thresholds were estimated by determining the threshold size of the critical spatial detail (the angular subtense of the single dark bar, of the vernier offset or of the bars making up the grating) by means of a 'forced-choice' Constant Stimulus Method. This means that the subject was forced to report a particular property of the test target on each trial. For the single line, its presence at any time during a test exposure was required to be reported, in other words, a single forced 'yes' or 'no'. For the vernier stimulus, the direction, right or left, of the lower segment was required; for the grating its orientation, horizontal or vertical. This kind of psychophysical methodology is still considered to be the optimal way of measuring thresholds, with the exception of the single line target, where its orientation would probably be varied in the same manner as the grating. These were, and remain, very ambitious experiments because a tightly fitting scleral contact lens had to be worn in sessions long enough to make measurements with normal, as well as with stabilized, targets – the design optimal for making valid inferences about effects of stabilization. The difficulties were not simply confined to running many long sessions under uncomfortable conditions. Using target orientation in the forced-choice procedure with the grating required that stabilization had to be achieved on the vertical, as well as on the horizontal, meridian – most previous and subsequent work with stabilized grating targets is less ambitious, confining itself to gratings with vertical bars and stabilization exclusively on the horizontal meridian.

It is not entirely clear from the publication that these stringent demands were completely met. The number of tests of each type of target were not described nor were the total number of replications so it is not possible to determine the number of tests contributing to each of the thresholds reported. There were probably only a modest number because the author found it desirable to report an 'additional' experiment to verify the main result. In this experiment only "two stimuli around threshold size were presented 60 times [in itself a modest number for this kind of measurement] for each of the exposure durations of 0.20, 0.75 and 1.00 sec" [7 durations were used in the basic experiment]. An additional complication can be seen in the fact that only tests with the vertical grating were used for estimating thresholds because there "was often a difference in thresholds between vertical and horizontal orientations of the grating lines." It is possible that these differences arose from the fact that stabilization was less complete on the vertical meridian (the meridian critical for a grating with horizontal lines) than on the horizontal meridian, where vertical lines are the critical detail. We suspect this because one of us knows from personal experience (Steinman, 1965) that it is difficult to orient a contact lens mirror to be exactly orthogonal to a line parallel to the line of sight even when it is 5 or more mm in diameter and mounted on a ball and socket joint at the end of a stalk cemented to the contact lens. It is difficult to imagine how this adjustment would be made and maintained from session to session with the 1.5-mm-diameter mirror embedded in the surface of the contact lens – the arrangement used in Keesey's experiments.

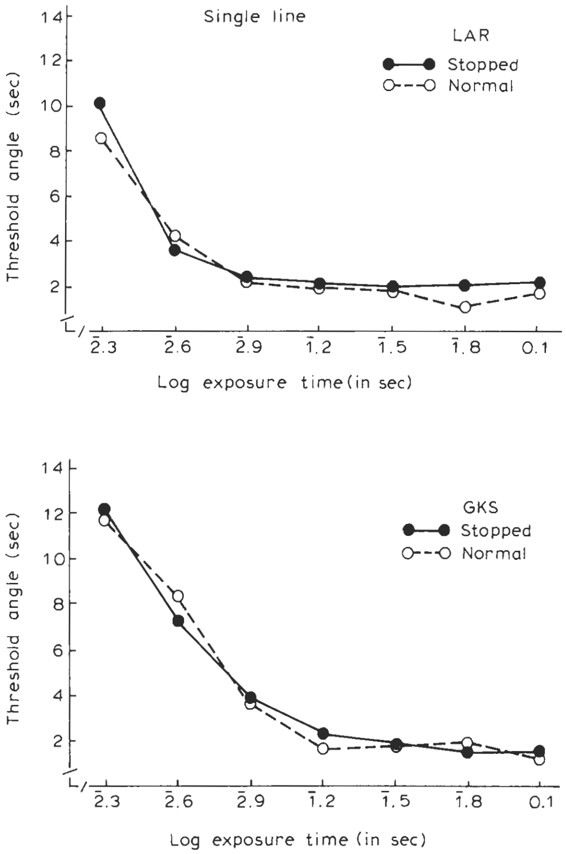

Fig. 9. Threshold curves for detection of single black lines as a function of log exposure time under stabilized (stopped) and normal viewing conditions. (From Keesey, 1960)

Quibbling aside, the results reported were very orderly, particularly for the single dark bar target. Keesey's graphs for this condition for both of her subjects are reproduced in Fig. 9, where it can be seen that stabilized and normal viewing were affected similarly by exposure duration across the entire range studied. It is also clear that acuity asymptotes to its best value at about 200 ms. Keesey "noted that for the relatively short stimulus durations used in this study no disappearance of targets was reported. A few isolated exceptions were the disappearance of very fine lines during the 1.280 sec exposure time with the stabilized image condition" (p. 772). Her results with the vernier and grating targets led to the same conclusion although the performance of her second subject, GKS, was less orderly than LAR with these test stimuli. At this point once it had been shown that acuity was the same with stabilized targets as it was when normal eye movement was permitted, it was possible to conclude that "acuity is mainly based on the discrimination of the spatial pattern of retinal illumination, regardless of any temporal changes of intensity pattern on the receptor cells" – clearly, a complete refutation of dynamic theory, Weymouth as well as Marshall-Talbot.

6.2.4. But were retinal images really stabilized?

Barlow (1963) was not convinced that they were and published a paper which stirred up appreciable controversy, both about the appearance of stabilized images and about the merits of alternative methods of stabilizing images. Prior to his report, both facts and methods had seemed to be relatively well-established. Barlow undertook his research because he felt that the results obtained by the main groups investigating stabilized images (the groups led by Ditchburn, Riggs and Yarbus) "are not in good agreement" with each other or with "information available as to what happens when a pattern of light is held unchanged on the retina". He pointed out that a variety of entoptic images (percepts arising from features located within the eyeball), such as the shadows of small retinal blood vessels made visible by moving a small point source of light or by light shining through the sclera, only remain visible for very short periods when the conditions required to render them visible in the first instance are maintained. Other entoptic phenomena such as Maxwell's spot and Haidinger's brushes show similar characteristics. The disappearance of such intrinsically stabilized images is rapid, complete and persistent. Observations such as these suggest that steady, unchanging stimulation produces only a short-lived percept which can only be reinstated by changing the stimulation once the initial percept has faded. Barlow also pointed out that the situation with visual afterimages is similar. They lose their fine details and then fade out completely. Periodic blinks or flashes of light are required to bring them back into view once they have disappeared. These well-known properties of entoptic images led Barlow "to the clear expectation that any arbitrary pattern of light held stationary on the retina will fade and cease to be visible within a few seconds of first presentation, but only Yarbus's results fit in with this." Both Ditchburn's and Riggs's group had reported that stabilized images disappear relatively slowly and reappear periodically, results at variance with the expectation from both entoptic phenomena and Yarbus's reports (1957a,b).

Very different stabilizing techniques were being used. Both Ditchburn and Riggs used a tight-fitting scleral contact lens to place a mirror at the fulcrum of an optical lever designed to stabilize targets located in the optical path. Their optical instruments usually incorporated provisions for convenient control of critical properties of test stimuli, such as their contrast, size or motion. These instruments were also typically laid out so as to make it convenient to alternate normal with stabilized viewing of the same target. Their research scleral contact lenses were fitted so as to rest on the limbus (the margin of the cornea and sclera). Pressing such lenses into place establishes suction between the lens and the eye, sometimes making the lens hard to remove after it has been worn for a half hour or so.

Yarbus's method was completely different. He used a rubber 'sucker' to establish negative pressure in a lightweight aluminum cone ('cap') whose base had serrated edges that rested on the margins of the cornea near the limbus. A high-power ('short focus') positive lens, mounted at the apex of the cap, allowed the subject to see a focused stimulus object located very near the eye. These stimulus objects were stabilized by attaching them to the aluminum cap by means of a girder. When the eye moved, the object moved with it, providing, of course, that the girder was rigid and the aluminum cap stayed in place on the eye (see Barlow, 1963, for additional details, or Yarbus, 1967, for a complete treatment of his method and findings). Yarbus's method, unlike the contact lens optical lever, does not allow convenient control of properties of test stimuli, but it may have at least one distinct advantage, namely, it stabilizes the stimulus far better, at least according to Barlow. Unfortunately, Barlow's conclusion rests on experiments in which he confounded characteristics of these quite different methods – a problem (explained below) that Riggs and Schick (1968) would subsequently suggest contributed importantly to Barlow's conclusion. Barlow's paper, however, includes a number of important observations about the appearance and possible physiological underpinnings of stabilized images as well as a number of criticisms of Yarbus's method. For these reasons Barlow's paper will be described in some detail despite the fact that his rejection of the scleral contact lens does not apply to its use in an optical lever.

Barlow prepared scleral contact lenses of "the type ordinarily used for stabilized image work, with a close fit at the limbus." He then painted an artificial pupil on the perspex (acrylic) contact lens, and molded the corneal portion so as to form a 40 D supplementary lens. A short aluminum tube, which was cemented around the pupil, was extended on one side to form a girder used to hold a target in focus 25 mm in front of the eye. These additions to the scleral lens, particularly the aluminum tube and girder with a target at its end, seem certain to apply much larger, potentially dislodging, forces on the contact lens than a beam of light falling on a small mirror at the fulcrum of an optical lever – the method introduced by Riggs and Ditchburn. Barlow's modifications of the Yarbus sucker cap were more modest, but important nevertheless. The basic cap was similar to Yarbus's. It was made of aluminum and had a rubber sucker glued over a hole pierced near its narrow anterior end. The cap carried a girder which was used to mount targets 15 – 25 mm in front of the eye. The main differences were the kind of short-focus lens (Barlow's were planoconvex with the flat side towards the target) and the fact that the cap was filled with a 1.5% solution of $NaHCO_3$ rather than with air, the filler used by Yarbus. The fluid filling made Barlow's caps heavier than Yarbus's (350 rather than 100 mg) but their optical quality was reported to be much better than could be achieved by following Yarbus's method.

An afterimage technique was used to examine the effectiveness of the two types of stabilizing method, first qualitatively, then quantitatively. The quality of stabilization was examined by making a persistent afterimage in the shape of an arrowhead by presenting a brief intense light in Maxwellian view through an aperture mounted on the contact lens or sucker cap. An occluder which had covered a portion of the visual field when the afterimage was made was then removed, exposing a transilluminated aperture in the shape of an arrowhead with its point facing and lined up with the point of the afterimage. If the contact lens or sucker cap did not move relative to the eye, the two points of the arrowheads should stay in alignment despite movements of the eye. Barlow found that relative movements were "unmistakenly detected" during attempts at maintained fixation, as well as when the subject looked around, when the contact lens was used to stabilize the transilluminated arrowhead. Relative movements were much less with the sucker cap and when they occurred "they are most often caused by the rim. . . touching the eye-lid." During large eye movements with the sucker cap, relative motion was observed, but at the end of each large eye movement the original relationship was restored.

Quantitative estimates of the effectiveness of stabilization were obtained by making two afterimages of a target, separated by several seconds, under conditions in which a failure of stabilization would produce a displacement of a feature in the composite afterimage whose extent could be measured psychophysically. Barlow accomplished this by mounting a target with a long straight edge on one or the other stabilizing device and first produced an afterimage of one half of the edge, waited 4 or 5 seconds, and then produced an afterimage of the other half of the edge. If the stabilizing device did not change position on the eye during the interval, an afterimage with a single continuous straight edge would be seen. If, however, the device changed position on the eye in the interval, the second afterimage would be offset from the first and a single long straight edge would not be seen. The size of the offset of the edge in the composite afterimage would indicate the extent to which the stabilizing device had slipped during the 4- or 5-s interval. A rotatable polaroid was used to deliver light to each half of the straight edge independently without disturbing the location of either the target or its stabilizing device.

When both afterimages had been made, the subject removed the stabilizing device from his eye and adjusted the projected image of a step of variable size to be the same size as the offset seen in his composite afterimage. The subject viewed the projected variable-step display and his composite afterimage simultaneously on a screen located at about arm's length. Using this technique, the estimated r.m.s. deviation of the contact lens relative to the eye was more than 3 minutes of arc. The sucker cap was much better. Its r.m.s. deviation was well under 1 minute of arc. Barlow concluded that "the full-fitting type of contact lens does not stabilize the image adequately" and proceeded to use his version of the Yarbus sucker cap to study the appearance of stabilized images. He also discussed a number of potentially important artifacts associated with this device. This discussion of artifacts will be summarized next because they bear on Yarbus's prior work and also because they provided a useful guide for subsequent investigators who used a Yarbus-type sucker cap (most notably Gerrits and his collaborators).

6.2.5. Artifacts in Yarbus's method

Before considering Barlow's concerns with potential artifacts in Yarbus's method, it is worth noting that Barlow reported that his version of the Yarbus sucker cap was not perfectly fixed upon the eye. Its variability (r.m.s. deviation) during 4- or 5-s intervals was somewhat more than a half minute of arc. Barlow's sucker cap performed much better than his scleral contact lens carrying the same loads, but the sucker cap only reduced retinal image motion. It did not eliminate retinal image motion entirely. How good does stabilization have to be before it is good enough to guarantee that functionally effective transient stimulation, sufficient to drive visual neurons, has been eliminated? Barlow did not deal with this question explicitly. His discussion, however, assumes that his version of Yarbus's method, unlike the method of Riggs and Ditchburn, would be good enough. Later, we will describe Arend and Timberlake's (1986) recent theoretical calculations, which show that Barlow's assumption, made by other contemporary and subsequent investigators as well, was unwarranted.

Barlow's other concerns, concerns which encouraged him to modify Yarbus's method, stem primarily from optical considerations. The inner surface of the supplementary lens mounted at the end of a sucker cap "steams up very rapidly unless the lens is first warmed, and this only delays it for about ½ minute." Also, the suction used to hold the cap to the eye changes the shape of the cornea. Changes in the shape of the cornea change the focus of the eye. If the target is in focus when the sucker is first attached, it will be out of focus somewhat later. Barlow eliminated both problems by filling his sucker cap with a solution of bicarbonate. Chromatic aberration can also be a problem with a sucker cap. It is is not significantly greater with a Riggs- or Ditchburn-type of scleral lens optical lever than with the unencumbered eye because this method does not require a short-focus supplementary lens. Chromatic aberration becomes a problem, however, when a short-focus lens is added to the optics of the eye, an eye which already has some chromatic aberration, because "its own aberration will be added to that of the eye, and is liable to become the limiting factor [in resolving details]. The additional error is greater the shorter the focal length, which is why lenses at ca. 20 mm have been used rather than 5–10 mm as favoured by Yarbus" (p. 42). Barlow used a zinc-crown glass planoconvex supplementary lens with its curved surface immersed in $NaHCO_3$ to reduce the chromatic aberration introduced by the positive surfaces of the glass. Yarbus sometimes used pinhole apertures to reduce aberrations, but Barlow points out that this technique severely compromises image quality because of diffraction effects.

There are other potential artifacts which cannot be easily circumvented with the Yarbus sucker cap. Barlow mentions trans-scleral light, which could reduce the contrast in the stabilized image. Trans-scleral light can be reduced with a scleral contact lens simply by painting the scleral portion with an opaque substance such as several coats of black acrylic paint. With a sucker cap the sclera is completely exposed and trans-scleral light could reduce contrast in the retinal image, depending on its intensity and the optical arrangements used to illuminate the target. Reducing contrast could have important consequences for the time-course and ultimate appearance of a stabilized image. It is probably worth noting, before going on to report Barlow's description of the appearance of stabilized images free from misting and serious optical aberrations, that he procured these improvements by adding appreciably to the forces applied to the sucker cap. His fluid-filled cap was 3 times heavier and the lever used to support the target was 2–4 times longer than Yarbus's.

6.2.6. The subjective appearance of stabilized images

Barlow, having convinced himself that previous reports were suspect because of either poor stabilization or poor optical quality and that he had reduced or eliminated both problems, undertook to provide a description of the appearance of stabilized images. His description is somewhat at variance with previous reports. In his words: "When one inspects an image of good contrast and optical quality, moderate retinal illumination (say 1 to 100 Trolands) and as well stabilized as we are able to achieve, it is seen with full clarity only for the first few seconds, at the end of which time it lacks some of the fine detail and contrast of the original. There is then a period of a minute or so during which its appearance fluctuates, disappearing and regenerating in a way that will be described later. Finally these fluctuations die out, leaving a stable appearance or a fog or grey sky with ill-defined dark and light clouds in it corresponding to the white and black parts of the original image. This final state, a very blurred, very low contrast version of the original image, seems to persist without fluctuation for as long as conditions are held unchanged. On the occasions when it has disappeared, the cornea has been found to be misted, or the lens smeared, on removing the contact lens [sucker cap]" (p. 43).

"In summary then, stabilized images both fade and regenerate, but they do not fade completely, for after several minutes a cloudy, low contrast form of the original persists, and they do not regenerate completely, for there is an initial 'blurring' or loss of detail and contrast that is never regained" (p. 45).

The main differences with previous reports were his failure to observe complete disappearance of all features of the stimulus, the persistent blurring of sharp details, which did not sharpen when the stimulus regenerated, and regeneration itself. His observations were most at variance with those of Yarbus, who had described complete disappearance without reappearance. The main differences with Riggs's and Ditchburn's groups was their observation of the reappearance of fine details and high contrast following regeneration. Barlow felt that these differences called attention to shortcomings in their scleral contact lens method. Barlow was able to repeat his observations after homatropinizing the eye, which showed that his observation of blurring and regeneration was not caused by fluctuations in accommodation of the crystalline lens. The differences with Yarbus are ascribed to the poor optical quality of Yarbus's stimulus, with some emphasis on trans-scleral light, both of which could reduce the contrast of Yarbus's displays. Barlow believed that Yarbus's failure to observe regeneration and to achieve complete disappearance resulted from the poor quality of Yarbus's stabilized retinal images. Barlow next has to explain his failure to satisfy his own criterion for achieving adequate stabilization, namely, the failure to observe complete disappearance, which he initially described as a characteristic of all entoptic phenomena. These intrinsically stabilized images disappear rapidly and completely (see section 6.2.4). He ascribes the failure to meet his initial criterion to the fact that the details in the entoptic percepts are very small, of low contrast and in some instances arise outside the central fovea – in many ways the same kind of reasons he uses to explain Yarbus's result which is congruent with the expectation from observations of entoptic phenomena.

Barlow next offers neurophysiological speculations for a number of his observations. He suggests a 'diffusion process' in the retina of the kind proposed by Brindley (1962) to account for the blurring and loss of detail observed in the stabilized image. A 'two-channel hypothesis', that is, 'a dual system of fibres connecting eye and brain' is proposed to explain other observations. The first is a rapidly adapting [transient] channel which signals changes in illumination in small retinal regions relative to the average signal arising from larger regions. "These units would signal an approximation to the spatio-temporal derivative of the light, $-\delta(\delta^2 i/\delta x^2+\delta^2 i/\delta y^2)/\delta t$, and would have properties similar to the on-centre and off-centre units of the cat's retina. . . The second channel is a slow- or non-adapting channel [sustained] which would serve as the mechanism subserving the 'dim clouds' which appear after prolonged stimulation, and are correctly related to real luminance of parts of the visual field" (p. 49).

6.2.7. Krauskopf's ring-disk experiment

In the same year that Barlow published his paper comparing the Yarbus and Riggs-Ditchburn stabilization methods, Krauskopf (1963) reported an experiment in which he showed that stimulation by a moving edge was critical for maintaining perception of the color within the region circumscribed by the edge. Krauskopf traced his interest in this problem to observations by Liebmann (1927), who had reported that equally bright patches of different color were unstable when viewed for prolonged periods, and to Ditchburn and his coworkers (Ditchburn, 1957; McCree, 1960; Clowes, 1962), who had reported that stabilization had different effects on lights taken from various parts of the visible spectrum. Krauskopf did his experiment by using a Lummer-Brodhun cube from a *Macbeth Illuminometer* to produce a ring of one color surrounding a disk of another color. Four relatively narrow-band chromatic stimuli, selected from the long, middle and short wavelength portions of the visible spectrum, were employed. A Riggs-type optical lever stabilized the light passing through the Lummer-Brodhun cube. A field stop (aperture), which was placed near the outer edge of the approximately 2 degree stabilized test field, destabilized the outer edge of the ring surrounding the disk. Neutral density wedges were used to match the luminance of the chromatic stimuli under normal fixation conditions, that is, conditions in which the ring and disk were not stabilized. All possible combinations of the four wavelengths as rings and disks were used. Each ring-disk combination was presented for 30 seconds, the subject pressing a key whenever he saw the disk and releasing the key whenever the disk disappeared. A disappearance-time fraction, based on these key-presses, provided the dependent variable Krauskopf used to evaluate reports that effects of stabilization varied with spectral locus.

Krauskopf "found that when the image is stabilized, the inner boundary which is stabilized disappears, the central disk taking on the color of the annulus. Thus in a case in which the observer saw initially a red disk on a green annulus he reported that after disappearance he saw a large green disk. . . Another mode of appearance was reported when the disk luminance was greater than that of the annulus. Under this condition the observer reported that the border between the disk and the annulus became uncertain and irregular in shape. The simplest description of the result is that the central color seemed to spill out through the border and invade the annulus" (p. 742). His analyses of the disappearance-time fraction data did not permit any simple generalizations about effects of stabilization on combinations of colored lights selected from different parts of the visible spectrum.

6.2.8. Gerrits's elaboration of Yarbus's method

Gerrits, De Haan and Vendrik (1966) introduced important refinements of the stabilization technique developed by Yarbus. They eliminated two of Barlow's criticisms of Yarbus's method, namely, fluctuations of accommodation and the accumulation of water on the inner surface of the air-spaced compensatory lens which Yarbus mounted on his sucker cap. Fluctuating accommodation would allow the target to come in and out of focus, making it less and then more likely to disappear when stabilized. Reports of fluctuating target appearances could be caused by fluctuations in accommodation rather than from more central visual processes associated with stabilization. Similarly, water condensed on the inner surface of the high-power compensatory lens would reduce the contrast of the test stimulus, making it more likely to disappear rapidly when stabilized and to remain invisible throughout a subsequent long period of stabilization. In short, condensed water could lead to reports of complete and persistent disappearance, which would not be observed if water had not accumulated during the experimental run. The first problem was eliminated, as Barlow had suggested, by instilling a mydriatic drug into the eye. This has the effect of dilating the pupil as well as almost eliminating changes in accommodation of the crystalline lens. The second problem was eliminated by flowing warm water, which kept the compensatory lens near body temperature, through a tube wrapped around the cone of the sucker cap.

Gerrits et al. (1966) also made a number of other elaborations of the Yarbus-Barlow technique. For example, they used a threaded tube, rather than an aluminum lever, to mount their target objects on the sucker cap. The threaded tube allowed the observer to make fine adjustments in the focus of the test object so as to render its image subjectively sharp. They also used ingenious hydraulic drivers to impose linear unidirectional and random motions on a test object mounted on the sucker cap. In subsequent research Gerrits and Vendrik (1970) developed a miniature synchronous electric motor, which was mounted on the sucker cap and used to impose rotational and 'jumping' motions. They also began to take care to eliminate trans-scleral light, which reduces target contrast, in line with another of Barlow's (1963) suggestions (see Fig. 10 for diagrams of this remarkable instrumentation).

As the reader might suspect, the instrumentation illustrated in Fig. 10 was heavy (almost 10 g by 1970) and Gerrits and his coworkers found it necessary to make observations only while in a supine position; these devices would move around and even fall off if the observer sat up. Inconvenience and unsuitability for a wider variety of visual stimuli eventually led to the retirement of the elaborate instruments diagrammed in Fig. 10. Gerrits and Vendrik (1972, 1974) devised a replacement by mounting the end of a fiber optic bundle on their version of the Yarbus-type sucker cap. The bundle, which provided a visual field subtending about 13 degrees of visual angle, contained 160000 fibers, each fiber subtending about 1.8 minutes of arc (about 5 times the diameter of a cone at the foveal center). Individual fibers had been dissected free of the relatively stiff protective outer plastic sheath for a distance of several centimeters so as to permit a reasonable range of eye movements when the dissected end of the fiber optic bundle, which was cemented into a metal band, was inserted into a tube mounted at the end of the sucker cap. The other end of the fiber optic bundle 'looked at' the face of a TV display upon which a variety of stationary or moving test stimuli could be presented. The weight and stiffness of the dissected end of the fiber optic bundle was such as to require the observer to maintain a supine position. It is important to note that in all the research with the fiber optic bundle the diameter of the individual fibers "is larger than the visual resolution of the cone system (fovea). Therefore the moving contours of the stimuli used had to be located out of the fovea, where individual fibers were no longer observable" (Gerrits and Vendrik, 1974, p. 176). Stabilization of targets confined to the foveal floor, or even more importantly to the 20 min arc foveal bouquet where the retinal mosaic is at its finest, has never to our knowledge been accomplished with the Yarbus-type methodology, the methodology believed by some to be capable of the best possible stabilization, at least since Barlow's report. More will be said about this later in section 6.2.10.

6.2.9. Main observations of Gerrits and his coworkers

They reported that "once the perception of a stabilized image has disappeared it does not come back as long as the subject is capable of preventing large rapid eye movements. Trained subjects can achieve an uninterrupted absence of perception for at least ten minutes" (Gerrits et al., 1966, p. 434). In

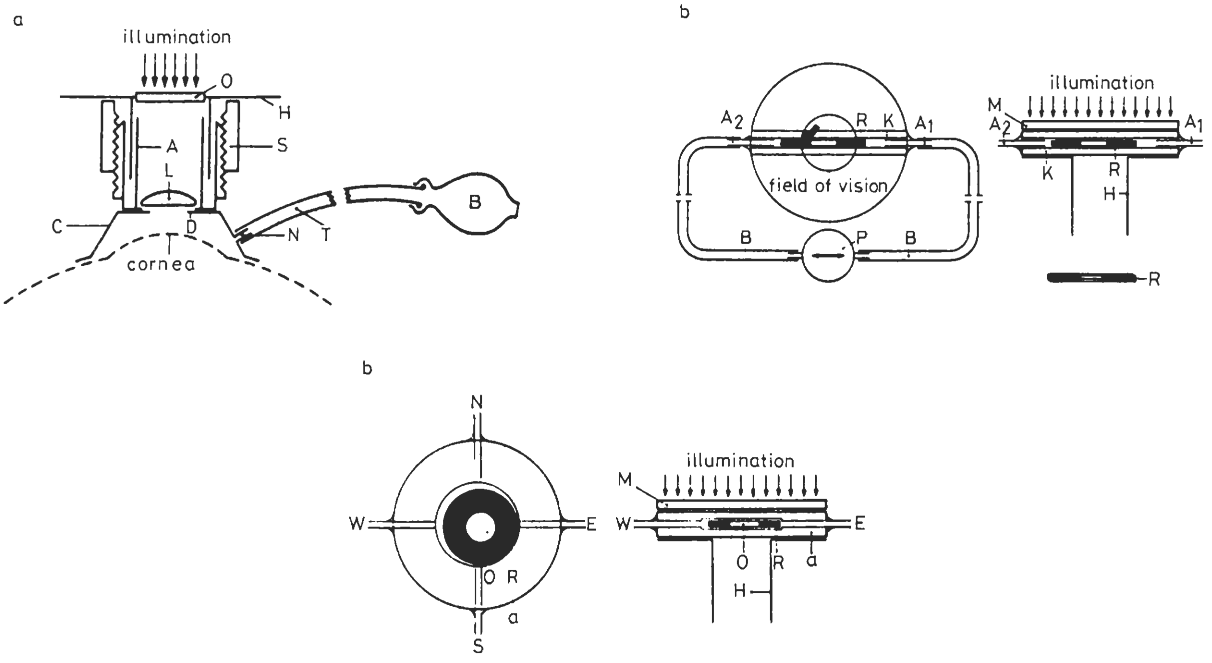

Fig. 10. (a) A standard cap sucked onto the eye (not drawn to scale). The object holder, H, with the object, O, can be changed during the experiment. The weight of the cap is 2.2 g and the object in its holder adds another 0.5 g. A is a non-reflecting black tube, B is a bulb, C is a dural cone, D is a diaphragm, L is a 50 diopter lens, N is a connector, S is a screw and T is a flexible tube. (b) System used to obtain movement in one direction of a stabilized object. The weight of the object in its holder is 2.1 g. A_1, A_2 are tubing connectors, B are flexible tubes, K is an acrylic (perspex) channel, M is an opaline plate, P is a volume pump and R is an acrylic (perspex) rod. (c) System used to obtain random movements in all directions of an otherwise stabilized object. The weight of the object in its holder was 2.7 g. N, E, W and S are four inlets. R is a black annulus, O is a hole, M is an opaline plate and G is an acrylic (perspex) leader. (From Gerrits et al., 1966)

their next experiment with the synchronous electric motor mounted on the sucker cap (Gerrits and Vendrik, 1970), they studied the effect of motion imposed on a stabilized target. They operated the motor in ways that they believed imitated "the drift, the saccadic and the tremor movements of the eye, and studied their influence on perception. It was found that drift-imitating movements regenerate (fill in) a disappeared object. The movements imitating saccades and tremor were never effective in restoring vision" (p. 1455). These conclusions seem reasonable for drifts and tremor, but are less clear with respect to a role for saccades. Gerrits and Vendrik (1974) made quantitative estimates of the drift motions optimal for preserving perception in a subsequent experiment in which they used their fiber optic stimulating bundle. They found that mean drift-like image speeds of about 22 minutes of arc per second were optimal for preserving the perception of a stabilized object. They duly noted that this speed, "which is very effective in preserving perception, is higher than the mean speed of [natural] eye drifts (1–6′/sec)" (p. 178). Remember that Riggs et al. (1953) also found that doubling (in their case) 'natural' retinal image motion was more effective than natural image motion in preserving perception (see section 6.2.1). We will encounter this mysterious fact again before its significance is appreciated.

By 1978 Gerrits found himself at variance with other investigators with respect to the effect of such stimulus properties as the order of fading in the peripheral and foveal retina, the importance of boundaries in the stabilizing field, the number of cycles and duration of tests when the target was a spatial frequency grating. His discussion of this problem is useful because it calls attention to persistent problems in stabilized image research – problems that are still unresolved. He pointed out that discrepancies between stabilized image researchers with respect to details of stimulus properties are often found in experiments in which the investigators also fail to obtain fading for long periods of time. He fears that such "discrepancies can be understood by an incomplete, insufficient stabilization (Barlow, 1963). One should be very careful before attributing effects to [targets stabilized on] foveal cells: an extremely high degree of stabilization is demanded to prevent conclusions to be drawn from small artefacts some of which prove to be very effective in the foveal area but remain invisible in the periphery. The difference in movement sensitivity [between the fovea and the periphery] is even a common observation in non-stabilized conditions. It is very easy to limit voluntarily one's eye movements, the larger amplitudes can be suppressed. As a result the image fades in the periphery, in the areas containing the largest elements [receptive fields] (Troxler's effect). After a short training smaller and smaller amplitudes of eye movements can be suppressed too, also by naive subjects (Winterson and Collewijn, 1976), even down to the microsaccades (Steinman et al., 1967, 1973). This results in further shrinkage in the visual field, down to the fovea. An image with contours in the foveal area never fades by voluntary effort because it is impossible to suppress the remaining miniature eye movements, e.g., the drifts. The influence of these, particularly effective in the foveal area, can only be fully cancelled by excellent stabilization. However, when no stimulus contours are present in the foveal area, e.g., in the case of a large stimulus covering the fovea, even the foveal (filled-in) brightness percept can be made to fade by voluntary effort. The percept is restored by a saccade of large amplitude" (pp. 239–240).

Gerrits, Stassen and van Erning's (1984) recent extension and summary of almost 20 years of research on stabilized images in Nijmegen concluded that "drifts are capable of preserving the perception of stimuli in and around the fovea but not stimuli with contours far outside the fovea". It is important to realize, however, that their conclusion of the necessity for saccades with large stimulus patterns, whose edges fall far away from the fovea, is based on observations made while the observer was lying as still as possible with his head supported artificially. We shall see later that this methodologically imposed constraint, immobility of everything except the eyeball, in stabilized image research obscured important properties of normal visual processing until very recently.

Before we close this section on the Nijmegen School of Stabilization, we feel that it is important for the reader to realize that no instrumentation so far devised has freed the stabilized image investigator from making a great many subjective decisions about the merit and purity of particular observations. Much of the stabilized image research literature rests on the finesse and skill of the observer, almost always one of the investigators. In many instances the observation of stabilized images developed into a highly stylized performing art. This was particularly true when subjective reports, rather than thresholds, provided the primary or only basis for discussion and theory (our nervousness about the role of artistry in visual science was discussed in section 6.1.1). Gerrits and his coworkers were aware of the highly subjective nature of many of their reports and went out of their way to give the reader indications of the kind of observational skills and judgement upon which their reports were based. Consider, for example, the following quotation from the method section of Gerrits (1978) in which he describes the use of the fiber optic stimulator:

"Most of the results to be described have been obtained from two highly trained subjects; one investigated these effects for 2 years, the other for 4 years. A number of other subjects participated occasionally in the experiments. When the cap is sucked on the eye of the subject, the image of the object does not, in most cases, fade within a few seconds after the light is switched on. The subject starts to bring the image in focus and thereafter looks for the most comfortable yet satisfactory position of his head as well as his eyes in his head. He changes his line of sight until no on- and off-borders (caused by small shifts of the cap over the eye) are generated any more and the image fades easily. Small shifts causing destabilization will occur if the object holder touches the subject's nose or eyebrow, if the small rubber tubes supplying warm water and the underpressure [suction] exert a pulling force on the cap and, particularly, when an unquiet subject moves his eyes too much. Just by looking to the relative position of the on- and off-borders a trained observer is able to correct a wrong direction within a few minutes and to keep this most comfortable" (Gerrits, 1978, p. 227).

The instrumentation in Nijmegen never progressed to the point where less skill and training was required. To illustrate, Gerrits et al. (1984) report that "The stiff protecting mantle of the fiber bundle was removed over a length of 30 cm in order to allow the bundle to follow the subject's eye movements. These eye movements did not affect the position of the image on the subject's retina as long as the suction cap adhered well to his eye. The optic fiber bundle had, however, a small braking effect and therefore could cause slippage.

"To enable the subject to distinguish between a percept generated by a genuine stimulus movement or by an unwanted slippage of the suction cap, a small black disc at the end of a non-moving stalk was placed in front of the color TV display. This disc functioned as a control spot relative to the moving square [the stabilized test target upon which movement would be imposed] and enabled the center of the square to be projected onto the fovea before the onset of the movement of the square" (pp. 447–448).

6.2.10. In defense of the contact lens

Barlow's (1963) paper on stabilization methodology did not force the contact lens into retirement despite the impetus it provided for Gerrits and his coworkers to undertake a long line of research based on the sucker cap method introduced by Yarbus. Barlow's comparison of the stability of the contact lens and the sucker cap was fair only when both of these devices were used to hang a relatively heavy optical stimulator at the end of a lever which extended a centimeter or two in front of the eye. Contact lenses were first used in this way by Ditchburn and Pritchard (1956). Their method had begun to gain popularity because of its relative simplicity when Barlow undertook his study. Barlow did not test Riggs's optical lever stabilization technique, in which only a lightweight mirror is mounted on or within the surface of the scleral contact lens. Free-standing optical elements provide the compensating path required for image stabilization in the Riggs-type apparatus. These arrangements meant that the forces applied to the contact lenses were very small as compared to the forces applied to the contact lens in the Ditchburn-Pritchard-type apparatus tested by Barlow. Riggs and Schick (1968) pointed out this limitation in Barlow's experiment when they borrowed his afterimage method to measure the stability of a scleral contact lens when it is used at the fulcrum of an optical, rather than mechanical, lever (see section 6.2.5).

They modified Keesey's (1960) visual acuity apparatus (see section 6.2.3) to allow vernier-offset measurements between an afterimage and a stabilized image. The configuration of these two images was arranged so as to provide a vernier acuity test target in which vernier acuity could be used to estimate changes in the position of the contact lens. The vernier acuity test target consisted of a dark rectangular afterimage, whose long side was vertical, and a stabilized dark bar of the same size and shape. The afterimage was made by viewing a strobe-flash through a bar-shaped aperture while the aperture was in vernier alignment with the stabilized dark bar. The horizontal alignment of the stabilized dark bar with respect to the afterimage could be adjusted by the subject so as to eliminate any observed offset in their vertical alignment. Data were considered to have merit during trials in which the dark negative afterimage was sharp and the vernier alignment was good at the beginning of the experimental run (the aperture and dark bar were aligned by the subject just before he fired the strobe). Experimental runs lasted as long as the afterimage remained clear (20 to 80 s after the strobe-flash). The afterimage was kept visible by flickering the background field illumination about once each second. If the contact lens moved with respect to the eye during the experimental run, the adjustable dark bar, which was stabilized by reflection from the contact lens mirror, would change its vernier-offset relative to the afterimage. The subject's task was to adjust the offset, whenever necessary, so as to bring the configuration back into vernier alignment.

Control measurements, made with a real unstabilized bar substituted for the afterimage, showed that the vernier tracking error (standard deviations of offset corrections for offsets introduced by the experimenter rather than by contact lens slippage) was less than 11 seconds of visual angle for each of the three subjects. This level of precision was clearly adequate for the proposed measurements of contact lens stability, which, based on Barlow's report (1963), should be 5 to 16 times larger (see section 6.2.4). Riggs and Schick's (1968) main measurements consisted of offset-error position corrections of the stabilized bar relative to the afterimage. These measurements were made while the subject attempted to minimize large eye movements – the mode of viewing subjects would employ during a typical experiment on stabilized images (see section 6.2.8). Data were also collected when the two more experienced subjects made saccades of known magnitudes (up to 6 degrees), the subject making alignment adjustments before and after each saccade. In this experiment the stability of the contact lens was inferred by calculating the differences between measurements made before and after saccades. Measurements were also made over extended periods of time by substituting an unstabilized bar for the afterimage and tracking the vernier relationship between the stabilized and unstabilized dark bars for a half hour with offset corrections made every 3 minutes.

Each subject made about 50 tracking records and the median standard deviations of image displacements were about 23 and 25 seconds of arc for the two more experienced subjects and about 35 seconds of arc for the less experienced subject. The average shifts in image displacement were slightly smaller. The shifts in image displacement associated with voluntary saccades were also less than a minute of arc. The long-term drift of the position of the contact lens was slow. Its speed ranged from about 6 to 15 seconds of visual angle/minute. Riggs and Schick (1968) suggested that the continuance of slow drift of the contact lens over long periods of time can be "explained in part by the fact that the eyeball changes shape over extended periods of wearing a tightly fitting lens". The authors go on to discuss the "extent to which numerous earlier studies of eye movements and stabilization may have been affected by errors of the magnitude reported" when they, conservatively, take the magnitude of contact lens slippage to be about half a minute of arc for relatively short-term stabilization. They conclude that error of this magnitude "is not sufficiently large to be of much significance in work with stabilized images" (p. 165). Their conclusion rests primarily on the report of Krauskopf (1957; see section 6.2.2) and Riggs et al. (1961), who reported that motions of the retinal image smaller than 1 minute of arc are not sufficient either to prevent disappearance of a stabilized target or to cause regeneration of a stabilized target which has disappeared. They also cite work by Ditchburn and his coworkers in support of the conclusion that retinal image motions smaller than one minute of visual angle have no consequences for vision.

Riggs and Schick include important comments on the expected appearance of entoptic phenomena in their paper. They note that complete disappearance of entoptic images, whose images are intrinsically stabilized, does not provide convincing evidence that failure of complete disappearance and periodic re-appearances of mechanically (extrinsically) stabilized stimuli does not necessarily result from imperfect stabilization. The differences associated with the fate of images stabilized with intrinsic and extrinsic techniques can be explained by other facts, most notably the poor focus, low contrast and extrafoveal location of most entoptic images. All of these factors would facilitate rapid and persistent disappearance of entoptic images. Such effects would not be expected when well-focussed, high-contrast targets are presented at the center of the fovea with an extrinsic stabilization technique (Barlow, 1963, had already called attention to such considerations).

So, having first shown that Barlow's (1963) rejection of the contact lens for stabilization was restricted to a special case in which large inertial forces were applied to the contact lens, Riggs and Schick (1968) went on to conclude, on the basis of prior contact lens optical lever research, that stabilization good to only one minute of arc is sufficient to study the effect of target motion on visual processing. This conclusion left open the possibility that complete, persistent disappearance of an intense, stabilized foveal target with good optical properties may not be characteristic of visual system performance (Barlow's, 1963, conclusion). Riggs and Schick (1968) did not comment either on Barlow's reported failure to achieve complete disappearance of a stabilized image which met these criteria, or on the contradictory report of Gerrits et al. (1966) in which complete, persistent disappearance was, once again, suggested to be the sine qua non of effective stabilization – a claim that Gerrits and his coworkers would continue to make in subsequent experiments published during the next two decades. The final fate of an adequately stabilized, intense, well-focussed, high-contrast display presented to the central fovea remained uncertain as the stabilization technique entered the 1970s – the third decade of research with stabilized images. At this time emphasis shifted away from tests of stabilization techniques towards the development of rather elaborate quantitative theory, which attempted to relate eye movements to basic visual processing. A step in this direction had been made ten years earlier by Bryngdahl in 1961.

7. Fourier Optics and the role of eye movements in spatial vision

We have already referred to the introduction of what we called 'Fourier Optics' into the analysis of visual system function shortly after techniques were developed to do research with stabilized images (see section 6). By way of reminder, Schade (1956) and de Lange (1957, 1958) were responsible for initiating this initial interest; the former providing an example of how linear systems analysis might be used to describe spatial factors in vision, the latter providing experimental evidence of the power of these techniques, as well as an example of how they might be used to study the operation of temporal factors in the human visual system. Their approach was immediately taken up by others and by 1959 Levinson had used Fourier Optics to show that in a "flicker-fusion experiment there is more to flicker than meets the eye" (p. 919). Levinson supported his timely dictum by Fourier-analysing the harmonic content of a complex pulse-sequence of flashes and thereby explaining puzzling aspects of Brown and Forsyth's (1959) experiment with flickering lights. Levinson's analysis of the Brown and Forsyth data led him to suggest that at fusion threshold all but one of the Fourier components of the flickering stimulus (the fundamental) were below threshold – a conclusion consistent with the treatment of flicker introduced by de Lange in his doctoral dissertation. Within a year, however, Levinson (1960) had found a flicker waveform (the addition of a near-threshold second harmonic of suitable phase) which showed that flicker threshold did not always depend exclusively upon a single fundamental Fourier component. Fourier Optics was proving to be a powerful tool for the study of temporal factors in vision, so much so that, within the decade, Levinson (1968) had developed and tested a multistage linear low-pass filter model of the response of the visual system to flickering lights. The application of Fourier Optics to temporal factors in vision had considerable merit. Could it also be used to analyse and model the spatio-temporal variations inherent in visual stimulation now that we knew that the eye was in continual motion despite all efforts to maintain fixation?

7.1. Bryngdahl's linear filter model of eye movement and visual acuity

Bryngdahl's (1961) paper was entirely theoretical. He described the impetus for his work as "recent developments of eye-movement recording techniques [provide] a way for an examination of the information channel between the eye and brain. Questions in this field can be treated by either information theory or linear filter theory" (p. 1). Bryngdahl chose the latter.* He built his theoretical treatment on the then-new understanding of the fine-grain characteristics of the fixational eye movement pattern and on recent demonstrations of the effect of artificial motion imposed on stabilized images. Both had been worked out by Riggs and Ditchburn and their coworkers during the preceding decade (see sections 6.2.1 to 6.2.3). Bryngdahl also leaned heavily on de Lange's (1957, 1958) treatment of the perception of flickering lights. Bryngdahl approached eye movements at the level of single cones as events designed to cause variations in the intensity of the stimulating light – variations essential to maintain the visibility of acuity test targets and to enhance the contrast of differences in the light distributions of acuity targets. He began by pointing out that "the visual system works logarithmically for large variations at low frequencies [of sinusoidally time-modulated illumination]. . . [but] for limited variations *the system appears to work linearly* [his italics]. The frequency characteristic for cone seeing (attenuation characteristic) shows a filter action for intermittent light [after de Lange]. The response function has a maximum at about 10 c.p.s. [Hz]. This opens a way to determine the constants of the visual transfer [function] and to explain a correlation between eye movements and CT-curves" [contrast sensitivity functions] (pp. 3–4).

A detailed treatment of Bryngdahl's model would occupy much more space than the subsequent success of the model would justify. The model is highly speculative, resting in large measure on details of the transduction process about which Bryngdahl freely admits "almost nothing is known" (p. 9). Nevertheless, the model does propose an eye movement-based contrast enhancement mechanism. In subsequent work Gilbert and Fender (1969) claimed that Bryngdahl's (1961) paper "has shown that the prediction [of such a contrast enhancement mechanism] was quantitatively plausible" (p. 192). We are not convinced that Bryngdahl actually succeeded in doing this and suggest that the interested reader should study Bryngdahl's model himself. It is sufficient, here, to appreciate that his treatment was consistent with a number of features that had already been observed in the eye movement and stabilized image research. In this he sets the stage for numerous subsequent efforts to correlate eye movement features with contrast detection and visual acuity.

The main relationships that he inferred from his calculations, which were consistent with the empirical work of others, were as follows: (1) "flicks [saccades] are capable of supporting normal vision", (2) "the amplitude of high frequency motion had to exceed 1 min arc in order to have a demonstrable effect on the contrast threshold" [i.e., only very rare large-amplitude components of physiological nystagmus would have visual consequences], and (3) because "the critical resonance frequency is

---

* Information theory (Shannon, 1948) has not figured prominently in the development of modern visual science, particularly when compared with the enthusiastic adoption of Fourier Optics by the visual science community. Information theory had a brief period of influence in higher-order processes such as form perception (e.g., Attneave, 1957), but very little influence on theories of more basic visual processing. Ditchburn and Drysdale (1973a,b) provide an exception in that they used information theory in their analysis of visual information obtained from flashes and from afterimages. St.-Cyr and Fender (1969) provide an exception in the eye movement literature. These authors attempted to explain very short phase-lags observed during smooth pursuit of predictable, relative to pseudorandom, target motions from an information theoretical approach. Kowler and Steinman (1979) and Kowler et al. (1984a) showed that there were historical as well as empirical problems with St.-Cyr and Fender's ill-conceived attempt.

known to be about 10 c.p.s. [Hz] (the eye movements try to make this signal as large as possible)... this frequency is transduced by the *drift* [his italics] for 1 sec arc targets" [i.e., acuity for a minimum visible dark bar depends on slow drifts] (p. 13). Another way to think about this third point is to realize that with high spatial frequencies slow drift eye movements could provide flicker in the critical resonance range of about 10 Hz. Bryngdahl's treatment of the correlation between eye movement and contrast sensitivity explicitly reduces the motion of inhomogeneities in the retinal light distribution to a problem of resolving temporal variations produced by motions of this light distribution. This approach simplifies the problem because it eliminates the need to consider target velocity qua velocity – a spatio-temporal interaction. Only temporal factors need be understood. Unfortunately, this simplification would be tested and rejected by subsequent researchers (see section 7.2.4).

7.2. Enhancement of visibility by motion

7.2.1. Van Nes's observations at low spatial frequencies

Van Nes (1968) reported "evidence for enhancement of visibility by regular motion of retinal images" just as Fourier Optics began to dominate the research activities of the visual science community. His evidence was obtained by comparing the contrast sensitivity of a human observer to stationary and to moving sinusoidal spatial frequency gratings. The observer's task was to maintain fixation at the center of a 2.4° × 1.2° television monitor which displayed a sinusoidal grating of a particular spatial frequency and average luminance. The psychophysical contrast threshold for the grating was measured by varying the modulation of the grating from trial to trial. The grating either remained stationary or moved at a constant velocity on a given trial. Trial duration was left to the discretion of the subject, most judgements being made in less than 15 s. The retinal illuminance of the display was varied over a ten-fold range from 8.5 to 850 photopic Trolands. The average light level of the display proved to be an important variable in Van Nes's experiments; the most important results were observed most clearly only at the highest light level. Van Nes's 850 td retinal illuminance was equivalent to about 130 mL at the entrance pupil of the eye – a value more than 3 times higher than the light levels likely to have been used in the experiments in which motion was imposed on stabilized images described in section 6.2. His highest light level was also almost 3 times higher than the light level used in more recent studies of retinal image motion imposed on a stabilized spatial frequency display (Kelly, 1979a,b).

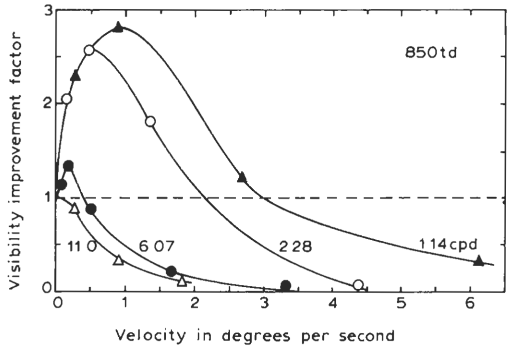

Fig. 11. Visibility improvement factors for four spatial frequencies as a function of the velocity of the grating. The average retinal illuminance was 850 trolands. (From Van Nes, 1968)

Van Nes used the Method of Limits, which he had found to be much more precise than the Method of Adjustment when used to measure psychophysical contrast sensitivity thresholds (Van Nes and Bouman, 1967). Spatial frequencies were varied from 0.64 to 11 cycles/° and constant velocity motions were varied up to a maximum of 13°/s. The main measurements were made with horizontal gratings moving vertically, but similar results were obtained in a smaller set of observations when the gratings were vertical and the motions were in the horizontal direction.

Van Nes found that "for low spatial frequencies and low velocities but rather high retinal illuminances, the grating thresholds for moving patterns were lower than for stationary patterns". He reported this "enhanced visibility as a 'visibility-improvement factor': the ratio of grating-threshold modulation at zero-velocity to grating-threshold modulation at velocity, *v*, for a given spatial frequency" (p. 369). Fig. 11 reproduces Van Nes's graph of the visibility-improvement factor at 850 td as a function of spatial frequency with grating velocity as parameter. In this figure, factors smaller than 1 mean that motion was detrimental and factors greater than 1 mean that retinal image motion was beneficial. Spatial frequencies of about 1 and 2 cycles/° benefited from retinal image motion when velocities were as high as 3°/s. Higher spatial frequencies, about 6 cycles/°, which were just beyond the peak of the human contrast sensitivity function, only benefited very slightly from motion and only when velocity was very low, below 0.5°/s. Visibility-improvement factors could be quite large for the lower spatial frequencies. For example, a 1 cycle/° grating required only a third of the modulation to be visible when it moved at 1°/s compared to when it was stationary, the condition in which retinal image motion was provided exclusively by normal fixational eye movements.

7.2.2. Assumptions about fixational eye movements

There is an important assumption underlying all research of this kind. Namely, it is assumed that the eye does not smoothly pursue when a subject is asked to maintain fixation of some stationary visual reference in the presence of a moving, structured visual background. Reflexive smooth pursuit eye movements, should they occur in this kind of experiment, would *reduce* the velocity of the grating's motion on the retina. If such reflexive smooth pursuit eye movements had been made in the Van Nes experiments, the visibility-improvement factor reported for low spatial frequencies during grating motion would actually have been obtained with lower retinal image velocities than he assumed, as were the visibility decrements caused by the retinal image motion of higher spatial frequency gratings. Without knowing what the eye is actually doing during experiments of this kind, it is difficult to draw firm conclusions about the effects of retinal image motion on contrast sensitivity. Conclusions rest entirely on the observer's subjective impression of the stability of his fixation, that is, he experiences the grating as moving while he experiences his line of sight as fixed with respect to the edges of the rectangular TV display, which in Van Nes's experiments were small enough to be almost entirely confined to his fovea. Such impressions, compelling as they are, could be misleading because the perception of motion of a particular visual stimulus need not be associated with displacements of its retinal image and retinal image motions are not necessarily correlated with perceived motion. For example, the perception of the 'induced motion' of an objectively stationary object is caused by the motion of visual frameworks around the object, and objects show 'position constancy' with respect to the environment when the retinal position of the image of the object is changed when the eyes move (see Wallach, Ch. 6 of this volume, for a discussion of the relationship of eye movement to the perception of motion).

This problem, the unknown contribution of eye movements to the retinal image velocity of a grating target, encouraged the development of instrumentation that would permit experiments of the kind reported by Van Nes, but avoid assumptions about the stability of fixation in the presence of moving grating targets (see sections 9 and 10 and Kelly, 1979a,b, for the direction these developments took.) The development of such instrumentation was not really necessary because moving gratings, particularly moving gratings whose contrast is near threshold, do not stimulate a smooth pursuit reflex, which captures the line of sight, when stationary objects are present in the field of view (Murphy et al., 1975; Kowler et al., 1984; Kowler, Ch. 1 of this volume), but it was not at all obvious, when Van Nes first reported his contrast threshold measurements, that eye movements had not contributed to his results.

To summarize, the Van Nes experiment was important in two ways for the development of our contemporary understanding of the role of eye movement in the detection of contrast. First, it showed that retinal image motion aids the visibility of low spatial frequencies while it hinders the visibility of high spatial frequencies. The former was news. The latter had been appreciated for a long time (see Crook, 1937; Ludvigh, 1948; for influential antecedent experiments). It also called attention to the potential importance of controlling eye movements during experimental procedures of this kind just as Fourier Optics began to be used widely to study spatio-temporal factors in human vision.

7.2.3. The role of eye movement in the 1970s – a necessary nuisance

At this point in time, eye movements were viewed primarily as an obstacle to be overcome rather than as a potential mechanism for enhancing visual acuity – the way eye movements had been viewed during the Marshall-Talbot decade. Most visual scientists no longer viewed eye movements as providing a mechanism for enhancing acuity by 'dithering' the target, but the status of eye movements had improved over what it had been in the pre-dynamic, static period of Selig Hecht. The eye movement recording and image-stabilizing experiments, which had removed the behavioral underpinnings of the Marshall-Talbot theory, had contributed new importance to the functional significance of eye movements for visual processing. It was now widely accepted that normal fixational eye movements (drifts and/or microsaccades) provided the retinal image motion required to preserve normal vision, complete with fine details, by producing transient stimulation – the kind of stimulation that had figured prominently in theories of visual processing ever since 1938 when Hartline reported a preponderance of 'on-off' phasic responses in frog retinal ganglion cells (see section 5.1). Physiological nystagmus was out but larger, lower-frequency eye movements were in.

Despite the generally accepted view that eye movements were necessary for the maintenance of normal vision, it was clear that eye movements could be an obstacle to research on spatio-temporal factors because eye movements could interfere with, or be confounded with, the control of stimulation to the retina. This became increasingly important when the traditional way of minimizing the effects of eye movement in psychophysical experiments by keeping test exposures very short, below 150 ms, were shown to produce undesirable transients, in themselves capable of clouding results as profoundly as uncontrolled eye movements (see, for example, Estevez and Cavonius, 1976).

For reasons such as these, it seemed clear, at least to some investigators, that it might be safest to do Van Nes's kind of experiment without making assumptions about the quality of the observer's fixation while he observed moving gratings. This could be done by imposing motion on a grating stabilized on the observer's retina. Once the display was stabilized, the observer could view it for as long as he wished while he adjusted its contrast to threshold or made judgements about the visibility of the grating as its contrast was varied by the experimenter. Renewed interest in this possibility represented a return to the kind of experiments that Krauskopf had initiated in 1957 and extended to contrast sensitivity in 1962, in which he imposed motion on a stabilized acuity target. In all his work, Krauskopf had stabilized the targets on his retina by means of a contact lens optical lever. This means of stabilization placed constraints on his experimental procedures (see section 6.2.2). Renewed interest in this kind of research took the form of trying to develop techniques to stabilize test targets with respect to the retina without using any attachments to the eye. Stabilization without attachments to the eye would have the obvious advantage of permitting long experimental sessions during which traditional psychophysical techniques could be used to provide very reliable estimates of threshold values. Attachment-free stabilizing instrumentation would also allow the participation of relatively large numbers of subjects because they would not have to wear tightly fitted, impermeable, research contact lenses. Such lenses can be uncomfortable, and can degrade visual acuity when they are worn for more than a half hour because they deprive the cornea of the oxygen it requires and normally obtains from its contact with air (Murphy, 1978, found that before oxygen depletion causes corneal clouding, research contact lenses can have minor beneficial effects on contrast sensitivity in emmetropic, as well as in myopic, observers. See section 6 for a description of early stabilization research, including problems arising from time-limitations inherent in this technique which make it hard to use during threshold experiments).

Attachment-free stabilizing instrumentation eventually achieved a degree of useful refinement. Data on the effect of retinal image motion on contrast sensitivity obtained with such new instrumentation will be described in section 9.3 after a new role for eye movement in theories of visual processing and new knowledge about the 'natural fixation pattern' have been described (here, 'natural' means the fixational eye movement pattern observed when the subject's head is not stabilized artificially). Before moving on, however, it is worth mentioning that Van Nes appreciated the fact that his observations 'were comparable' to Krauskopf's (1957) in that they both showed a beneficial effect of image motion on visibility. Krauskopf had oscillated a bright vertical bar sinusoidally and found a visibility improvement factor of 1.4 when its frequency was 4 Hz and its peak-to-peak amplitude was 12 minutes of arc. This beneficial, sinusoidal image motion would have a peak speed of about 5°/s and an average speed of about 3°/s. Krauskopf reported beneficial effects with oscillations as high as 5 Hz where peak and average speeds would be even higher, about 6 and 4°/s. Oscillation frequency had to exceed 8 Hz before motion became detrimental. Beneficial retinal image speeds such as these will interest us later when we consider 'natural retinal image motion' caused by imperfections inherent in oculomotor compensation for motion of the unrestrained head. Remember, however, that there were important differences between the Van Nes and Krauskopf experiments. Krauskopf's stimulus, a bright bar contained all spatial frequencies and its average light level was probably much lower, less than a third of the average light level of Van Nes's sinusoidal spatial frequency display. Recall that Van Nes found the greatest visibility improvements only with low spatial frequencies and high light levels – with spatial frequencies below 6 cycles/° and 850 td illuminating the retina. Higher spatial frequencies were adversely affected by motion. A truly 'comparable' finding, therefore, would require Krauskopf's bar to be seen without its sharp edges when its visibility was improved by target oscillation, that is, its high spatial frequencies would be missing from the percept. There is no mention of this in the published report and we have no way of knowing whether it underwent such appropriate changes in appearance. Recall also that Van Nes's grating moved at constant velocity; it always moved in the same direction. Krauskopf's targets oscillated sinusoidally. The difference in these patterns of motion may have important consequences for the effect of retinal image motion on visibility, as will be pointed out in section 8.2 when Kelly's measurements are compared with ours.

Research on the relationship of retinal image motion to contrast sensitivity was a timely undertaking in 1968, with Ercoles and Zoli reporting that constant-velocity motion of a bright Landolt ring, seen as a luminance increment on a bright background, enhanced the visibility of gaps in the ring of various sizes when target speed was about 2–3°/s. This visibility improvement, to use Van Nes's term, was observed both when the ring was presented at the fixation point and also when it was displaced 1–2° vertically from the fixation point. A decade earlier, Fiorentini and Ercoles (1957) had shown that sinusoidal oscillation of a test field at 1–3 Hz enhanced the visibility of Mach Bands – results similar to Krauskopf's in the same period.

Westheimer (1965) and Lit (1968) published representative reviews of the status of research and theories of 'visual acuity' during the early years of Fourier Optics. Both made reference to Fourier Optics but both still dealt mainly with the traditional problems of visual acuity present in Helmholtz's influential treatment of the subject a century earlier. To illustrate, slightly over 1 of the 13 pages of Westheimer's text is devoted to 'Fourier Theory and Resolution' – most of it sceptical about its application in visual science, while Lit devoted about the same space in his 23 pages of text.

7.2.4. Eye movements in Arend's model of spatio-temporal processing

In 1973 Arend published a theoretical paper on differential and integral operations in the human visual system in which eye movement was afforded a major role. Arend's treatment of spatio-temporal processing represents the most recent attempt, to our knowledge, to use eye movements in a general theory that includes the perception of brightness, color and contour in a single model. Arend's model will be described in some detail, more because of its novelty and scope than because of its influence on the visual science community in the decade and a half since it was published. Our discussion will emphasize his treatment of the role of eye movements only with respect to the detection of contrast and contour. His treatment of their role in the discrimination of hues and in the perception of absolute levels of hue and brightness is beyond the scope of this chapter.

Arend ascribes the source of his interest in this problem to the preceding two decades of research on stabilized images in which "experiments have consistently shown that the retinal stimulus must continually change over time if color perception is to continue ['color', as used here, includes the hue, brightness and saturation of a light]. In normal viewing, if the stimulus light itself is not temporally modulated, the principal source of temporal changes on the retina is excursions of contours on the retinal surface as the eye moves. Under these extremely common stimulus conditions, image-movement-generated responses carry essential information to the central visual system about the light falling on the retina... The pattern of temporal changes on the retina is a function not only of image characteristics, that is, the spatial rate of change of the stimulus illuminance and chromaticity but also of the pattern of motion of the image relative to the retina, a pattern dependent upon both eye movements and movements of objects within the external object distribution itself" (p. 374). This idea is shown schematically in Figs. 12 and 13.

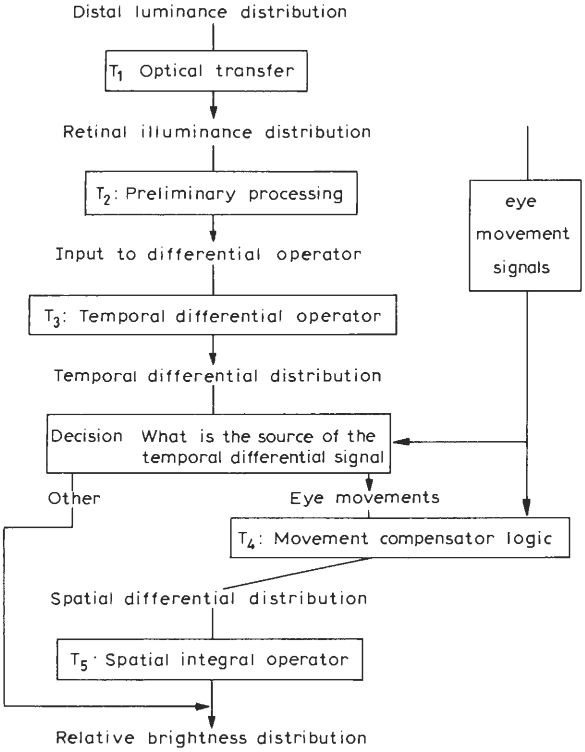

Fig. 12. Arend's (1973) block diagram of the stages of processing and intermediate spatial distributions of responses for responses generated by movement of a retinal image. See the text for an explanation.

The stages in the model up to the box labeled DECISION in Fig. 12 are described as "not radically different from previous models" (p. 378). We concur and will only summarize them very briefly here. Arend's treatment of the early stages of visual processing would currently be called a retinal 'center-surround organization, opponent process' model. Models of early retinal processing of this kind are derived from ideas developed by Ernst Mach towards the end of the 19th century. Ratliff (1965) brought Mach's ideas back to prominence as he developed models of lateral inhibition in the compound eye of the horseshoe crab. Arend was well aware of Ratliff's influences and cites them accordingly. Such models are also derived from Kuffler's (1953) observations of the functional properties of the retinal ganglion cells of the cat. This influence can be seen in the theoretical weight-

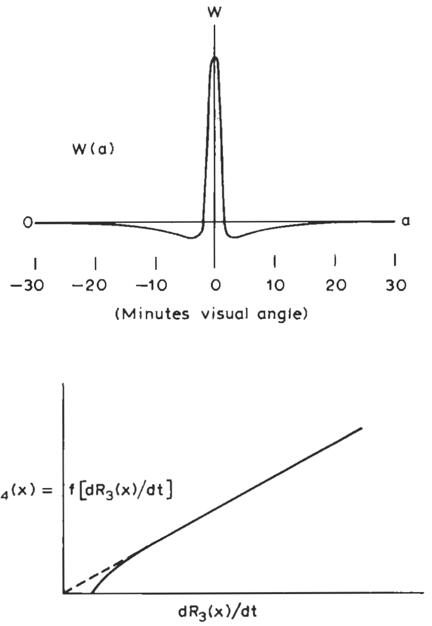

Fig. 13. (Top) Weighting function for a transfer from the luminance distribution to $R_3(x)$. The abscissa represents distance on the retinal surface in units of visual angle, *a*; the ordinate is $W(a)$. (Bottom) The function relating $R_4(x)$ to $dR_3(x)/dt$. The solid line is this function. The dotted line is the extrapolation of the line, $R_4(x) = kdR_3(x)/dt$. See text for an explanation. (From Arend, 1973)

ing function plotted at the top of Arend's model shown in Fig. 13 where inhibitory flanks are shown adjacent to an excitatory central region. For Mach (and Arend), what mattered for seeing was not only retinal illumination, an obvious retinal requirement, but it was also necessary to have the second derivatives in *both space and time* of the retinal illumination.

In Fig. 12 the OPTICAL TRANSFER and PRELIMINARY PROCESSING boxes refer to optical, photochemical and neural processing. These processes are assumed to be modeled by an approximately logarithmic transformation followed by lateral inhibition of the kind proposed by Ratliff (1965, pp. 77–142) to account for Mach bands and related phenomena. The equation for these processes on a single dimension along the retinal surface is:

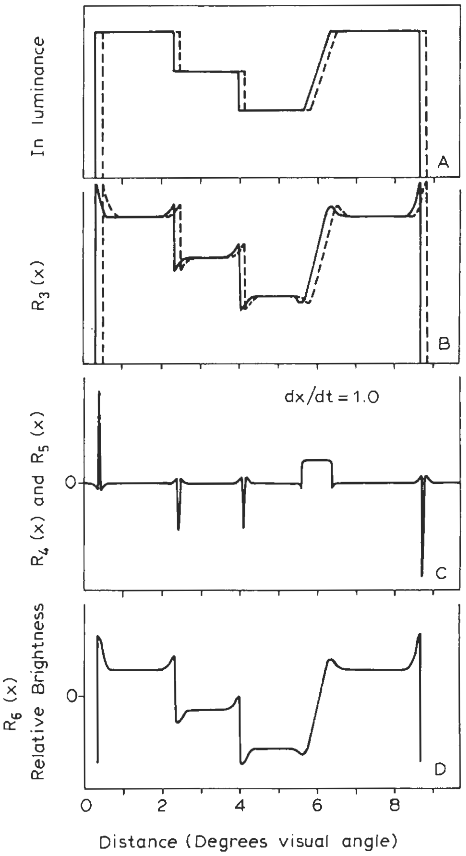

Fig. 14. Stages of processing of the spatial distribution of $\ln L$ is shown in part A of the figure. The solid line in part B shows the $R_3(x)$ for values of *x* along the abscissa at an instant when the eye is moving so as to displace the R_3 distribution towards lower values of *x*. The dotted line shows $R_3(x)$ a moment later after a displacement. Part C shows $R_4(x)$, and assuming that $dx/dt = 1$, $R_5(x)$, Part D shows the relative brightness distribution, $R_6(x)$, obtained by integrative processing of $R_5(x)$. (From Arend, 1973)

$$R_3(x) = \int_{-\infty}^{+\infty} W(a) \ln L(x-a)da$$

where $R_3(x)$ is the value of response distribution at Location *x*, $W(a)$ is the value of the line response function shown at the top of Fig. 13, at point *a*, and $L(x-a)$ is the luminance at location $x-a$. It is assumed that the temporal properties of box, T_2: PRELIMINARY PROCESSING, are complicated and, for simplicity, it is proposed that the luminance term in the equation above is a running average of the luminance taken over 10–20-ms intervals. In other words, it is assumed that there "is temporal as well as spatial blur". T_3 is the TEMPORAL DIFFERENTIAL OPERATOR. It represents the phasic response properties of the visual pathways which may be due to either a "passive adaptive process" or to "some active, opponent process". T_3 is described by the function:

$$R_4(x) = f[dR_3(x)/dt]$$

where $R_4(x)$ is defined in the same way as $R_3(x)$, $dR_3(x)$ is the momentary temporal rate of change of $R_3(x)$, and f is the function shown at the bottom of Fig. 13. Arend includes an important feature of his model in Fig. 13, which is not singled out for emphasis in his text. Namely, he shows the rate of change of the luminance distribution on the retina dropping to zero, close to but not at the origin of the abscissae. This means that he is assuming that there is a threshold velocity below which things will not be seen; some motion of the luminance distribution on the retina is assumed to be essential for vision. We will consider his justification of this assumption later when we evaluate the model, and it is sufficient here for the reader to notice that it is explicit and important to the success of his model. This velocity threshold provides the basis for explaining perceptual 'filling in', one of the main predictions of the model proposed in this paper. The same assumption will be made in all of Arend's subsequent work, which continues to emphasize the assumption of a velocity threshold and, thereby, Arend's persistent insistence on transient stimulation as the progenitor of all visual processing, an insistence likely to favor an important role for eye movement in visual processing (see section 9.5 for Arend and Timberlake's recent computation of the probable upper limit of this velocity threshold and Krauskopf's stabilization experiment described in section 6.2.7 for an example of 'filling in').

The later stages of the model, starting with the box labeled DECISION, are more novel and also of greater specific relevance to the topic of this chapter. The DECISION box is "required to separate eye-movement-related R_4s from R_4s produced either by object motion or by temporal modulation of the external stimulus". The proposed basis for the separation is "the temporal correlation between eye-movement-generated R_4s and responses specifying the pattern of movement of the eye" (p. 378). Arend does not describe the basis upon which retinal and extraretinal sources of position information ("inflow and/or outflow" signals) actually allow us to discriminate the motion of our eyes, head and body from motions of objects contained within the visual scene (see Wallach, Ch. 6 of this volume, and Skavenski, Ch. 5 of this volume, for a treatment of this and related problems). Instead, Arend simply makes reference to the fact that we can discriminate self-motion from object-motion reliably and, therefore, assumes that the information required for such discrimination "is available within the visual system". Having made this assumption, it becomes possible to go on to assume that "those R_4s not temporally correlated with eye movement signals undergo transformations different from T_4 and T_5, [transformations] which distort their information" (p. 378). Arend considers "temporal changes of the retinal stimulus not produced by eye movements beyond the immediate scope of this paper" (p. 392) and provides only a relatively terse treatment of how they might be included in a subsequent more elaborate model that includes such stimulation. In essence, he proposes that the rate of motion of objects relative to their backgrounds could provide the same kind of information about temporal changes as is provided by the generation of eye movements (the interested reader should consult pp. 392–393 of Arend's paper for details of his treatment of this problem. We will not go into it in any detail here because it is germane to the perception of motion and position constancy rather than to our topic, the perception of contrast and detail).

Next, we return to details of Arend's model, considering only eye-movement-generated responses.

T_4: MOVEMENT COMPARATOR LOGIC in Fig. 12 is a process in which eye-movement signals are used to evaluate the spatial differential responses (R_4) with which they are correlated temporally. Every R_4 that is above threshold is associated with a unique $dR_3(x)/dt$. Because the rate of eye movement which produces $R_4(x)$ is known, $dR_3(x)/dx$ may be determined by:

$$R_5(x) = R_4(x)/(dx/dt) \approx kdR_3(x)/dx$$

where dx/dt is the momentary rate of movement of the image relative to the retina, and k is the slope-constant of the linear portion of the function, f. An approximation symbol is used because of the threshold in this function.

T_5 is a discrete analogue of indefinite integration. Its output can be computed by arbitrarily assigning a value of zero to the output of the integration, $R_6(x)$, at any specific x. Then:

$$R_6(x+d) = R_6(x) + R_5(x+d)$$

where d is a very small unit increment of x. $R_6(x)$ is a relative brightness distribution with interval scale properties. The computation does not necessarily involve a 'scanning' mechanism that sweeps the distribution of R_5 over time. The brightness scale could be constructed by means of simultaneous operations. The initial arbitrary assignment of zero to the integration output, which was made for the purpose of computation, can be reassigned to an identifiable locus in the visual field, specifically, the locus whose R_6 value is the midpoint of the range of the R_6 distribution.

7.2.5. Predictions of Arend's model

The performance of Arend's model, when presented with step-changes of luminance across the retinal surface, is illustrated in Fig. 14. The spatial luminance distribution is stationary with respect to the physical world but its position changes on the retinal surface as it is "viewed with careful voluntary fixation" (p. 379). Such a displacement of the light distribution is represented by the differences between the solid and dashed lines shown in Fig. 14A. Fig. 14B–D shows the successive responses of the successive processing stages of the model which are produced by this kind of displacement of the stimulus. Fig. 14D represents the predicted relative brightness distribution resulting from the processing schematized in Fig. 12. This predicted brightness distribution resembles the actual appearance of a light distribution with such spatio-temporal properties inasmuch as it shows characteristic border-enhancement effects, that is, peaks and troughs, which correspond to the bright and dark bars perceived near regions of abrupt changes in luminance distributions. These peaks and troughs had, until recently (Ratliff, 1984), been widely believed to be caused by the operation of lateral inhibitory sensory processes at *abrupt* luminance borders, the same kind of inhibitory sensory mechanisms believed to be responsible for the perception of bright and dark Mach bands in retinal regions where there is a change in the rate of luminance change (that is, they were an example of the visual system responding to the second derivative of the retinal light distribution, Mach's original explanation of his bands). This property in a retinal light distribution is represented in Fig. 14A, which includes one typical Mach band stimulus (the relatively gradual slope drawn, one step-change in, on the right side of the abscissae).

Arend goes on to support his model mainly with reference to results from stabilized image experiments, primarily results of Yarbus and Gerrits and his co-workers, who reported that complete disappearance of stabilized targets will occur once stabilization is complete and that nothing will be seen as long as the stabilized image is not disturbed. He also demonstrates that his model predicts the appearance of the Craik-O'Brien-Cornsweet illusion. This illusion consists of creating the perception of an 'artificial contour' between areas of equal luminance. The artificial contour consists of an abrupt edge of luminance flanked by luminance gradients, which are below threshold, and return the luminance to the same level on each side of the abrupt edge. The percept is of an ordinary edge, that is, a contour separating two areas of uniform, but different, brightness. It is an illusion because these areas actually have equal luminance. Arend's model predicts this illusion and similar perceptual effects rather well, at least with respect to their qualitative properties. Effects such as these all require that "for a carefully fixating subject... there is a threshold spatial rate of change which must be present for any spatial brightness change to be perceived at that location" (p. 381). For Arend, image motions which exceed this threshold rate of change are provided by motions inherent in the eye movement pattern of the subject when he fixates carefully.

7.2.6. Evaluation of Arend's model

Arend's model, as might be expected, was particularly successful in explaining phenomena acknowledged to inspire its development, namely, 'artificial contours' in the Craik-O'Brien-Cornsweet illusion, and, more generally, 'filling-in' phenomena of the kind described by Krauskopf (see section 6.2.7). His explanation of these phenomena is derived largely from the assumption of a retinal image velocity threshold below which visual neurons are deprived of the stimulation they require to transmit information. Here it is assumed that *all* visual neurons signal transients, all are incapable of signalling the presence of a *stationary* discontinuity in the retinal light intensity distribution. This assumption can be viewed as controversial – it was not compelled by all data available when the model was proposed (see below).

Arend's model not only rested on a potentially controversial assumption, its support was entirely qualitative; for example, a demonstration that the perception of an 'artificial contour' can be predicted for an extension of the Craik-O'Brien-Cornsweet illusion. Quantitative estimates of the assumed retinal image threshold velocity were not provided, nor were specific eye movement characteristics associated with 'careful voluntary fixation' related to an estimate of this velocity threshold. There are hints, implicit in Arend's selection of results from the stabilized image research literature, which suggest that the threshold velocity might be rather low – probably in the range of retinal image velocities that would be generated by drifts, rather than by microsaccades, during 'careful voluntary fixation'. A selection of drift-like velocities for generating visual information in the model, a selection quite different from Ditchburn's (1973, 1980) life-long preference for the occasional, high-velocity transients produced by fixational microsaccades, would have been prescient because current thinking favors the continual, lower velocities associated with drifts (see section 6.2, or Ditchburn, 1980; Kowler and Steinman, 1980). But our inference that Arend actually preferred drifts over microsaccades as the functionally significant fixational eye movement is only a guess, because Arend avoided explicit treatment of fixational eye movement characteristics in his model. In this respect, Arend's model is very different from the Marshall-Talbot model, which proposed a specific fixational eye movement, physiological nystagmus, as the functionally significant behavior. Their proposal included specific assumptions about the dynamic properties of this high-frequency fixational tremor, as well as assumptions about its amplitude relative to the grain of the retinal mosaic. Unfortunately, the breathing-space Arend gained by being vague about the functional significance of particular fixational eye movement characteristics makes it difficult to evaluate how the model works during what Arend calls 'careful voluntary fixation'. Are the retinal image velocities typically produced by fixational drifts suprathreshold, or is effective stimulation confined to the very brief, infrequent transients produced by fixational microsaccades, or are all fixational eye movements equally effective once they exceed the velocity threshold? These, and other possibilities, are left entirely open.

Other features of Arend's model can be evaluated more directly. Arend's paper contains a footnote of considerable importance – a footnote in which Arend completely discounts Barlow's (1963) description of the appearance of the visual field after stabilization has been maintained for several minutes. Barlow observed what Arend (1973) describes as "residual hazy gradients of brightness in the field, roughly corresponding to the distribution of illuminance of the retina" (p. 375) (see section 6.2.6

for Barlow's description). Arend ascribes this observation to Barlow's failure to achieve adequate stabilization. Arend rests his interpretation of Barlow's result on "the weight of evidence against this view", specifically citing Yarbus's and Gerrits's reports of complete, persistent disappearance of their stabilized images. Here, of course, Arend is ignoring some important facts. Namely, Barlow had shown that Yarbus's technique (1) had very poor optical properties, producing a good deal of chromatic and other aberrations which precluded sharply focussed target images, and (2) allowed water, condensed from the air in contact with the cornea, to accumulate on the inner surface of the air-spaced compensating lens attached to the front of the sucker. Such fogging-up of the compensating lens greatly reduces the contrast in the retinal image of the target, as does the addition of trans-scleral light, which can find its way to the retina if the stimulating light is intense and the sclera is not screened. Yarbus did not screen the sclera. In short, Yarbus's technique was elegant in its simplicity but it was crude. It reduced the image quality and contrast of the stabilized retinal image to such an extent that, from Barlow's point of view, it produced conditions favorable for complete, persistent disappearance for much the same reasons that entoptic phenomena, such as retinal blood vessels, disappear and remain invisible. They are also of low contrast and fuzzy. They fade easily and remain invisible because of these properties. The trick in stabilization research is to find out what happens when a well-focussed, high-intensity, high-contrast target is stabilized and confined to the central fovea where cellular density is highest and receptive fields are very small.

Gerrits and his co-workers recognized the second flaw in Yarbus's technique and went to great lengths to eliminate condensed water in their elaboration of Yarbus's method, in the process loading down the sucker with tubing to circulate warm water around the air between the cornea and the compensating lens. They did not, however, satisfy all of Barlow's objections to the method. Gerrits and his co-workers also never succeeded in providing a retinal image with excellent optical quality (they always used simple, air-spaced compensating lenses and often permitted appreciable trans-scleral light as well) nor did they study targets stabilized on and confined to the central fovea. All of their stabilized targets had a relatively coarse grain and/or extended into perifoveal retinal areas where the complete disappearance of a somewhat degraded visual stimulus would be much more likely to occur than with a stimulus of better optical quality confined entirely to the central fovea. In short, Arend's complete rejection of Barlow's observation of a degraded but persistent visual stimulus after stabilization, in our view, rested rather clumsily on a controversial interpretation of the literature on vision during stabilized viewing. (See sections 6.2.4–6.2.9 for additional details of the Barlow-Yarbus-Gerrits research on the appearance of stabilized images. The situation has not changed since 1973 when Arend first published his model. We still do not know what would happen to the appearance of a high-contrast, high-luminance target, with 'normal' optical quality, which was stabilized, perfectly, and confined to the foveal center.)

It is hard to understand why Arend was, and has been throughout the ensuing years, so committed to discounting completely a role in visual perception for the 'sustained, non-adapting neural processes, carrying information about the retinal illuminance distribution' (Arend, 1973, p. 375), which were implied by Barlow's result and posited by him. This is particularly puzzling when it is remembered that the modern emphasis on phasic responses in the visual nervous system can be traced, from the Marshall-Talbot theory on down (see section 5), to Hartline's (1938, 1940) recordings from frog retinal ganglion cells which included a population of tonic units (20%), whose responses would provide precisely the kind of non-adapting sustained responses Barlow's observations would require (see section 5.1). It is true that Hartline reported more phasic units (50%) in these seminal papers, but it is unclear why the potential visual significance of these tonic units should be discounted *entirely*. Arend, in our view, seems to have sought simplicity in his model at the expense of plausibility. But, as we shall see, Arend went on to show recently that sustained units, should they have functional significance for visual processing, would only come to operate *alone* at retinal image velocities far below a value likely to obtain in any natural viewing condition or even under the best possible stabilizing conditions (see section 9.5).

To summarize, Arend afforded eye movements a critical role in the detection of spatial detail. Their importance stems from Arend's insistence on transient stimulation of phasic visual neurons as the exclusive source of significant visual input; such neurons are completely blind to stationary illumination gradients. This idea is incorporated in the model by the assumption of a retinal image velocity threshold. Also central to Arend's model is the idea that the visual brain must be able to distinguish between retinal image velocity signals associated with movements of the eye, head and body and retinal image velocity signals generated by motions of external objects. In other words, the visual brain must be able to distinguish between intrinsic and extrinsic sources of inescapable temporal and spatial blur. This is accomplished by feeding an eye movement signal into a 'comparator', which manages somehow to separate eye-movement-produced transients from other transients. The way the brain might accomplish this is not described. Arend's support for the idea rests entirely on perceptual accomplishments which are themselves still in need of additional psychophysical observations that suggest plausible neurophysiological explanations (see Skavenski and Wallach, this volume, for the current status of work on this problem.) The relationship of the proposed retinal image velocity threshold to characteristics of normal fixational eye movements was also left open.

These are large matters to leave unresolved. The model, notwithstanding these important omissions, has merit in our view, primarily because it is a 'dynamic' model, and as such encouraged consideration once again of the importance of eye movement to theories of spatial vision. As we shall see in later sections, the importance of incorporating some kind of dynamic processing of retinal image motion produced by eye movement continues to be an important problem – we will argue later, *the* most important current problem. In this, Arend, in our view, was clearly on the right track. Unfortunately, his model does not describe the functionally significant input to the comparator logic or how the comparator operates – both descriptions are required to allow us to move beyond an insistence on the need to consider eye movement, either to drive phasic units as Arend supposed or, more likely from our point of view, to keep wildly fluctuating binocular retinal images in some kind of functional registration so that details of edges can be extracted from these inputs, which are varying, very appreciably, in space and time.

7.2.7. Motion can enhance the visibility of low spatial frequencies

Shortly after publishing his model, Arend (1976) showed that retinal image motion can be beneficial for the detection of low spatial frequency sinusoidal gratings. His experiment was similar to Van Nes's (1968) earlier report, which had already shown that retinal image motion made it easier to see low spatial frequency gratings (see section 7.2.1). The main difference between the two experiments was the manner in which the grating was moved on the retina. Van Nes moved the grating. Arend moved the eye. A comparison of some of the details of the two experiments will be helpful before dealing with their significance for our topic. Namely:

Van Nes required the subject to maintain fixation at the center of the stationary 2.4° × 1.2° TV display while a sinusoidal spatial frequency grating remained stationary or drifted at a constant velocity, ranging from 0 up to 13°/s (effects of image velocities up to 39°/s were reported previously by Van Nes et al., 1967). The grating was drifted either up or down on a given trial while its contrast threshold was measured by the Method of Limits at 1 of 3 luminance levels, spanning a hundred-fold range. The Method of Limits, with contrast steps of 0.1 decade, was used because the Method of Adjustment had been found to be much less precise in a prior experiment with monochromatic sinusoidal gratings (Van Nes and Bouman, 1967). Van Nes et al. (1967) reported substantial detrimental effects of imposed, relatively low velocity drifts (>0.5°/s) on relatively high spatial frequency gratings (>6 cycles/°). None of Van Nes's reports was cited by Arend, who also failed to take note of Krauskopf's work on image motion (1957, 1962, or sections 6.2.2 and 7.2.3), work that Van Nes credited as the first demonstration of a beneficial effect of relatively low velocity retinal image motion on visibility as well as providing an early, accurate indication of the detrimental effects produced by faster image motions.

Arend's subjects fixated a luminous spot when it remained stationary at the center of a 3° × 4° display or smoothly pursued the spot when it moved back and forth, horizontally, at a constant speed; speed varied from trial to trial up to a maximum of 5°/s. The display contained a spatial frequency grating whose contrast was set to threshold by the Method of Adjustment. The subject made two kinds of threshold setting during smooth pursuit when spatial frequency was 5 cycles/° or higher. One setting was made at the middle of each track when eye speed would be expected to be fastest. A second setting was made near turnabouts when the eye would be expected to be pursuing more slowly, either slowing down or speeding up. All experiments were performed only with a single, relatively low, space-average luminance (about 7 mL). This light intensity was near the lowest level used by Van Nes, who, you will remember, had obtained his most striking results at his highest light level, about 130 mL.

Inferences about retinal image velocity in both the Arend and Van Nes experiments depended entirely on assumed characteristics of maintained fixation and smooth pursuit under conditions that had not been studied when their research was performed (see section 7.2.2 for a discussion of this point). Arend, for example, assumed that fixating a stationary spot at the center of more-or-less structured backgrounds (gratings with different numbers of 'bars') or smoothly pursuing the spot when it moved back and forth across one of these backgrounds would lead to essentially normal fixational or smooth pursuit eye movement patterns. Both assumptions did, in fact, prove to be reasonable when they were tested in eye movement experiments (Murphy et al., 1975), but they were by no means obvious before these experiments were performed. Even now, we know that these assumptions are not applicable to all subjects – an occasional subject with poor eye movement control slows down his smooth pursuit and increases his saccade rate when the pursuit target moves across a highly structured background (Collewijn and Tamminga, 1984). Van Nes had an analogous problem. He was forced to assume that a subject could fixate at the center of a relatively small, stationary frame and avoid pursuing a grating when it moved within the frame. Van Nes's assumption also proved to be reasonable when tested in eye movement experiments but, once again, this outcome was fortuitous, inspired entirely by subjective impressions about the fixational eye movement pattern under rather special experimental conditions (see Kowler et al., 1984; Kowler, this volume, for a discussion of maintained fixation and smooth pursuit in highly structured visual fields).

There is, however, an additional and even more fundamental assumption about the relationship of eye movements to visual processing that is raised by the Van Nes and Arend experiments. This assumption is not confined to their research; it permeates the visual science literature. Recent developments have brought it into question. Note that in both experiments it was assumed that paying attention to maintaining fixation at the center of a TV display, while *at the same time* paying attention to the visibility of striations within the grating (Van Nes's experiment), or paying attention to smoothly pursuing a spot moving back and forth across a grating, while *at the same time* paying attention to striations within the grating (Arend's experiment), does not influence the threshold measurements of contrast. It was also assumed that paying attention to the contrast of the displays did not influence the stability of fixation or the effectiveness of smooth pursuit. This assumption can be described in the terminology of cognitive psychology as follows: attentional resources can be allocated fully and independently to a visually guided oculomotor task and to a visual psychophysical discrimination task when they are performed simultaneously. Khurana and Kowler (1987) tested and rejected this assumption recently (see Kowler, this volume, for details). The conclusion of their research, once again, in the terminology of cognitive psychology can be described as follows: attending to visual stimuli for purposes of visual information processing and, *at the same time*, attending to another visual stimulus in order to exercise oculomotor control draws on the same reservoir of attentional resources. In other words, visual attention is unidimensional. 'Looking' in order to see and 'looking' in order to move the eye are not independent cognitive capacities. The significance of these experiments goes far beyond the scope of our chapter. Potentially, they have broad implications throughout visual psychophysics. Whenever an observer is asked to fixate or to pursue a visual feature extraneous to the subject matter of the particular psychophysical investigation, his performance on one task may affect his performance on the other. The magnitude and generality of such oculomotor and psychophysical interactions must still be determined (see Murphy, 1978, experiment 3, p. 527, for a fine example of this interaction reported 9 years before its significance was fully appreciated).

Despite these potential ambiguities in the interpretation of the role of retinal image motion in contrast sensitivity, the main results of the Van Nes and Arend experiments have proven to be relatively robust in the light of present knowledge – a fact that encouraged us to include them in this chapter. A somewhat detailed discussion of the goal and significance of Van Nes's experiment was presented in section 7.2.1. A similar treatment of Arend's (1976) experiment is presented next.

7.2.8. Image motion and the shape of the CSF

Arend undertook his experiments in order to clarify the mechanism responsible for the shape of the spatial MTF function in the light of "the discovery of psychophysical visual mechanisms responsive to narrow bands of spatial frequency (Blakemore and Campbell, 1969; Sachs et al., 1971), [which had become a problem when] Graham [1972] demonstrated with adaptation techniques that changes of MTF shape, resulting from temporal modulation, may not be attributed to drastic changes in the tuning of the narrow band channels. . . While some of the variation of thresholds for stimuli of different spatial frequencies may be attributable to differences in the structural properties of the underlying psychophysical channels, it is clear that factors which do not reflect differences among the channels may also affect the shape of the MTF (e.g. optical blur). It is essential, therefore, that other possible sources of variation be examined in detail; the only current means of obtaining psychophysical evidence concerning the relative sensitivities of channels tuned to different spatial frequencies is progressive correction of the spatial MTF for mechanisms common to all the channels" (Arend, 1976, p. 1035) (see sections 6 and 7 for the 'channel hypothesis' in Fourier Optics).

Arend (1973) had already pointed out in his model paper (described in section 7.2.4) that the temporal rate of change of retinal illuminance would be such a factor that should affect the shape of the spatial MTF. The shape of the spatial MTF is 'strongly influenced' by temporal variations, particularly for spatial frequencies below the region of maximal sensitivity to contrast, i.e., 3–6 cycles/°. Once allowance is made for differences in sensitivity to the temporal rate of change of the retinal illumination gradient, low spatial frequency 'channels' do not differ in peak sensitivity. Thus, the well-known, progressive loss of sensitivity to increasingly coarse gratings need not, in itself, be an indication of differences in contrast sensitivity. It could indicate just as well that all low-frequency 'channels' require the same *amount* of contrast to reach threshold, but this contrast threshold depends on temporal, as well as on spatial, variations of contrast, and that temporal variations must occur more rapidly when spatial frequency is lowered.

Arend (1976) made two assumptions about these temporal variations. First, temporal variations are assumed to be essential for vision. This assumption is based on the somewhat controversial fact (see section 7.2.6) that all visual images (including intense, high-contrast visual targets confined to the central fovea) will fade completely and remain invisible when all temporal variations are prevented by an effective image-stabilizing technique. It is also assumed that these temporal variations must exceed some *temporal* threshold value. In Arend's words, "the temporal change of retinal illuminance must exceed a criterion magnitude if the subject is to detect spatial nonuniformity or pattern of any spatial frequency" (p. 1035). Three additional assumptions are made in order to explain the "linear decline of contrast sensitivity commonly found at low spatial frequencies with steady stimulus presentation" (p. 1035). First, it is assumed that there is a monotonic relationship between the response of the visual system in a local retinal region to the rate of change of the illumination falling upon this local retinal region. Second, it is assumed that there is a *spatial pattern* threshold, that is, the detection of the spatial pattern occurs when the local response exceeds some criterion level. Finally, it is assumed that, in Arend's words, "the population of eye movements of a subject, viewing sinusoidal stimulus patterns near the contrast threshold, is approximately independent of the spatial frequencies of the patterns, at least when the subject is fixating a small spot" (p. 1035).

This last assumption is quite important for developing the role of eye movements in the detection of contrast because it leads to the conclusion that the decline in sensitivity for low spatial frequencies should be ascribed to characteristics of fixational eye movements. "If the population of eye movements is constant across spatial frequencies, the stimulus modulation amplitude must be increased in direct proportion to reductions of pattern spatial frequency if the same level of response is to occur. As a result there is a linear rise of threshold with decreasing spatial frequency, a prominent characteristic of the steady-state MTF. According to this model, then, the rise of threshold with decreasing spatial frequency merely reflects the steeper spatial slope of high frequency sinusoidal patterns relative to lower frequency patterns" (p. 1036). In short, once it is assumed that eye movement characteristics do not adjust themselves to the spatial frequency of the grating being viewed, low spatial frequencies may not produce the suprathreshold temporal and spatial variations required for detecting the striations in the coarse grating pattern.

Of course, the argument given above can only explain why very low spatial frequencies require more contrast than higher spatial frequencies. When spatial frequencies are relatively high (i.e., above the peak sensitivity region usually found to lie between 3 and 6 cycles/°), contrast must also be increased. When sinusoidal spatial frequency gratings as high as 50–60 cycles/° are viewed through the normal optics of the eye, striations cannot be seen at any contrast level. When this occurs, it is said that the 'high-frequency cutoff' of the spatial MTF has been reached. At relatively high spatial frequencies progressively more and more contrast is required. Arend ascribes part of this loss of sensitivity at high spatial frequencies to eye movements, explaining that "even the limited eye movements of fixation will produce rapid, multiple-cycle temporal changes at all points on the retina illuminated by the display" (p. 1036). Here, he is proposing that eye-movement-produced temporal blur adds to spatial blur caused by diffraction and other optical characteristics of the normal human eye (inherent dioptric errors). These aberrations become important when high spatial frequencies are viewed. Arend's main results are summarized in Fig. 15.

Visibility improvements were observed with spatial frequencies as high as 5 cycles/° when the stationary display was moved on the retina by smoothly pursuing a point that moved across the display at 0.5°/s. This improvement was relatively modest, about 0.05 log units, a bit more than 10%. 'Visibility improvement', here, as in Van Nes's experiments (see section 7.2.1), refers to increased contrast sensitivity produced by moving the grating of a

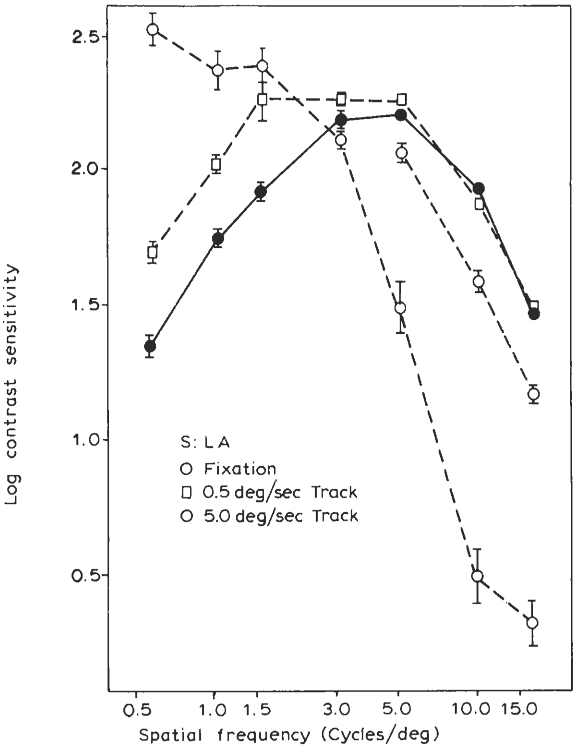

Fig. 15. Mean log sensitivities for seven spatial frequencies at three rates of target motion, namely, a stationary target (filled circles), 0.5°/s (open squares) and 5°/s (open circles). Contrast sensitivities are larger by 2 log units than those published in Arend (1976) due to a labeling error in his original figure. (From Arend, 1976)

particular spatial frequency on the retina relative to the contrast sensitivity observed with the same grating when its movements on the retina were produced exclusively by fixational eye movements. The improvement was about the same when thresholds were measured with the 3 cycle/° grating but increased to about 0.35 log units, a factor of 2.25, when spatial frequency was reduced to 1.5 cycles/°. At 1.5 cycles/° and below, the improvement in visibility reflects the progressively greater loss of sensitivity to fixated gratings relative to the sensitivity to gratings moved relatively slowly on the retina (0.5°/s). Still larger improvements in visibility, actual increases in sensitivity to low spatial frequency gratings rather than reduced decreases in sensitivity, were observed when the grating moved faster, that is, when nominal smooth pursuit velocity increased to 5°/s. At this velocity the improvement with 1.5 cycle/° grating was about 0.49 log units, a visibility improvement factor greater than 3. It increased further to about 0.65, a factor of 4.4, with 1 cycle/°, and further still to 1.2 log units, more than a factor of 10, with the 0.5 cycles/° grating. We are inclined to be somewhat sceptical of this last value because Arend's display only subtended 4° of visual angle, which means that Nyquist's criterion for estimating the frequency of a sinusoid was just met when the grating's spatial frequency was only 0.5 cycle/°.

The situation at 1 cycle/° is less ambiguous but here we note a somewhat troublesome difference in the results reported by Arend (1976) and Van Nes (1968). Namely, Van Nes's greatest visibility improvement was observed when a 1.14 cycle/° grating drifted at about 1°/s. This improvement was substantial (0.64 log units), essentially the same as Arend's greatest improvement once his lowest spatial frequency is ignored for reasons described just above. There were, however, very large differences in the conditions under which each author obtained his maximal improvement. Van Nes only got his large improvement when the space-average luminance of his display was at its highest level (about 130 mL). He obtained no measurable improvement when his display was set to be near the level Arend used (7 mL). Furthermore, Van Nes reported his maximum improvement when the 1.14 cycle/° grating drifted at 1°/s. This grating showed virtually no improvement when it drifted at 3°/s and suffered an appreciable adverse effect of motion when velocity was only 6°/s. Arend reported his maximum improvement with his 1 cycle/° grating when it moved at 5°/s. At 0.5°/s, the visibility improvement was less than half as great. There are, then, some clear differences in the outcomes of these experiments. Detailed speculation about possible reasons for these differences, such as differences in the way in which retinal image motion was produced or the effects of large differences in retinal illumination levels, will not be attempted here. Many variables could be responsible. Arend (1976, p. 1040) devotes considerable space to a discussion of the potential importance of the number of cycles in the grating displayed, and to the size of the display, including the diameter of its homogeneous surround, when he compares his results with other reports. Rather, we will call attention to other features of interest in Arend's summary graph. These features will be emphasized in subsequent sections when we reproduce contrast sensitivity functions measured in the presence of known 'natural' retinal image motion (see section 8.2). Note the large adverse effect of imposed image motion on contrast sensitivity when spatial frequency was at or above the 3–5 cycle/° peak of the spatial MTF. Five degree/second motions were very detrimental above 3 cycles/° with the high-frequency cutoff lying somewhere in the region of 10 to 15 cycles/°. Sensitivity falls by 1.6 log units, a factor > 40, as spatial frequency increases from 3 to 10 cycles/°. The crossover spatial frequency, that is, the spatial frequency where image motion helped below and hindered above, occurred at 3 cycles/°. This spatial frequency represented the lower end of the region of peak sensitivity observed during fixation or when the relatively slow 0.5°/s motion was pursued.

Arend provided a very nice demonstration figure which permits the reader to examine for himself the beneficial effects of image motion on low spatial frequency sinusoidal gratings and at the same time to examine the detrimental effect of the same motion on high spatial frequency gratings. His demonstration plate is reproduced in Fig. 16. It will be used to illustrate these phenomena now and will be used later to illustrate a number of related points. Place this figure, which contains the sum of a relatively low and a relatively high spatial frequency sinusoidal grating, at arm's length and fixate the small centered dark square (hold it in only one arm, leaving the other arm free to provide a moving stimulus). You should be able to see a relatively high spatial frequency sinusoidal grating clearly and a much lower spatial frequency grating, somewhat vaguely, at the same time. The figure contains about 10 cycles of the easily seen, high-frequency grating to a single cycle of the harder to see, coarse grating. At arm's length (about 75 cm) the entire display subtends about 11° horizontally and about 8° vertically. At this distance the low-frequency component of the display is about 0.4 cycles/° and the high-frequency component is about 4 cycles/°. The high frequency is near the peak sensitivity of the spatial MTF and the low frequency would fall just below the lowest spatial frequency Arend studied – the region receiving the greatest benefit from imposed motion (see Fig. 15).

Fixate the tip of your finger (or the end of an unsharpened pencil) and hold it at arm's length so that it rests just at the surface of Fig. 16. Now, move your finger tip back and forth from one side of the figure to the other at about 5°/s. It will take about 2 seconds to get across when you move your finger tip at the correct speed. Now, shift your attention to the appearance of the grating as you pursue your finger tip. Shift attention when your finger tip is near the center of the figure where your eye movements should be up to speed. Be sure, while you shift your attention to the appearance of the grating, that you are smoothly pursuing the tip of your finger and also be sure that your finger is moving back and forth, uniformly, at the correct speed. When you have this just right, the low spatial frequency component of the grating will stand out as broad dark bands. The high spatial frequency component, which stands out quite clearly when you fixate the small centered dark square, appears blurred while you smoothly pursue your finger tip and attend to the, now prominent, dark bands. These effects illustrate, quite compellingly, the visibility improvement retinal image motion can produce for low spatial frequencies while retinal image motion interferes with your ability to see relatively fine gratings – the findings first reported by Krauskopf (1957, 1962), and confirmed by Van Nes (1968) and then by Arend (1976).

Now that you have made these observations yourself, you will probably have been struck by how hard it was to know exactly what you were doing while you made them. Actually, you probably noticed that the broad dark bands popped out when you were not at all sure that you had the conditions exactly right. It is not easy to attend to how each of

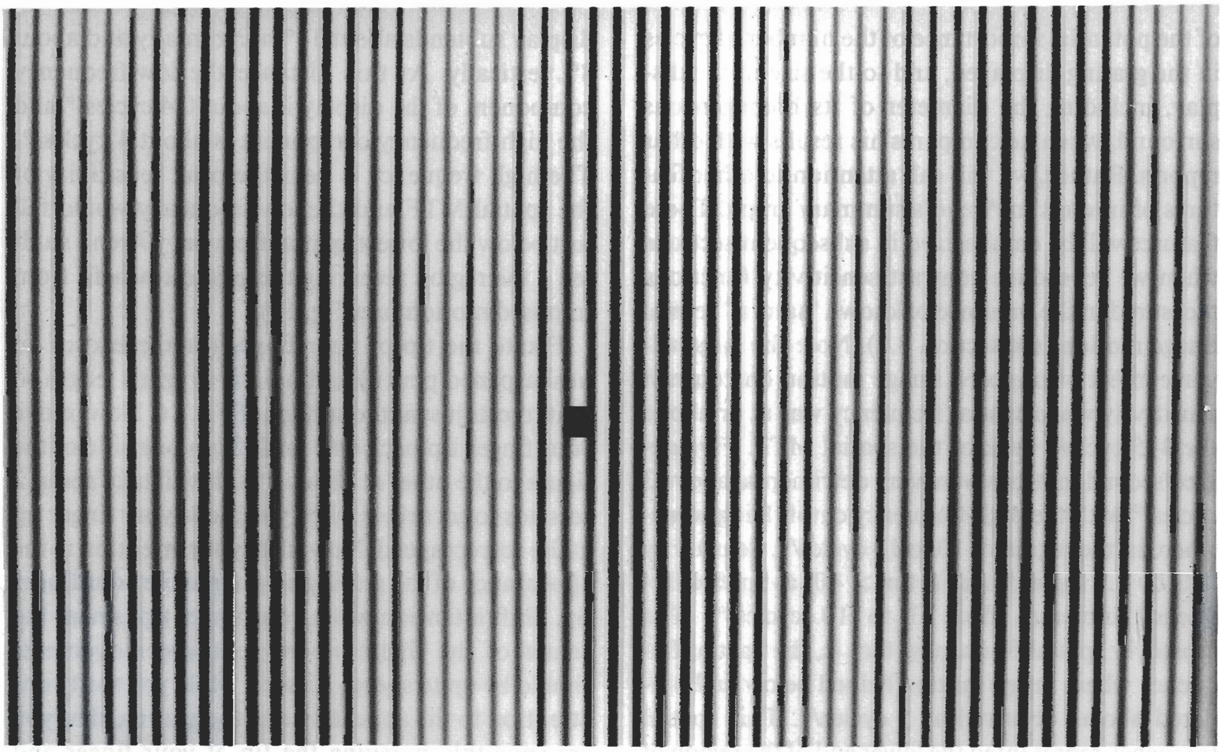

Fig. 16. Photograph of the sum of two sinusoidal spatial frequency gratings. See the text for the manner in which this figure should be viewed. (From Arend, 1976)

the two spatial-frequency components look while you attend to pursuing smoothly at the same time. Concerns such as these were raised, just above, when we alluded to recent experiments by Khurana and Kowler (1987) on the allocation of attentional resources between visual information processing and oculomotor control. Putting aside this still unanswered problem, it was well-established by 1968 (Van Nes) that retinal image motion can affect the detection of both low and high spatial frequencies, helping one and hurting the other.

How do these well-established experimental observations, which establish relationships between image-motion and the visibility of patterns of different spatial frequencies, relate to the visual processing of spatial patterns in everyday life? Specifically, is the effectiveness of oculomotor compensation for movements of the head sufficient to explain how we maintain pattern vision over the range of retinal image motions the human being typically encounters? We will consider this possibility first in the relatively simple case in which the observer's head is supported as he tries to make out details in a moving target by tracking the target with smooth-pursuit eye movements (one of the compensatory oculomotor behaviors). Here, we ask over what range of target velocities is the effectiveness of smooth pursuit sufficient to keep the residual retinal motion (retinal slip velocity) of the target image low enough for its details to be discerned? Once this question has been answered, we will then consider the more complex, but more natural situation, in which the observer is free to move about. In this situation, the observer can use both visual and vestibular signals to drive the full gamut of his oculomotor compensatory repertoire. Here, we would expect, on teleological grounds, that he would stabilize the retinal image of objects which catch his interest just enough to be optimal for vision, that is, to allow sufficient retinal slip to facilitate his detection of the low spatial frequency content, while at the same time keeping retinal slip low enough to preserve sufficient high spatial frequency content to allow the moving retinal images to provide good signals about the significance of the distal objects.

7.2.9. Slip limits contrast sensitivity during smooth pursuit

Murphy (1978) showed that retinal slip limits contrast sensitivity during smooth pursuit by making psychophysical contrast threshold measurements concurrent with accurate recordings of smooth-pursuit eye movements. In one condition, the subject tracked a constant-velocity motion of a 1.36° × 1.36° sinusoidal grating display which moved in synchrony with a 2-bright-point acuity test target located at its center. The acuity test target was adjusted to threshold (for both subjects, a 4′ separation between two bright points of light) just prior to the measurement of the threshold contrast of the grating. The composite test grating–fixation target display could be moved, horizontally, within a 5.38°-wide part of the CRT face where the raster was set to the same space-average luminance as the grating. An acuity target was used as a guide for fixation in order to establish and maintain the subject's crystalline lens focussed at the plane of the grating, which provided poor cues for accommodation when its contrast was near threshold. The moveable test portion of the display contained 7 cycles of a 5.14 cycle/° grating, whose contrast could be adjusted to threshold. A high criterion for 'threshold' was used, that is, the subject was required to set contrast so as to establish a "just visible pattern of bright and dark bars" rather than the lower criterion in which the subject sets threshold until he sees a "smudge (frequently described as an inhomogeneity in the display)" (p. 524). Murphy's high-criterion instruction was similar to the instruction given by Van Nes and Arend to their subjects, that is, they required the threshold to be set to 'stripes' or 'striations'.

A contact lens optical lever was used to record eye movements, a choice which allowed very accurate indications of smooth pursuit velocity but severely restricted the length of experimental sessions (see section 6.1.1). This restriction encouraged the use of the Method of Adjustment because of its relative speed, despite the fact that this psychophysical method was known to be influenced by starting-point biases and could be influenced readily by the subject's knowledge about the likely (or desired) outcome. Murphy went to some lengths to prevent such influences from contaminating his threshold measurements, e.g., the relationship between physical settings of the subject's potentiometer and the contrast it delivered was changed from trial to trial, and his subjects were not given any information about their performance ('receive feedback') while the experiments were under way.

Murphy included elements of both the Van Nes and the Arend experiments in his design (see section 7.2.7), that is, he produced retinal image motion by moving the grating and also by moving the eye. There were, however, three important differences between his and the previous experiments: namely, (1) Murphy used only a single spatial frequency, 5.14 cycles/°, a value near the high end of the region of peak sensitivity of the typical subject's spatial MTF, (2) he recorded his subjects's eye movements very accurately rather than making assumptions about the likely characteristics of their eye movements, and (3) his grating and fixation target moved together while the contrast of the grating was adjusted to threshold – this procedure might have simplified cognitive aspects of his subject's task, relative to the cognitive demands of Arend's task, because attention to smooth pursuit (a visuomotor task) was not divided between it and the concurrent visual psychophysical task (a visual pattern threshold judgement) as much as it was divided under the conditions employed by Arend. Arend left the grating stationary and the subject tracked the fixation target back and forth across the grating. The subject in Murphy's tracking condition (half of his trials) set the contrast of a moving grating while he pursued a special feature at its center which moved in precisely the same way as the grating.

The other half of Murphy's trials were similar to Van Nes's in that in both experiments the subject maintained fixation on a stationary target while he set the contrast of a moving grating to threshold. There were differences, however; namely, Murphy's subject maintained fixation on a centered 2-bright-point acuity target while Van Nes's subject maintained fixation at the estimated center of a 2.4° × 1.2° display with only the edges of the display, which fell on perifoveal retina, available to maintain fixation. Fixation was probably maintained better under Murphy's fixation conditions than under Van Nes's condition (see Steinman, 1965), particularly when grating contrast was relatively high (Murphy et al., 1975). The pattern of grating motion was also different in the two experiments. Van Nes drifted his grating at constant velocity in only one direction (mainly up) while Murphy's constant velocity oscillations went back and forth horizontally. This difference in the patterns of retinal image motion might be quite important – a possibility considered later when we discuss issues in current research.

When the subject fixated the stationary target and the grating oscillated in Murphy's experiment, the effect of retinal image motion on the contrast threshold could be related directly to the speed with which the grating moved. Murphy could do this because he found that "the moving grating did not interfere with the ability to maintain a steady line of sight" (p. 526). Specifically, when the grating was moved at speeds which ranged from about 0.5°/s up to about 2.4°/s drifts of the eye during maintained fixation ranged only between 0 and 0.07°/s. Moreover, no correlation between the direction of grating motion and the direction of drift was observed at any grating speed when the subject maintained fixation on a stationary target while the grating moved.

In Murphy's other condition in which the subject pursued the grating, which oscillated along with the fixation target at its center, retinal image motion was produced by the failure of the eye to exactly match the velocity of the moving composite display (the grating with the acuity target at its center). In other words, the retinal image of the composite display slipped during smooth pursuit because smooth pursuit did not stabilize the display on the retina perfectly, even when the display moved in a highly predictable pattern. When Murphy performed these experiments, the now well-established fact that there is usually appreciable retinal image slip during smooth pursuit was still somewhat controversial, a circumstance which contributed in part to Murphy's interest in this experimental condition (see Kowler, this volume, for current concerns about smooth pursuit, which have taken a very different turn during the past decade). Murphy was able to achieve useful values of retinal image slip by setting the peak-to-peak amplitude of the moving display to about 3.5° and its speed to values ranging from 0 up to about 7°/s. These smooth-pursuit conditions produced retinal image motions (slip speeds) ranging from 0 to over 1°/s. Murphy's graph showing retinal image speed as a function of target speed is reproduced in Fig. 17. The results he obtained when image motion was imposed on a grating, while fixation was maintained on a centered stationary fixation target, are also reproduced in Fig. 17.

Murphy found that retinal image motions which were produced by moving a grating while fixation was maintained on a stationary target or by slip of the display during smooth pursuit had the same effects on contrast thresholds. In both cases, retinal image motions up to almost 2°/s had only very modest adverse effects, allowing Murphy to conclude that "it is possible to see pattern despite appreciable retinal image motion" (p. 529). Westheimer and McKee (1975, 1978) published similar observations for Landolt C and hyperacuity targets (vernier acuity and stereoacuity) at about the same time.

8. Natural retinal image motion

8.1. Its origin and probable extent

Why do human beings have this capacity to tolerate appreciable retinal image motion? A decade ago, it was rather widely believed that compensatory eye movements usually made tolerance of retinal image motion unnecessary, providing only that the full repertoire of compensatory eye movements was available. Availability here implies not merely a normal oculomotor system, it also implies that this sytem is provided with the stimulation, necessary and sufficient, to allow the oculomotor system to operate in a natural way. The experiments described thus far did not meet this condition because they precluded natural oculomotor compensation by immobilizing the head on chin rests or biteboards. Under these conditions, only smooth pursuit could be used to reduce the retinal image motion of moving displays. Saccades could also be used to correct position errors, which accumulated as eye speed lagged behind target speed, but velocity errors (retinal image slip) would persist because smooth pursuit was only able to keep up with relatively slow motions of the target. Even when target motion was relatively slow, smooth pursuit could only provide partial compensation because it often 'overcompensated' by starting to move or by changing direction *before* the target started to move or to change direction. Both of these anticipatory actions produce or increase retinal image slip rather than reduce slip (see Chs. 1 and 2, this volume, for discussion of these 'predictive' properties of smooth pursuit).

Other sources of oculomotor compensation come into play once the head is free from artificial supports. Once the head is free, the effectiveness of natural oculomotor compensation can be observed because the VOR can come into play as accelerations of the head, acting on the fluid and hair cells in semicircular canals, generate neural signals proportional to the angular velocity of the head. These signals can drive the eye in the direction opposite to the direction of the head and thereby compensate for retinal image motions that would be produced when the head rotates. The VOR was assumed to operate synergistically with smooth pursuit. Smooth pursuit handled the relatively slow motions of images on the retina while the VOR took care of the faster, higher frequency, motions resulting from head movement. The scheme just described was entrenched in the neuromythology of previous decades because of the coherent, user-friendly, textbook-ready, message it contained. Like all mythology, however, the synergy of smooth pursuit and vestibulo-ocular response rested on common-sense beliefs rather than on verified observations. In this common-sense (ostensibly 'theoretical') oculomotor world, the oculomotor system, operating in its natural mode, could, by itself, eliminate the need for tolerance of retinal image motion except under relatively extreme conditions of natural bodily movement. In this period, it was customary to demonstrate the 'virtual perfection' of compensatory eye movements (Wilson and Melvill Jones, 1979, p. 287) by asking the reader to make the following kinds of observations.

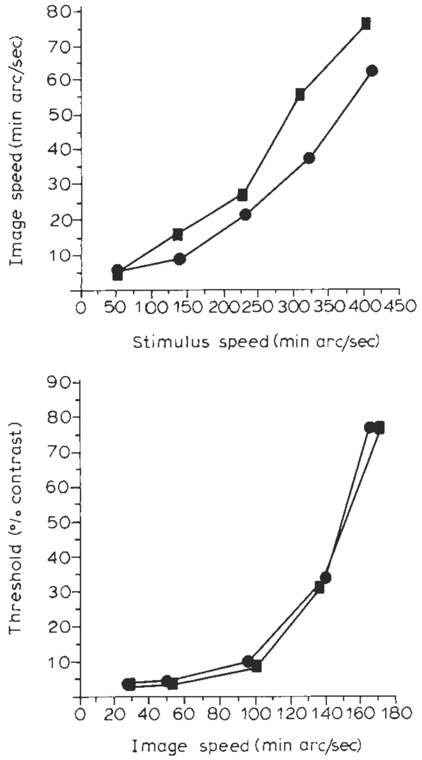

Fig. 17. (Top) Mean retinal image speed during smooth pursuit as a function of stimulus speed. The circles show subject RMS, the squares subject EK. (Bottom) Mean contrast threshold for a pattern when fixation was maintained on a stationary target and the grating moved. (From Murphy, 1978)

First, hold this book at arm's length with both hands and fixate a section of text at the center of the page. Now, move the book left and right at constant velocity through a comfortable angle (probably about 20 or 30°). First, try to keep your eyes centered in your head and your head as still as possible as you move the book. Note how quickly the text blurs, as retinal image motion smears the details in the proximal stimulus.

Having noted this effect, now allow your eyes to smoothly pursue the text as you continue to move it left and right. Note particularly the extent to which you can make out fine details near the center of your eye tracks when your smooth pursuit velocity would be at its highest and, consequently, the retinal image slip would be at its lowest value because your eye has reached its best possible speed and has not yet started to slow down in anticipation of your intention to reverse the direction of the moving book. Now, move the book left and right, faster and faster, reducing amplitude as you increase frequency, until the text blurs at the center of each track despite your best effort at smooth pursuit. You will find that you cannot move the book very fast and change direction very frequently before the text blurs despite your best efforts at smooth pursuit.

Now, having convinced yourself that you have a pretty good idea of the extent to which smooth pursuit can compensate for the motion of objects in the physical world, hold the book still at the center of your visual field and start oscillating your head about its vertical axis, all the while maintaining fixation of some portion of centered text. Move slowly at first and reduce the amplitude of your head oscillations as you increase their frequency. Keep speeding up, all the while trying to blur the text to the same extent it blurred when the text, rather than the head, was moving. Remember always to note what you see when your head is near the center of each oscillation.

You should now be convinced that oculomotor compensation can be virtually perfect until quite violent motions of the head are made. You will see motion, a jitter of the text when your head oscillates very rapidly. The best way to get very high frequency motions is to clench your teeth and strain the muscles of your jaws and anterior part of your neck and upper chest (tense your platysmas muscles). When you do this, you will notice that the finest details in the text will remain sharp until its jitter is near its maximum possible, naturally produced, value. The difference between moving the book with the head still and moving the head with the book still is the difference between compensation by smooth pursuit alone and compensation by the VOR supplemented by smooth pursuit. Observations such as these left oculomotor specialists (including one of us, RMS) feeling very smug. It was their system, the system they had chosen to study, which allowed human beings to see as they moved about in the real world. The belief that eye movements prevented unwanted retinal image slip made it plausible for the visual scientist to study vision conveniently after stabilizing his subject's head. In other words, observations and discoveries made with the head immobilized will generalize to the 'real world' precisely because of the effectiveness of oculomotor compensation. There was good reason for such oculomotor chauvinism 10 years ago; there is much less reason today now that the degree of oculomotor compensation has actually been measured rather than being inferred from informal perceptual observations of the kind just described.

8.1.1. Requirements for studying natural image motion

Some of the difficulties inherent in the development of techniques to make these kinds of measurements accurately and precisely in both space and time were not fully appreciated when this line of research began. Mistakes were made. The most embarrassing, in retrospect, was the failure to appreciate the extreme sensitivity of Robinson's (1963) magnetic field sensor coil technique to translations of the head within the small, usually about 0.6 meter, simple Helmholtz coils used by all in the 1960s and 70s (see Steinman, 1975, 1976; for measurements of head and eye movements with the sensor coil that confounded head translations with head and eye rotations).

The requirements for making accurate, translation-free measurements, while allowing head movement, were quite demanding. All prior work done with the head supported artificially had shown that the accuracy and precision of maintained fixation was of the order of 2–6 minutes of arc and that saccades, the fastest of the compensatory eye movements, had peak speeds ranging from about 5°/s to more than 500°/s when saccades shifted gaze from 6′ (the size of the average fixation microsaccade) to 75° (the size of the largest possible saccade that will bring the line of sight to fall directly on a small, very eccentric target). Performance such as this required an eye movement recording technique with accuracy and precision better than 1 minute of arc and bandwidth in the neighborhood of 200 Hz or better. It also required that rotations of the eye and head could be measured, unconfounded with spurious signals arising from translations of the head in space, at least when the targets were relatively far from the eye. When targets are far, translations of the head do not require the eye to rotate appreciably in order to maintain fixation (see Steinman et al., 1982, for details of this relationship).*

The difficulties inherent in making valid measurements of oculomotor compensation emerged quite quickly when the first measurements of fixation were made with the head free from artificial restraints (Skavenski et al., 1979). These authors found it necessary to build Helmholtz coils 2 meters on a side (3 times larger than customary), locate the 2 cm region of relatively homogeneous magnetic flux within these large coils, and provide a framework to confine the subject's free head within this small homogeneous region while recordings were made. After all these modifications of the Robinson magnetic field sensor coil technique, it was only possible to study oculomotor compensation while the subject maintained fixation on a target at optical infinity while he sat or stood *as still as possible.* Natural oculomotor compensation, even in this most limiting condition, was far from perfect, as is illustrated in Fig. 18. The recordings reproduced in Fig. 18 show eye and head position relative to an earth-fixed coordinate system. If the head or the eye did not change orientation with respect to this coordinate system, the position traces would not change position in these records. Note that the horizontal and vertical head traces did not change position; they produced horizontal straight lines when the head was supported by a biteboard. The horizontal and vertical eye traces were also relatively straight and only varied moderately in this condition. The variations which can be seen in the position of the eye traces are produced by the fixational eye movements described and discussed throughout the earlier parts of this chapter.

The story was quite different when the head was removed from the biteboard and the subject tried to sit or to stand as still as possible. The head moved a great deal on both meridians. The eye traces moved less than the head but they moved more than the head when it was supported by a biteboard. This shows that there is oculomotor compensation for small, irrepressible head movements. If there was no compensation whatsoever, the eye and head traces would move together, in the same direction and by the same amount. Compensation was not complete. The eye moved when the head was free. Complete oculomotor compensation would have produced eye traces that looked like traces of a head supported by a biteboard. The eye movement records reproduced in this figure may be interpreted so as to represent motions of the fixation target on the subject's retina as readily as they can be interpreted so as to represent movements of the eye with respect to earth-fixed coordinates. This important fact was established by Ferman et al. (1987), who discussed and examined empirically all known and

---

* We have just begun to study natural compensatory eye movements properly, i.e., so as to infer retinal image motion accurately from recordings of eye position when the head is free and the targets are nearby. Such inferences require techniques for measuring head translations to about 0.1 mm. It is not impossible to make these measurements but it is not a trivial undertaking and it must still be done. See Steinman (1986b) for a description of the development of the revolving magnetic field–silicone annulus sensor coil technique currently being used to measure eye and head orientation in free-headed human subjects, viewing distant targets, and Collewijn et al. (1990a) for the first accurate measurements with the head free and the target nearby.

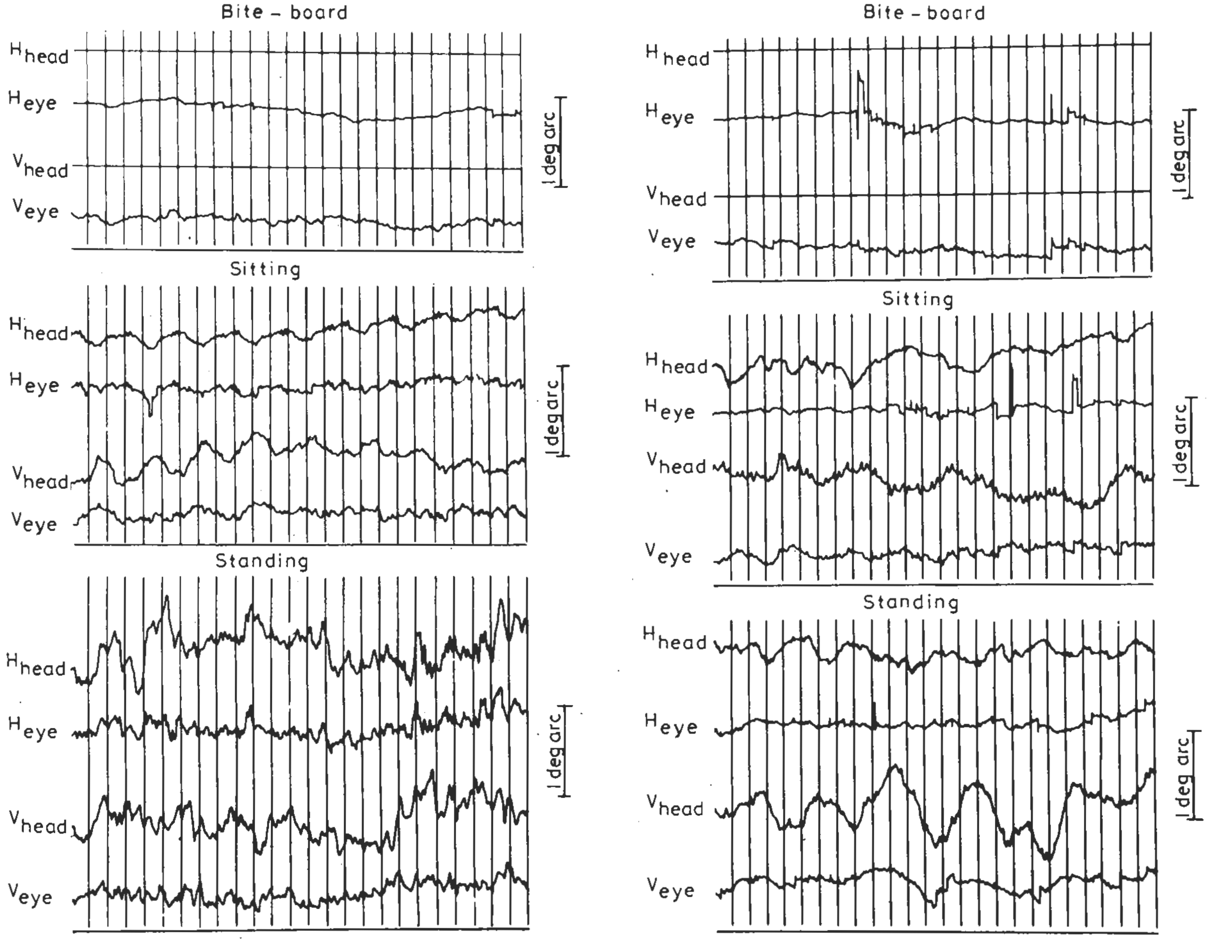

Fig. 18. Representative horizontal (H) and vertical (V) head position and gaze (retinal image position) of subject AAS (left) and subject RMS (right) while they fixated a target at optical infinity with their heads supported on a bite-board, or while they were sitting or standing as still as possible. Records begin on the left. The vertical time-lines show 1-s intervals and the vertical scales on the right side of each record show 1 degree of visual angle. Upward changes in these position traces signify rightward or upward rotations. (From Skavenski et al., 1979)

suggested potential artifacts in such recording methods. Before the Ferman et al. report, this claim was questioned by some (Duwaer, 1982; Stark, 1983).

Quantitative analyses of the retinal image motion observed in this experiment showed that motion increased, over measures obtained with the head supported on a biteboard, by factors of 2–4 when the head was free and retinal image stability depended entirely on compensatory eye movements. Compensatory eye movements corrected only a modest portion (on average, about half) of these irrepressible natural head movements. Skavenski et al. (1979) also measured the VOR to very small amplitude (0.5° and less) passive sinusoidal oscillations (0.1–15 Hz) and reported that compensation was better than when the subject remained as still as possible but still far from perfect (at best about 75%). Taken together, these results encouraged Skavenski et al. (1979) to conclude that "the degree of compensatory oculomotor response is actively adjusted downwards so as to guarantee sufficient retinal image motion to prevent perceptual fading when the body is relatively stationary and is actively adjusted upwards, so as to guarantee sufficient retinal stability to prevent blurring when the body moves actively. Seen this way, the goal of oculomotor compensation is not retinal image stabilization, but rather controlled retinal image motion adjusted so as to be optimal for visual processing over the full range of natural motions of the body" (p. 675). But what is optimal? This had actually been known, at least with respect to the optimal lower limit, since Riggs et al.'s (1953) influential early study of the effects of image stabilization on vision. In 1979, however, the fact that it was known had not yet been appreciated (see section 6.2.1).

Ten years ago, eye movement data obtained with the head stabilized on a biteboard were treated as if they represented an ecologically 'normal' oculomotor activity, that is, a behavior honed by evolutionary pressures to some value that would be optimal for vision. There was, however, long-standing, often replicated, evidence from stabilized image research that should have raised concern about this assumption. Before the Skavenski et al. (1979) report, the average 'normal' fixational drift speed on a single meridian was taken to be about 5′/s, the observed average value of intersaccadic drift (difference in eye position between the end of one saccade and its position at the beginning of the next saccade). Intra- and intersubject fixational saccade rates of the experimenters, who served as the main subjects in these contact lens optical lever experiments, ranged from 2/s to 0.5/s and lower, which means that that the estimate of 'normal' 5′/s drift speed was based on long sampling intervals, 500–2000 ms or more (5–20 times the length of the 'critical duration' for visual processing, i.e., the interval during which the intensity and duration of stimulation can substitute completely for one another). Estimates of fixational drift speeds based on somewhat more appropriate, shorter, sampling intervals (200 ms) gave slightly higher values (e.g., 6.5 to 9′/s in Nachmias, 1959, 1961; Steinman, 1965). Skavenski et al. (1979), with the still shorter sampling interval of 50 ms, obtained average drift speeds of about 14′/s with the head on a biteboard and 2–4 times that value when the head was free.

But which among these values was the 'normal' retinal image speed with respect to optimizing visual processing? Here, we can turn to results of stabilized image research. Recall the three conditions in the Riggs et al. (1953) experiment in which the effect of exposure duration on visual acuity was examined (section 6.2.1). Tests were made under normal, stabilized and 'exaggerated' image motion viewing conditions. They found that once exposures were longer than 200 ms, exaggerated image motion (twice 'normal' in their experiment) was clearly optimal for vision. 'Normal' fixational drift eye movements (i.e., drifts with the head on a biteboard) were *too slow*. Krauskopf (1957) made similar observations. He imposed oscillatory motion on a stabilized acuity target and also reported that higher than 'normal' retinal image motion was better than 'normal' motion, but, like Riggs and his co-workers, Krauskopf did not discuss the implications of this finding for natural visual processing (see section 6.2.2). The beneficial effect of higher than 'normal' fixational drift-like eye movements, imposed during stabilized viewing, permeates the stabilized image literature from 1953 until 1978, but it was only after Skavenski et al. (1979) had measured fixational eye movements under natural conditions that their significance became apparent. (See Kowler and Steinman, 1980, for a review of the history of this problem and a discussion of the visual significance of the new notion of 'normality'.) At present it seems most appropriate to infer 'normal' fixational eye speed for both the visual and oculomotor systems from the relatively high retinal image speed observed during natural oculomotor compensation, when the head is free, rather than to continue the older practice of inferring 'normality' from the relatively low speeds observed when the head is immobilized on a biteboard. This speed, observed with the head off the biteboard while a subject tries to sit as still as possible, is the *lowest* natural limit. It provides a reasonable estimate of the least amount of retinal image motion ordinarily available to provide transient stimulation to visual neurons. There is no deficiency of retinal image motion once the head is free despite intentions to remain as still as possible. But human beings rarely sit as still as possible. Most useful activity is accompanied by appreciable movement of the head and body. How effective is oculomotor compensation under these conditions?

8.2. Vision in the presence of natural retinal image motion

8.2.1. The extent of retinal image motion during active head movement

The first accurate measurements of this kind were made by Steinman and Collewijn (1978, 1980). They were made binocularly – a serendipitous undertaking which had unforeseen consequences for our understanding of binocular visual processing because each eye performed quite differently, an outcome that could not have been anticipated from the literature available a decade before. The significance of head movement for binocular oculomotor and visual performance is described by Collewijn and Erkelens (this volume) and will be mentioned from time to time only briefly in this chapter.* We will emphasize retinal image motion and its visual consequences only for a single viewing eye. Steinman and Collewijn (1980) asked the subjects (themselves and two co-workers in the Rotterdam laboratory) to maintain fixation on a distant object (5 or 35 kilometers away) while oscillating the head about its vertical axis at progressively increasing frequency; essentially the same thing the reader was asked to do as his last observation in section 8.1. There were only two differences. First, the reader maintained fixation on the nearby text while the subject in Rotterdam looked at a distant target. Second, both the reader and the subject in Rotterdam were asked to notice what he saw while the head oscillated, but the subject in Rotterdam wore binocular silicone annulus sensor coils (Collewijn et al., 1975) and his binocular, horizontal eye and head movements were recorded by means of a revolving magnetic field sensor coil monitor (Collewijn, 1977). The main results obtained in Rotterdam are illustrated in Fig. 19.

The four rows of position vs. time graphs show representative performance of each of the four subjects studied. The columns show head frequency, increasing from left to right. The data were actually collected during a continuous 30-s trial in which the subject started oscillating his head very slowly (leftmost graphs), continued to increase head frequency, reducing amplitude as he did so (the middle graphs), finally achieving his maximum head frequency with very much reduced amplitude, which is shown in the performance at the extreme right. These data were grouped into 3 separate graphs merely for convenience of plotting. These recordings show head and eye positions with respect to an earth-fixed coordinate system, which means that the significance of these records is the same as the significance of the records shown in Fig. 18 taken from Skavenski et al. (1979). There are only three differences: (1) the head movements in Fig. 19 are scaled down to 1/10th of their value because the subject moved his head through relatively large angles rather than kept it as still as possible, (2) recordings were only made of movements along the horizontal meridian and (3) binocular, rather than monocular, eye movements were recorded.

Virtually perfect compensation would cause the eye traces to approximate horizontal straight lines. Such lines were seen from time to time in *one* of the subjects's eyes. Note, for example, seconds 5 to 10 of subject RS's right eye in the record on the left and the first 4 seconds of subject LK's left eye, also on the left. In these instances, and in similar instances

---

* Discussions of binocular eye movements and their consequences for stereopsis, stereoacuity, fusion and 'hysteresis' can also be found in Collewijn et al., (1990b). A discussion of the precision of monocular and binocular gaze with a free head and its potential significance for vision can be found in Steinman et al. (1982). In this review, it was argued that natural failures of oculomotor compensation have more significance for binocular than for monocular vision, which had already been shown to tolerate retinal image motions up to about 2°/s (Westheimer and McKee, 1975; Murphy, 1978; see section 7.2.8).

in subjects studied subsequently, the compensation in the companion eye was always far short of perfection. This guaranteed considerable variation of the retinal position of the target image in at least one of the eyes and also guaranteed considerable variation of the absolute disparity between the retinal images present in each of the eyes. In other words, the oculomotor vergence response was not stable. The retinal image of the fixated distant target in one or the other or both eyes moved while the subject maintained fixation, binocularly, on an object whose distance and direction relative to the subject were not changing. Retinal image motion within each eye arose from both under- and overcompensation of head rotations. In these records (as in Fig. 18), undercompensation occurred when the eye trace moved in the same direction as the head trace and overcompensation occurred when the eye trace moved in the opposite direction. The reader should inspect the eye traces in Fig. 19 closely and note the degree of complexity and inconsistency over time of the binocular retinal image motion patterns that subjects tolerate, or perhaps even prefer, when they move their heads while using both eyes. The inter- and intrasubject variation of oculomotor compensatory activity, illustrated in Fig. 19, is characteristic of all subjects studied thus far. Subjects, despite all of these perturbations of the positions of their retinal images, reported 'normal vision', by which they meant that the perceived visual world was unitary and that fine details within it remained clear and stationary. It was only at the highest possible frequency of head movement that some jitter and slight degradation of fine detail was noticed. Try it yourself now that you know what is likely to be going on on your retinas. Use a distant target because it eliminates the need to consider translations of your head (see footnote on p. 177).

8.2.2. Residual image motion is the goal of oculomotor compensation

Collewijn et al. (1981, 1983) showed that the compensatory subsystems *prefer* appreciable retinal image slip by requiring the oculomotor system to adapt to novel optical arrangements (magnifying or minifying spectacles) that changed the amount of eye rotation required to compensate for a head rotation of a given size. "These adaptation experiments were undertaken to determine whether the observed departures from virtually perfect compensation arose from limitations inherent in the compensatory subsystems or from the desire of the compensatory subsystems to maintain retinal image motion at some nonzero value that might be optimal for vision" (Collewijn et al., 1981, p. 312). This second possibility had been raised as was noted above (section 8.1) by Skavenski et al. (1979) when they reported the first accurate measurements of natural retinal image motion, that is, image motion measured with the head free from artificial supports (see section 8.1). Collewijn et al. (1981) showed that the adapted compensatory response reestablished the deviation from perfect compensation that was characteristic of the individual subject when either no or normal minifying spectacles were worn in the case of myopic subjects. This occurred despite the fact that the optical arrangements required the compensatory eye movements to encounter a state in which the retinal image had zero slip during head oscillation. In the words of the authors, "it is important to realize that when we pushed the compensatory subsystems from their natural low gain to an unnatural high gain, gain moved through values that would have allowed virtually perfect stabilization of the retinal image. Had virtually perfect stability been the goal of the compensatory subsystems, they should have stopped adapting at this time. They did not. This result permits us to conclude that the compensatory subsystems seek some appreciable nonzero retinal image speed rather than virtually perfect image stability (p. 327)... We must now study individual visual capacities with known gaze velocities [retinal image slips] and show that the gaze velocities preferred by an individual are optimal for that individual's visual requirements" (pp. 328–329).

8.2.3. Visual psychophysics with known natural retinal image motion

The first actual psychophysical measurements,

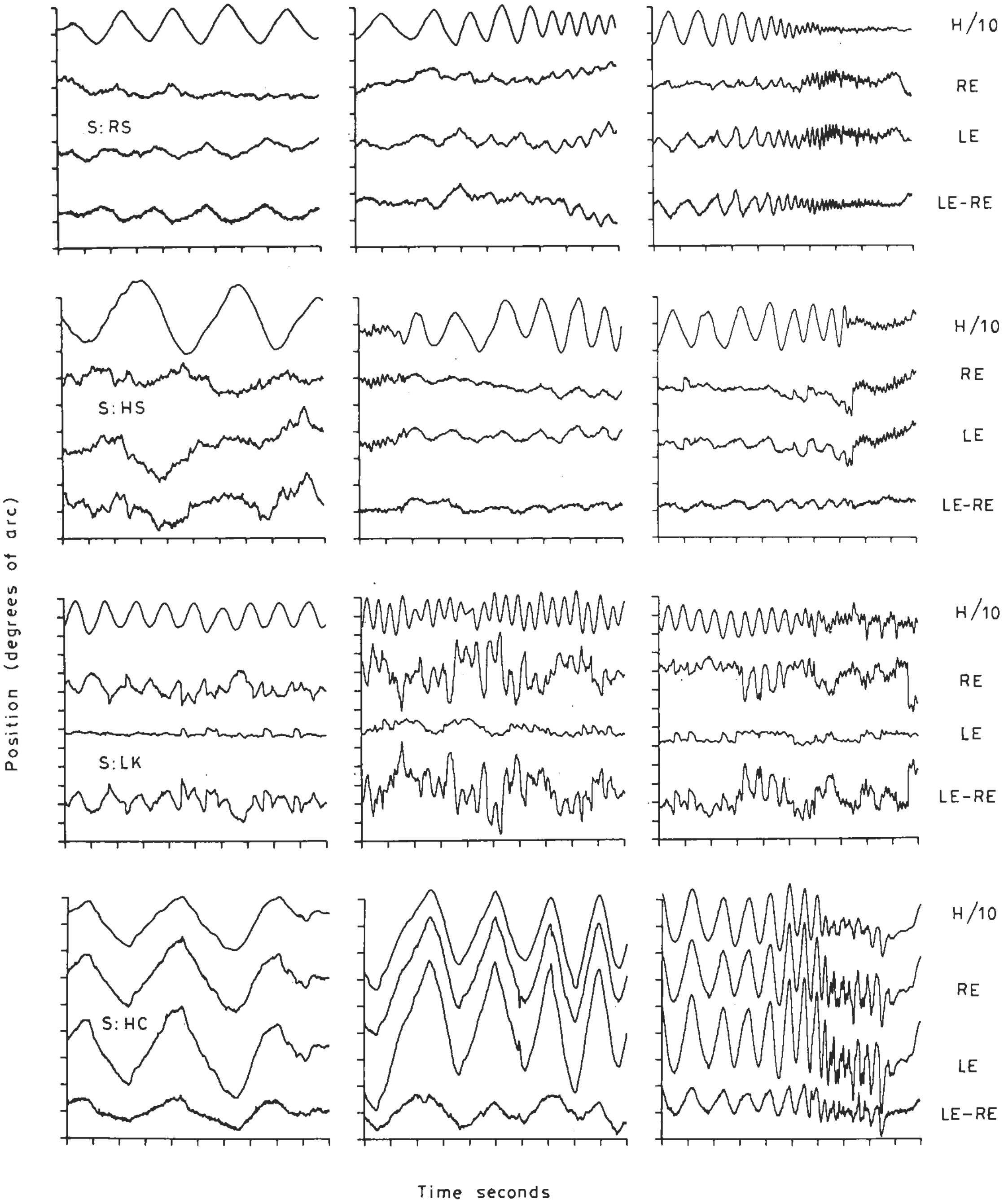

Fig. 19. Representative recordings of horizontal head (H) and binocular gaze (retinal image position) of four subjects while they fixated a distant target and oscillated their heads about the vertical axis. Each of the 12-s records begins on the left, and tick-marks on the abscissa indicate 1-s intervals. Tick-marks on the ordinate indicate 1 degree of visual angle. The movements of the head have been scaled to 1/10 of their actual value and the trace LE–RE signifies vergence eye movements, with convergence shown by upward changes in position. Upward changes of the right eye (RE), left eye (LE) and scaled head traces (H/10) signify movements to the right. (From Steinman and Collewijn, 1980)

rather than informal introspective descriptions, of stereoacuity and contrast thresholds in the presence of natural retinal image motion produced by head movements of the sort studied by Steinman and Collewijn (1980) (section 8.2.1) were reported by Steinman et al. (1983, 1985). These reports also included demonstrations which showed that random dot stereograms could be fused during violent head movement and also that fusion persisted despite heroic efforts to break fusion by shaking the head. Here, we will discuss only the work on contrast sensitivity, considered to be based primarily on monocular processes – binocular contributions to the detection of contrast are believed to arise exclusively through probability summation, essentially the parallel processing of two independent monocular processes. (The reader should consult Collewijn and Erkelens, this volume, for a discussion of the role of eye movement in unique phenomena such as fusion, stereopsis and stereoacuity, and Collewijn et al. 1990b, for a treatment of binocular processes when the head is free from artificial restraint.)

These experiments (Steinman et al., 1983, 1985) rested on previous information about the retinal image motion of three very experienced, eye movement subjects (namely, Collewijn, Kowler and Steinman), who had served in the adaptation experiments described just above. Their monocular contrast sensitivity was measured under three conditions: (1) with the head supported by a biteboard, (2) while they oscillated the head about its vertical axis through an angle of about 34° at 0.33 Hz, or (3) oscillated the head through the same angle at 1.33 Hz. Timing was paced by a metronome. Such oscillations were known to produce retinal image speeds that ranged from about ⅔ to 3°/s in these subjects under these conditions. Subjects were instructed to make their contrast judgements (determine whether they could they see the grating pattern) near the center of each swing of the head when retinal image speed would be at its highest value, essentially the same instruction Arend had used when he used smooth pursuit to move a grating across the retina (see section 7.2.8). A variant of the Method of Adjustment was used in which the subject indicated verbally whether contrast should be increased or reduced, continuing to indicate the change that should be made, until satisfied that contrast was set to its threshold value. This variant of the Method of Adjustment was used because the subject was too busy to twiddle a potentiometer knob and at the same time: (1) keep time with the metronome, (2) keep the size of head movements within the desired range and (3) make judgements about the appearance of the display near the center of each swing of the head.

Here, the reader surely realizes that the subject was required to do something akin to what the reader was asked to do in section 8.1, differing only in that: (1) the experimental display contained a sinusoidal spatial frequency grating which was far away (about 6 m), rather than the text in this book which was held at arm's length, and (2) the frequency of head movement was kept uniform in the experiment while contrast was varied until threshold was reached for the particular spatial frequency under test, rather than keeping contrast at the same, relatively high level (black letters on a white page) and varying the frequency of the head movement from very slow to as fast as possible while the reader judged the clarity of the text in a complex display, containing a wide range of spatial frequencies. The contrast sensitivity functions obtained for each of the subjects are reproduced in Fig. 20.

Steinman et al. (1985) reported that "the results for all three subjects were qualitatively similar. Head movement, with its concomitant retinal image motion, produced a need for more contrast at high spatial frequencies and reduced the need for contrast at low spatial frequencies. The crossover in the functions (where moving the head causes high-frequency attenuation and low-frequency enhancement of contrast sensitivity relative to the function obtained with the head on a chin rest) occurred at about 10 cycles/degree for subjects EK and RS and at about 6 cycles/degree for subject HC. Note, however, that the deleterious effects of image motion on high spatial frequencies were modest for all three subjects, the differences in all cases being less than a factor of 2 of contrast. Also note that the extrapolated high frequency cutoffs for each of the subjects would be well above 30 cycles/degree" (Steinman et al., 1985, pp. 227–228). These results are illustrated in Fig. 20, as is the average of these results (geometric means), which is also compared with the effects of retinal image motion on contrast sensitivity modeled by Kelly (1979b) for the same retinal image speeds. (Kelly's experiments will be described in detail in section 9.) Kelly's model was derived from threshold data, obtained with his eye, when constant-velocity unidirectional motion was imposed on a sinusoidal, relatively low spatial frequency display (<13 cycles/°), which had been stabilized on his retina by means of a *Double Purkinje Image Tracker, Stage III.* This, then-novel, non-invasive eye position monitor will be described and evaluated after the results of the naturally imposed and artificially imposed retinal image motion on contrast sensitivity have been compared. In other words, after effects of slip of the retinal image of the display during head movement have been compared with the effects of drifting a stabilized display.

There were several differences. Kelly's extrapolated high-frequency cutoffs would lie between about 9 and 18 cycles/°, whereas the Steinman and coworkers's extrapolated high-frequency cutoffs all fall above 40 cycles/°. Furthermore, Kelly's crossover points occur below 3 cycles/°. Steinman and coworkers's crossover points were above 6 cycles/°. Both of these results differ by at least a factor of 2, which led Steinman and his coworkers to conclude that the differences had potential theoretical implications.

In their words, "there are a number of important differences between our experiments that could contribute to the difference in results. Kelly imposed constant velocity displacements of the gratings – his stimulus moved continually in only one direction. Kelly's relatively low contrast sensitivity at high spatial frequencies and his relatively low crossover frequencies may reflect the effects of retinal velocity adaptation caused by continually moving the stimulus in the same direction. In other words, his constant-velocity technique could prevent normal processing by the visual system. Our free-head movements were periodic, resulting in retinal image motions of about the same frequency as the head. These periodic oscillations are the normal inputs to the visual system, which, even with the head restrained on a biteboard, occur, more or less sinusoidally, at frequencies predominantly in the range 2 to 5 Hz. Even such small-amplitude (<10′ peak-to-peak) oscillations are sufficient to prevent fading of targets located in the central fovea. Increasing the oscillation amplitudes by head movements improves contrast sensitivity at low spatial frequencies. But, of course, large-amplitude oscillations cannot but degrade the visibility of high spatial frequency gratings. The loss in acuity that we have found for image motion obtained with oscillatory head movements should be compared with the loss found with comparable oscillatory image motion imposed on a stabilized display" (p. 228).

Steinman et al. (1985), after discounting differences in the light levels and viewing conditions (binocular vs. monocular) in their and Kelly's experiments, suggested that image motion might be processed differently when it is associated with head movement than when motion is imposed while the head is immobilized. They speculated that the vestibular signal might provide the basis for a visual neural remapping that keeps track of changes of the position of details contained in the retinal image caused by the normal insufficiencies of the compensatory eye movements. The subject's visual system 'knows' about the motions of the head from the vestibular signal, already accepted as providing the oculomotor system with the information it requires to control compensatory eye movements. If the visual system also 'knew' the individual's preferred degree of retinal image slip (the 2–8% normal individuals allow reliably, albeit idiosyncratically), the visual system would be in a position to make allowance for the retinal slip expected to be associated with a head rotation of a given size. In effect, slip in the neural message would be less when the slip results from normal characteristics of com-

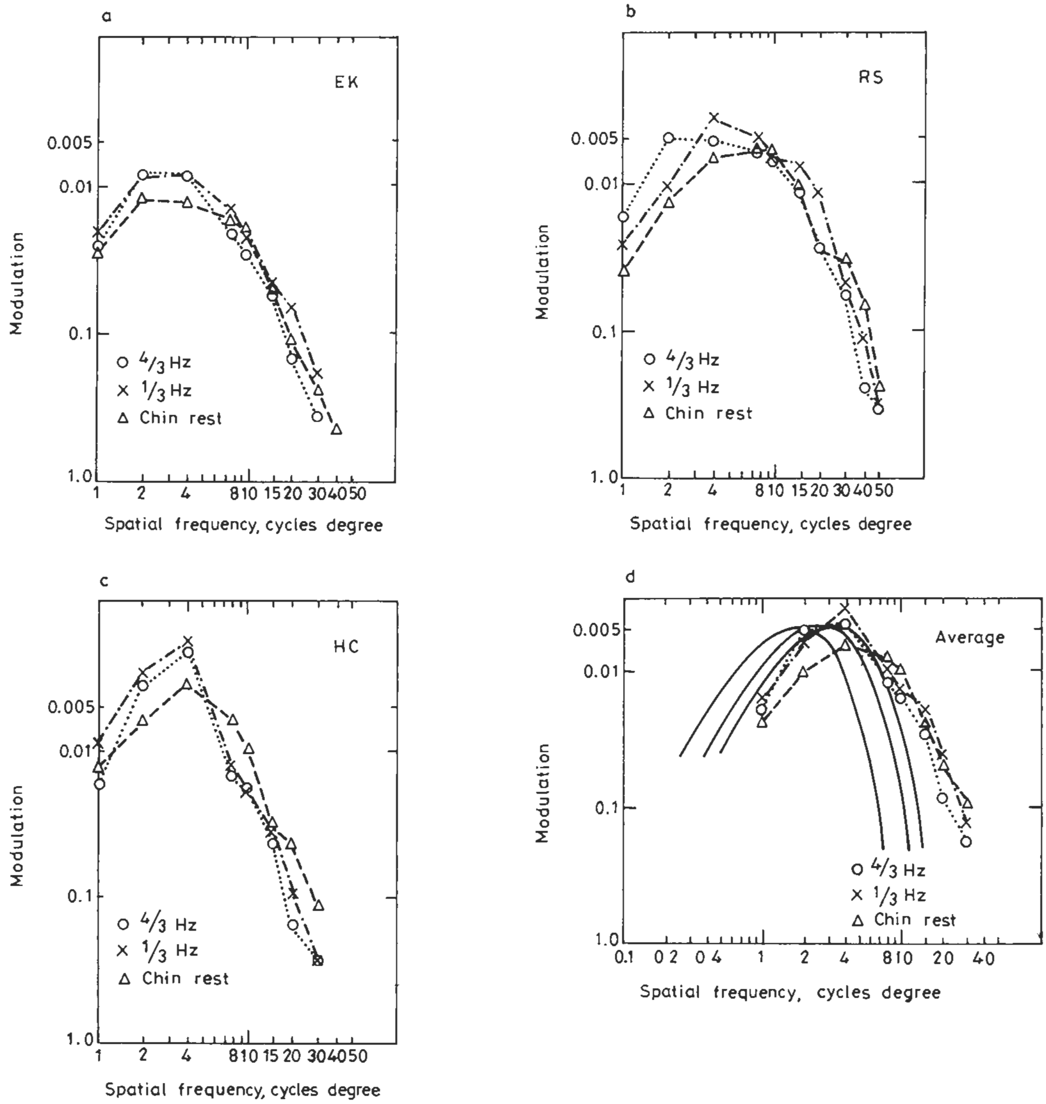

Fig. 20. Threshold grating contrast-modulation settings as fractions of 100% contrast at various spatial frequencies. Three frequencies of head oscillations about its vertical axis were used, namely, near 0 with the head supported by a chin-rest or oscillating at 1/3 and 4/3 Hz. (a) shows the CSFs obtained under these conditions for subject EK, (b) for subject HC and (c) for subject RS. (d) shows their average CSF (the geometric mean) as broken lines. Values and extrapolations from Kelly's model of the spatio-temporal transfer surface are shown as solid lines. These hypothetical functions are based on unidirectional motion of similar speed imposed on a grating 'stabilized' with a Stage III SRI eyetracker. See the text for a discussion of these functions. (From Steinman et al., 1985)

pensatory eye movements than when slip is generated by moving the visual stimulus while the head is immobilized. If this were to be the case, the well-known detrimental effects of image motion on high spatial frequency components of visual displays would be less when the slip is associated with concomitant head movements than when vestibular signals produced by head acceleration are not available. If additional principles do not come into play, the proposed mechanism for tolerating head-produced retinal image slip might, however, reduce the well-known visibility improvement for low spatial frequencies – an effect that was demonstrated originally when the head is stationary and the display is moved.

The reader should recognize that this is an elaboration of the 'comparator' function Arend (1973) included in his model described in section 7.2.4. Arend assumed there was such a comparator because the human being can distinguish between changes in retinal image position produced by rotations of his eye and changes in position produced by movements of external objects. Steinman et al. (1985) are proposing that vestibular information may have particular significance for the processing of retinal image motion under natural conditions. This change in emphasis is a natural outgrowth from the demonstration that virtual perfection of vision cannot be accounted for by the virtual perfection of oculomotor performance. There is, however, an interesting difference between the conclusions of Steinman et al. (1985) and the model of Arend (1973). Namely, Arend distinguished between the effects of image motion produced by moving the object and the effects of image motion produced by moving the eye. But Murphy (1978) had already shown that effects of image motion produced either by moving the object (stimulus) or by smooth pursuit were equivalent.

It was possible at this point in time to believe that the differences between the effects of retinal image motion reported by Kelly (1979b) and by Steinman et al. (1983, 1985) most probably arose from differences in their experimental procedures. Kelly (1979b) imposed an unnatural pattern of retinal image motion on his stabilized spatial frequency display, that is, a constant-velocity unidirectional drift, whereas Steinman et al. (1983, 1985) used the natural retinal image motion arising from imperfect oculomotor compensation associated with oscillations of the head. These natural, rather sinusoidal, retinal image motions consisted of relatively low frequency oscillations (most below 5 Hz). This experiment involved a potentially important artifact, notwithstanding the advantage it gained by using a natural pattern of retinal image motion: the subject might have ignored the instruction to make threshold judgements only near the middle of his head oscillation and actually made judgements when retinal image speed was low shortly before or shortly after the head changed direction at the end of each swing. Judgements made near turnabouts could produce very modest differences between thresholds obtained with an oscillating, as compared with a stationary, head at the high spatial frequency end of the CSF. An explanation along these lines would hardly be flattering to the dedication and competence of the three highly experienced subject-experimenters who served in the Steinman et al. experiment, to say nothing of its potential implications for their integrity. This possibility is being raised here not because we consider it to be a likely explanation of the failure to find large effects of natural retinal image motion on the contrast thresholds for high spatial frequencies but, rather, because such concerns had been expressed, and even supported experimentally, in the literature. Arend (1976) engaged this issue when he produced retinal image motion by smoothly pursuing a point that moved back and forth across a stationary spatial frequency display. He reported that he could measure reliable thresholds and that they were very much higher when he explicitly tried to make his judgements while the moving point was near the center of the display than when he tried to make his judgements near turnabouts, where his eye should have been moving more slowly (see Fig. 15 and section 7.2.7).

An artifact like this can explain how the bars in a relatively low contrast *high* spatial frequency display can be seen while the head moves but it muddies up simple interpretations of the observed beneficial effects of retinal image motion on thresholds for *low* spatial frequencies. Recall, Steinman et al. (1983, 1985) reported crossovers of the moving and stationary CSFs at 6–10 cycles/degree, while Kelly (1979b) reported crossovers at 2–3 cycles/degree. This means that the subject in the Steinman et al. experiment had to make threshold judgements, reliably, at two different parts of the head trajectory. He had to shift his threshold judgement interval from the ends to the middle of the head trajectory when the test grating was changed from a high to a low spatial frequency. Such a dual observational strategy would permit: (1) lower thresholds with low spatial frequency displays when the head moved than when the head was stationary and (2) similar thresholds with high spatial frequency displays when the head moved and when the head was stationary. The reader can get some feel for the ease with which this kind of dual observational strategy can be adopted and maintained by oscillating his head while fixating at the center of Fig. 16 (Arend's demonstration plate, which contains both a high and a low spatial frequency). Note the appearance of the low spatial frequency near the center of your head trajectory and also the appearance of the high spatial frequency when you change the direction of your head movement. The reader should also consider whether a subject using this dual observational strategy would be likely to be using these two different observational intervals, both reliably and inadvertently. In other words, try to form an opinion, based entirely on the subjective impression you form while engaged in this task, about the likelihood that a subject might consistently shift his observational interval without realizing that he was doing so. This may allow you to anticipate the outcome of an experiment (described in section 10) in which a subject making contrast threshold judgements while moving his head was prevented from using this kind of dual observational strategy.

In summary, Kelly's results are suspect because he used an unnatural pattern of retinal image motion, while Steinman and his colleagues's results are suspect because, despite the fact that they used a natural pattern of retinal image motion, they turned control of the experiment over to the subject, who might have been careless or cheated while making his observations. These issues are not completely resolved but some progress has been made. The final two sections of this chapter review this progress. The first will review both technical and theoretical developments underlying modern image-stabilization work.

9. Stabilized image research without attachments to the eye

9.1. A new technique for retinal image stabilization

In section 6, the development of techniques for stabilizing retinal images was described, from their introduction by Ditchburn, Riggs and their coworkers in the 1950s to their use in the hands of Gerrits and his coworkers, well into the late 1970s. All of these techniques were 'invasive', that is, they required attachments to the eye by means of either a tightly fitted scleral contact lens or a Yarbus-type sucker. It was pointed out in section 6.1.1 that all such techniques placed serious time-constraints on the investigators, constraints that precluded psychophysical methods for estimation of thresholds, or at least compromised the manner in which they could be used. An appreciation of the importance of this limitation encouraged Cornsweet and Crane to develop a stabilizing instrument which did not require attachments to the eye. This kind of instrument would make it possible to run long psychophysical sessions with the position (and therefore motion) of the retinal image of the test stimuli completely under the control of the experimenter. Their stabilization technique made use of what has come to be called the SRI Double Purkinje Image Tracker.

The Tracker has gone through five 'generations' or 'Stages' since its development began in 1967. It became a useful research instrument in its Stage III configuration after about 10 years of development. The subsequent two Stages concentrated on changes intended mainly to make it easier for inexperienced personnel to use the instrument. The development of the SRI tracker is described in two instrumentation publications (Cornsweet and Crane, 1973; Crane and Steele, 1978), neither of which provides detailed *quantitative* treatments of such important features as its frequency response, cross-talk between horizontal and vertical channels or the noise spectrum of its voltage output. Both of these publications do, however, describe quite clearly the problems to be overcome in the development of an instrument of this kind and the potentially effective solutions to these problems that have been incorporated into the instrument at its several stages of development. These publications, however, only provide selected eye movement records to illustrate the useful properties of the instrument. The reproducibility of these properties, within and between subjects, is not addressed quantitatively. The properties illustrated can be difficult to approximate in routine practice. The instrument can be used, but its performance and susceptibility to a variety of artifacts depend to no small degree on the characteristics of each subject's eye and the subject's ability to minimize head movement (even when the head is supported by a biteboard). Appreciation of these limitations, discussed in detail in the following sections, depends on the instrumentation skills and knowledge of the normal eye movement pattern that the experimenter brings to his evaluation of the SRI tracker from work with other 'state of the art' eye position monitors (viz., contact lens optical levers or the silicone sensor coil stationary or revolving magnetic field techniques).

9.2. Basis of our opinion of the SRI tracker

The sections which follow contain evaluative comments that call attention to technical limitations, which may be considered controversial by some Eyetracker users. Our comments are based on our personal experiences, and the personal experiences of our collaborators and acquaintances who have worked with Stage III or later instruments. In most cases, problems in using the instrument and its inherent limitations have not been emphasized in published papers. In our view, it is desirable to emphasize these problems here because inferences based on stabilization experiments done with the SRI tracker, as well as inferences based on all other available stabilization results, depend critically on the technical limitations of the device used to track movements of the eye. We believe that the problem of reliably stabilizing retinal images almost completely for periods of more than a very few seconds, without attachments to the eye and in the presence of natural behaviors (i.e., blinks, saccades and inescapable, albeit modest, changes of head position on a biteboard), has yet to be solved. **N.B.**, we are not proposing that the tracker cannot be useful in oculomotor or stabilized image research. Rather, we are calling attention to the inherent limitations in the use of the eyetracker as a stabilizing tool, and to the possibility that recent versions of the instrument, which have made it relatively simple to use, have risked increasing the likelihood of confusing artifacts with valid observations. We do not dispute that the SRI Double Purkinje Tracker, beginning with Stage III, has been used effectively in a number of publications, concerned both with eye movements and with stabilized retinal images. As noted earlier in section 6.2.10, the inability to obtain or to maintain perfect stabilization does not preclude drawing inferences about the effects of retinal image motion from observations made with a less than perfect instrument.

One central problem is that the validity of these observations rests primarily on the committed, highly experienced, subject's ability to separate observations that count from observations that must be discounted. This fact means that in SRI tracker stabilization research, just as in all previous stabilization research, experienced stabilized-image investigators must serve as subjects if relatively reliable results are to be obtained. In short, technology has not yet overtaken art in stabilized image research. We believe that the reader will find this view implicit in our quotations from the publications of Kelly – the SRI tracker's main champion and primary user for stabilized image research.

9.3. Retinal image stabilization with a Stage III SRI tracker

In this section, we will first describe the principle of operation of the SRI tracker and then describe an experiment (Kelly, 1979a) in which its relative effectiveness for image stabilization was reported. This section will show that inherent limitations make the Eyetracker useless for measuring sensitivity to high spatial frequency stabilized gratings even by highly sophisticated and committed subjects. This section will be followed by a discussion of recent theoretical and experimental work by Arend and Timberlake (1986), who claim that it may never be possible to stabilize images completely enough to drive 'on' ('sustained') neural units exclusively in a living eye (see sections 5.1, 6.2.4 and 7.2.6).

The SRI tracker measures changes in the distance between the first Purkinje image (the light reflected from the front surface of the cornea) and the fourth Purkinje image (the light reflected from the concave surface of the back of the crystalline lens). When a collimated beam of near infrared light is incident on the cornea, the convex surface of the cornea forms a virtual image (the first Purkinje) of its source within the eye. The location of this virtual image is nearly coincident with a real image (the fourth Purkinje) of this source formed by the concave posterior surface of the crystalline lens. Each of the reflecting surfaces forming the first and fourth Purkinje images is at a different distance from the center of rotation of the eye, which causes the distance between them to change as the eye rotates. "For small angles, this distance is a linear function of the rotation of the eye" (Kelly, 1979a, p. 1267). If the eye translates rather than rotates, the distance between the first and fourth images does not change. The degree, size and locations of the regions of approximate linearity of eye rotation indications depend on the anatomical structures of each subject's eye – these features vary within, as well as between, subjects. Calibration factors, within the same subject, can vary by about 10% between sessions. Some eyes are not suitable for use with this instrument – they cannot be brought into or maintained in a useful alignment for a variety of reasons, for example, very faint fourth images or subclinical cataracts.

The main virtue of the Eyetracker, and its great superiority over other non-invasive instruments, which use only the first Purkinje image or the amount of light diffusely reflected from the limbus, is its potential for separating signals arising from head translations from signals arising from eye rotations. As indicated above, in the SRI tracker, the distance between the first and fourth Purkinje images does not, in principle, change when the head or eye translates. It changes only when the eye rotates. In practice, the distance between the first and fourth images can change quite a bit whenever the tracker makes a focus adjustment, made necessary by a head movement (head rotations and translations cannot be completely eliminated by even the best combined biteboard and forehead supports).* The SRI Double Purkinje Tracker should make it possible to run large numbers of inexperienced subjects in lengthy psychophysical experiments with stabilized images because stabilization is accomplished without attachments to the eye. This eliminates the need for attaching tight-fitting, scleral contact lenses, suckers, or silicone annuli (described later) to the eye and for instilling topical anesthetics – procedures closely scrutinized by institutional committees for the protection of human subjects. In practice, however, considerable skill, cooperation and knowledge are required of the subject in stabilized image experiments and the investigators have continued to serve as the primary subjects for this kind of research. In the experiment to be described next (Kelly, 1979a), the subjects

---

* This is a significant problem in Stage IV and V instruments, which were designed to adjust focus of the first and fourth images as head position varied over a relatively large range – an 'improvement' in the direction of user-friendliness found to be less than completely successful by users who had made their own tests of tracker performance. Some such users have circumvented this problem by discounting data obtained during experimental intervals in which an appreciable change in focus was made by the tracker (e.g., 1 mm, personal communication A.A. Skavenski).

were L.A. Riggs, U. Tulunay-Keesey (see sections 6.1–6.2), and two SRI researchers, M. Clark and D. Kelly. All were highly experienced stabilized image researchers and would, therefore, be accustomed to selecting meaningful results, i.e., what they felt were steady-state thresholds in the presence of unavoidable artifacts, specifically, threshold changes produced by destabilization associated with blinks, saccades and head movements, and, as we shall see in the next section, even heartbeats.

Highlights in the SRI tracker stabilized-image literature will be illustrated by means of quotations from the publications of Kelly – its primary user. This is the same approach we took when the results of 'invasive' stabilization experiments were described in sections 6.2.4., 6.2.5 and 6.2.8–10. Quotations will also, as in previous sections, contain interleaved and footnoted evaluative comments. We begin with Kelly's description of how he used his SRI tracker to stabilize spatial frequency displays, and compared its performance to published data obtained with traditional invasive methods.

"We have found that the subject's task is easier (and his data are less noisy) if we stabilize only that aspect of the stimulus that is directly involved in his judgements. The eye tracker measures both horizontal and vertical eye movements, but we only use the horizontal signal to stabilize the image, since our targets contain only vertical lines. The same psychophysical method was used for moving or stationary gratings, stabilized or unstabilized. The subject found his steady-state threshold contrast by adjusting a multiturn potentiometer [Fechner's Method of Adjustment, see section 6.1.1]. He was instructed to choose whatever criterion he thought he could best remember and reproduce for all spatial frequencies and experimental conditions (rather than, e.g., adopting a 'high' or 'low' criterion level). Our unstabilized viewing technique was the same one used in previous studies. . . the subject. . . was permitted to move his eye 'normally', in any way that helped him see the target. He was instructed to judge his threshold in the *steady state* [Kelly's italics], i.e., to take enough time for each setting so that he felt the result would not change if he took longer. With this procedure, onset transients have no effect, and the contrast threshold is quite large at very low spatial thresholds. . . The same procedures were used in the stabilized condition, with one important exception. Since our purpose was to study the effects of normal eye movements by eliminating them as completely as possible, the subject was instructed to cooperate with the stabilizing equipment, rather than try to defeat it . . . a blink or a large saccade can disrupt the stimulus with any type of stabilizer*. After a few practice trials, most subjects learned how to find a steady-state threshold in less than half a minute. Even with experienced subjects, however, it always took longer to reach steady state in the stabilized than in the unstabilized condition. . . In any stabilized-image technique, the movements of the retinal image produced by the apparatus should match the corresponding movements of the retina. In our experiments, this correspondence is very sensitive to small errors and must be adjusted as precisely as possible. The adjustment also varies between subjects. . . the optimum gain of our stabilizer varies from subject to subject, because it depends on the optical geometry of the subject's cornea and lens. . . Each subject must establish his own optimum setting, which he does in the following way:

"With the eye tracker controlling the position of a vertical line, the subject first adjusts the stabilizer as well as possible by voluntary fixation. Using two unstabilized fixation marks, he attempts to make

---

* This may no longer be true. A silicone annulus sensor coil, which can signal eye orientation when placed in a suitable magnetic field, will, unlike a research contact lens, a Yarbus-type sucker or the SRI tracker, follow the eye during saccades of all sizes (Steinman et al., 1982; Collewijn et al., 1983) and also during blinks made with the eye open or closed (Collewijn et al., 1985). In principle, this device could provide very effective long-term stabilization. So far, it has only been used for long-term stabilization of the vergence component of binocular eye movements (see Collewijn and Erkelens, this volume). The effectiveness of the stabilizing device used in conjunction with the annulus in these experiments has not yet been described sufficiently to allow comment on its potential for eliminating virtually all image motion during saccades and blinks.

the stabilized line move from one fixation mark to the other, as he looks from one to the other. To achieve this condition, he uses a knob that controls the gain and one that controls the bias of the eye-movement signal; i.e., one knob merely translates the stabilized line across his visual field to the desired position, while the other magnifies or minifies the effect of his eye movements on the position of the line. The task is not difficult but gain settings made in this way are seldom repeatable to better than about 10%. The reason for this lack of precision seems to be that the subject's line of sight is seldom where he thinks it is. . . once gain has been roughly set by the voluntary-fixation method, it can then be optimized much more precisely by means of an afterimage technique. Again the subject views a vertical line, which is considerably brighter than the rest of the CRT screen. Now if the line were turned off, he would see a prominent dark line at the same position (due to the fatiguing of that region of his retina). This negative afterimage provides a retinal landmark that can be used to adjust the gain, by comparing the positions of the bright and dark lines (i.e., the stimulus and the afterimage). While viewing the stabilized, bright line, the subject swings his eyes back and forth over a small, horizontal excursion of a few degrees. . . Now if the gain is perfect, the dark line [the afterimage] will not be visible. . . With this technique, or variations of it (e.g., using a step instead of a line), an experienced subject can repeat his optimum gain setting to better than 1%*. . . The factors that keep a stabilized image at or below threshold are in such delicate balance that almost any disturbance, a large saccade or even a loud noise, can disrupt our measurements. If the subject is aware of the source of destabilization, he can simply refrain from making a judgement until his steady-state returns" (Kelly, 1979a, pp. 1269–1271).

Kelly goes on to compare his stabilized and unstabilized CSFs with results obtained in three other experiments, two of which used contact lens optical levers and the third of which used a limbus monitor, hardly believed by anyone to be capable of beginning to stabilize a target. He reported that the best results of the contact lens optical lever stabilization experiments showed stabilization reducing contrast by factors of only 5 or 6 near the 2–4 cycles/° peak of the CSF, whereas his SRI tracker produced a reduction in contrast by a factor of 20 in both his and Riggs's eyes. The SRI tracker, in his hands, was capable of much better stabilization than had been reported in the experiments chosen for comparison with the contact lens optical lever, one of the two traditional invasive stabilization techniques. In Kelly's words "our threshold-elevation ratio is an order of magnitude greater than any of the others. . . moreover, we can mimic the results of others simply by decreasing the precision of our stabilizer. . . thus it would seem that we have achieved more precise stabilization than any of the other techniques" (p. 1272). Note, however, that, even with the SRI tracker, displays only lost contrast when stabilized. It seems here that Kelly has a confirmation of Barlow's (1963) main result, which is described in section 6.2.6, that well-stabilized displays with good optical properties and high contrast, lose contrast and high spatial frequency content, rather than disappear completely as Yarbus and Gerrits had claimed (see sections 6.2.4, 6.2.8 and 6.2.9).

Kelly next attempts to explain the reasons for the reduction of contrast after stabilization in terms of a 'sensitivity mask' – a kind of neural, rather than photochemical, afterimage which develops when patterned stimulation remains in the same place on the retina. "It takes about 10 s for the sensitivity mask to form or to dissipate – i.e., to reach steady state – with any stabilized, high-contrast stimulus . . . it is the sensitivity mask that makes sta-

---

* The reader should recognize that Kelly has adapted the afterimage technique, used by Barlow (1963) and then by Riggs and Schick (1968) to evaluate the effectiveness of various stabilizing methods, to help him optimize stabilization gain settings. It should be noted, however, that it has been shown recently that there are serious problems with using the perceived location of an afterimage to infer eye positions or retinal image motions, despite the superficially compelling logic of this procedure (see, for example, Ferman et al., 1987; or Collewijn et al., 1990b for similar failures of nonius line procedures).

bilized images disappear" (p. 1273). He then measures the CSF of this mask by creating afterimages of spatial frequency gratings and then varying the contrast of the display so as to bring the bars back to threshold after they have faded. Given the admitted difficulty of establishing and maintaining 'steady-state' thresholds (quoted above), it is clear that making these measurements was an arduous, as well as tricky, task. CSFs were obtained, however, allowing Kelly to point out that in this kind of measurement, in which a 'sensitivity mask' is overcome by a contrast increment, "the more steadily the stimulus is held in position during mask formation, the higher the resolution of the mask. We sometimes use this high [spatial frequency] cutoff as a final criterion for our stabilizer gain setting. The fact that some subjects can resolve a 12 cycle/° afterimage [of a grating] argues that the noise in the eyetracker must be less than the (2.5′) width of one bar in this grating" (p. 1273).

This last point is important for our present, and future, discussions, because it calls attention to the influence of tracker noise on the high-frequency cutoff of the CSF. 'Tracker noise' is actually determined both by characteristics of the instrument and by idiosyncratic characteristics of each subject's eye. The former can be estimated more or less accurately by recording reflections from an appropriately configured artificial eye. The latter cannot. It can only be estimated from behavioral measurements, at best capable of suggesting the lower limit for stabilization noise measurements. The significance of this limitation in inferring visual functions from 'stabilized' data can be illustrated by considering Kelly's sensitivity mask CSFs for himself and his colleague, M. Clark (see Fig. 10 in Kelly, 1979a). Kelly's high-frequency cutoff was about 12 cycles/°. Clark's high spatial frequency cutoff was about 7 cycles/°, almost a factor of two lower than Kelly's cutoff – in itself as high as any reported tracker-stabilized high-frequency cutoff. So, two subjects stabilized with the same instrument and participating under presumably identical experimental conditions showed very different high spatial frequency cutoffs, a fact which "argues that the noise in the tracker" can be very different in different subjects (see quote from p. 1273 just above). Now look at the high-frequency cutoffs of the CSFs reproduced in Fig. 20. All range from more than 30 cycles/° up to about 50 or 60 cycles/°. These are normal expected values for an unstabilized CSF, i.e., the high-frequency cutoffs measured in many subjects in many laboratories. This means that SRI tracker noise in the living eye, inferred from Kelly's measurement of the CSF of the sensitivity mask, attenuated the high spatial frequency response of his visual system by a factor of 3 and attenuated the high-frequency response of Clark's visual system by a factor of 8, assuming only that Clark's unstabilized high-frequency cutoff falls in the normal range as Kelly's cutoff does (see DK's unstabilized data plotted in Kelly, 1979a, Fig. 6).

The implications of tracker noise, as measured in the living eye, and its potential influence on the shape, as well as on the high-frequency cutoff, of the CSF, will be considered further after we have reviewed a recent paper by Arend and Timberlake (1986) in which the importance of SRI tracker noise figured prominently once again. Arend's interest in the significance of noise during retinal image stabilization was kindled, at least in part, by Kelly's (1979a) concluding remarks: "It is not certain that maximum threshold elevations much greater than our ratio of 20 could be obtained with still more precise stabilization. The only way to be sure of this would be to increase the precision and try the experiment, which was not possible at this writing. On the other hand, our stabilized thresholds seem to be tuned to the sustained receptive fields of the fovea. If these sustained responses cannot be silenced even by perfect stabilization, then we may already have reached the point of diminishing returns" (p. 1273). As early as 1973, Arend had proposed a model of contour perception which had depended on Yarbus's and Gerrits's results to discount the significance of 'sustained' retinal neural responses (see sections 7.2.4–7.2.6). Kelly's conclusion was contrary to this claim, which encouraged Arend and Timberlake to determine, on the basis of Kelly's (1979b) measurements of threshold motions imposed on a stabilized display, the minimum amount of retinal image slip that would be sufficient to drive 'transient' retinal neural units. Only displacements below this limit would properly fall in Kelly's region of 'diminishing returns' – the region within which better stabilization could not affect vision because 'sustained' retinal elements would take over visual processing.

9.4. What is psychophysically perfect image stabilization? Do perfectly stabilized images always disappear?

Arend and Timberlake (1986) raised these two questions in the title of their paper, which set out to evaluate and extend Kelly's (1979a,b) reports. Two things of particular relevance to this goal had happened in the intervening 7 years. First, the SRI tracker had advanced to Stage IV, claimed by its developers to have half the noise of the Stage III instrument used by Kelly. Second, Steinman et al. (1985) had reported effects of retinal image motion on contrast sensitivity to both high and low spatial frequencies quite different from those Kelly had observed and had suggested that these differences might have been caused by Kelly's use of unidirectional, constant-velocity retinal image motions, which do not resemble the natural oscillatory retinal image motions associated with fixation when a human being sits still or moves his head. Arend and Timberlake attempted to resolve this and related problems in a paper whose text, as well as title, began with a question. Namely,

"Does spatial pattern vision require temporal change of the retinal image? Definitive experiments on this question must involve temporally constant retinal stimuli, i.e., stabilized retinal images. Three decades of stabilized-image research have not provided a clear answer. Virtually all researchers have reported that stabilization raises luminance-grating contrast thresholds, but some report that residual detection of high contrast luminance gratings remains, in the form of either continuous visibility or fluctuating appearance and disappearance. On the other hand, several researchers using elaborate suction contact-lens stabilization report that even high-contrast, high luminance patterns disappear completely" (p. 235). Such introductory remarks should seem familiar to the reader, as should Arend and Timberlake's claim that "Kelly argued that all important retinal-image motion had been eliminated in his experiments. Accordingly, he attributed his residual contrast sensitivity to psychophysical mechanisms capable of detection of a stationary retinal image in the absence of temporal modulation from any source . . . It is difficult to assess adequacy of stabilization. If one's stabilized pattern disappears, one obviously has a strong argument that all temporal changes important for perception of the particular pattern under that particular prevailing experimental condition have been eliminated. If the pattern does not disappear, arguments for stabilization adequacy rest on descriptions of technique . . . In order to attribute residual pattern detection to static [sustained or tonic] psychophysical detectors, one must convincingly argue that all psychophysically significant motion of the retinal image has been eliminated. What is the smallest psychophysically meaningful retinal image motion? There has been no defensible criterion for judging the psychophysical importance of residual motion except image disappearance. Some of Kelly's observations indicate that the visual system is sensitive to extremely small local temporal-luminance changes . . . drifting an otherwise stabilized luminance grating at a constant velocity of 0.012°/s [43″/s] . . . raised contrast sensitivity by more than 0.5 log unit over a broad spatial-frequency range. Assuming a summation time for contrast thresholds of about 0.1 s, the patterns move only one fifth of the intercone distance [within the 20′ diameter foveal bouquet] in one critical duration. A 100% contrast grating need move even less than these threshold gratings to produce identical local temporal changes" (p. 235).

Arend and Timberlake go on to use a Stage IV, SRI tracker to replicate some of Kelly's (1979b) results for motion imposed on a drifting, stabilized spatial frequency grating and to develop "a method for calculating threshold *retinal-image motion* [their italics] as a function of grating contrast and spatial frequency from existing flicker data". They also try, less successfully as we shall see, to test this method empirically in new psychophysical experiments.

Arend and Timberlake's (1986) comparison of their contrast sensitivity results with Kelly's (1979b) is reproduced in Fig. 21. Both sets of results are for constant-velocity, unidirectional, drifting spatial frequency displays. The maximum reductions of contrast sensitivity, resulting from SRI tracker stabilization of an objectively stationary grating, are included in both graphs. The subjects shown are Arend (top) and Kelly (bottom). In the authors's words, "The effect of pattern velocity and the overall shape of the curves agree with Kelly's. Under these stabilization conditions the 0-deg/s grating threshold was 50 times [1.7 log units] the 0.15-deg/s threshold at 1 cycle per degree (c/deg).* At higher spatial frequencies the elevation factor was smaller, decreasing to about 7 [0.84 log units] at 8.8 c/deg. As in Kelly's experiment, a drift velocity of only 0.012 deg/s (=43.2 sec of arc/s) increased sensitivity from the stabilized value by as much as 0.75 log unit" [a factor of 5.6] (p. 236). Arend and Timberlake go on to show that these impressive differences between stabilized and what they consider to be velocities similar to fixational drifts (0.15°/s) were, to no small degree, the result of their particular psychophysical threshold procedure, which was similar to the procedure Kelly had used 7 years before. Both experiments used the traditional Method of Adjustment, which exposes the test display continuously while the subject adjusts contrast during 'descending', as well as during 'ascending', trials. This means that half the trials began with contrast set to a high level and that contrast remained, more or less, above threshold until the subject decided that he had reached contrast threshold. Such continual exposure to relatively high contrasts will produce afterimages whenever the grating is relatively stationary on the retina (see Kelly's treatment of the 'sensitivity mask' in section 9.3). Such afterimages would have the effect of exaggerating the apparent loss of contrast sensitivity when a grating is stabilized as compared to contrast sensitivity measured when a grating was moving on the retina. This would be true regardless of the means by which the grating was moved, both when the stabilized grating was moved on the retina by experimentally introduced drifts and by natural fix-

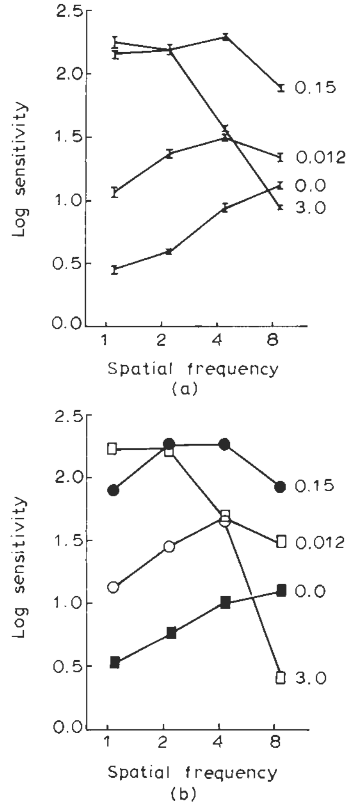

Fig. 21. Mean log sensitivity for stabilized gratings with superimposed constant-velocity drifts. Top panel shows data for subject LA and the bottom panel shows data for Kelly (1979b). Drift velocity is given as degrees/second on the right side of each graph. Bars in the top graph = 1 standard error of the mean. (From Arend and Timberlake, 1986)

---

* Note that this factor is larger than Kelly's (1979a) factor of 20 for the same condition. This difference probably reflects the improvement in tracker noise from Stage III to Stage IV. However, the fact that Arend and Timberlake confined all of their observations to spatial frequencies below 10 cycles/° might imply differences between the noise spectra of Stages III and IV, as well as differences in the noise properties of the various subjects' eyes.

ational drift eye movements when the display was not stabilized.

Burbeck and Kelly (1984) considered this problem and reduced the potential contribution of the 'sensitivity mask' to contrast sensitivity threshold measurements by using a 'reduced-exposure technique'. This psychophysical threshold procedure reduces the exposure to high contrast stimuli by employing only ascending tests and by only allowing contrast in the display during specific test intervals (a homogeneous display of the same space-average luminance was presented in the intervals between the tests). In effect, the contribution of the 'sensitivity mask' was reduced by using the traditional Ascending Method of Limits in which discrete test stimuli are used and the sequence of discrete tests always begins well below the expected threshold value. Arend and Timberlake (1986) used this reduced-exposure technique to compare the contrast sensitivity under stabilized and normal viewing (fixation with the head supported) and obtained much smaller effects of stabilization with a Stage IV SRI tracker than Kelly had reported for his noisier Stage III instrument. Arend and Timberlake's largest difference between stabilized (0.0°/s) and unstabilized (about 0.15°/s) contrast sensitivity thresholds was observed with a 1 cycle/° grating. The difference in stabilized and unstabilized viewing thresholds was now only 0.8 log unit (a factor of about 6 rather than the factor of 50 illustrated in Fig. 21). This means that the sensitivity mask reduced the contrast sensitivity of the Arend and Timberlake stabilized display by a factor of more than 8 (0.9 log unit). This suggests that Kelly's (1979a) reduction of contrast sensitivity with the stabilized image, if measured *without the sensitivity mask*, would have been less than a factor of 3 (0.4 log unit) rather than the factor of 20 (1.3 log unit) described in his original paper.* The difference between stabilized and unstabilized thresholds fell to about 0.3 log unit (a factor of 2) at 8 cycles/°, near the highest spatial frequency studied by Arend and Timberlake, using the reduced-exposure technique. This restricted spatial frequency range suggests that Arend's eye in the Stage IV tracker, like M. Clark's eye in Kelly's (1979a) earlier report with a Stage III tracker, contributed more noise to tracker performance than Kelly's eye in his Stage III tracker. Remember, Kelly could use his eye to study much finer spatial frequency displays, as fine as 12 cycles/°.

Arend and Timberlake's demonstration that stabilization with the SRI tracker is far less deleterious to contrast sensitivity when the contribution of the 'sensitivity mask' is reduced by using the reduced exposure technique was assumed (above) to apply to Kelly's (1979a) earlier report as well. This would bring tracker stabilization very much in line with previous reports based on contact lens optical lever stabilization. Kelly's claims of superiority over this method rest on a comparison with two experiments (see Kelly, 1979a, Fig. 9). The first, Tulunay-Keesey and Jones (1976), is not a fair comparison because the particular optical lever arrangements employed did not make provision for mounting the contact lens mirror to be normal to a line parallel to the visual axis, which means that eye torsions and translations would be confounded with eye rotations. Stabilization would have to be relatively incomplete and it is not surprising, therefore, that only modest differences between stabilized and normal viewing would be observed. The other comparison was with Gilbert and Fender (1969), whose stabilizing technique was not subject to this problem. These authors also did not keep their gratings in view continuously, as Kelly had, which means

---

* Burbeck and Kelly (1984) observed just this kind of reduction in the effectiveness of Stage IV SRI tracker stabilization when they used their 'reduced-exposure technique' to study local adaptation effects (0.16 log units for DK's eye at 3 cycles/°, a factor of 1.4, rather less than the factor of 20 reported for this spatial frequency and his eye when he used his noisier Stage III tracker). Note, however, that Burbeck and Kelly did not report the CSF during normal unstabilized viewing when they used their reduced exposure technique. It seems likely, on the basis of results obtained by Arend and Timberlake with the reduced-exposure technique, that the reduction in sensitivity produced by stabilization would be somewhat greater, perhaps 2.4, when local adaptation effects are prevented by an appropriate psychophysical procedure as well as by the drifts which occur during normal fixation.

that Gilbert and Fender's estimates of reductions in contrast sensitivity after stabilization were probably exaggerated less by the presence of a 'sensitivity mask' than Kelly's estimate. Arend and Timberlake (1986) found this to be quite important when estimating stabilization effectiveness from reductions of contrast sensitivity (see above). There are additional problems with Kelly's comparison of his with Gilbert and Fender's results. Kelly chose to describe the superiority of his over their results by comparing contrast sensitivity at only 3 cycles/°, a comparison most favorable for the conclusion he preferred. Gilbert and Fender's greatest loss of sensitivity was, as would be expected from Van Nes's (1968) and Arend's (1976) prior work (see Figs. 11 and 15), at lower spatial frequencies. In Gilbert and Fender's experiment, contrast sensitivity was reduced by a factor of about 6 (0.78 log unit) at 0.3 cycles/°, which is the same as Arend and Timberlake's factor when they used the reduced-exposure technique, and by a factor of about 4 (0.62 log units) at 1.2 cycles/°. At 3 cycles/°, the spatial frequency Kelly chose to compare, it was down to a factor slightly less than 2 (0.25 log units), but this factor is hardly representative of maximum loss of contrast sensitivity reported by Gilbert and Fender. Once it is assumed that Kelly's (1979a) reported reduction of contrast sensitivity with SRI tracker stabilization is exaggerated because over 90% of the reduction arises from the psychophysical procedure employed, as Arend and Timberlake showed was the case for their data, Kelly's maximum reduction due to stabilization falls from a factor of 20 to a factor of 2.4. This is less than half the maximum reduction in contrast sensitivity observed by Gilbert and Fender (1969).

It seems reasonable, then, to claim that Stage IV SRI tracker stabilization may be as good as, but not really better than, a properly designed contact lens optical lever stabilizing technique and that Stage III SRI tracker stabilization is worse. It is important to realize, moreover, that Gilbert and Fender were able to study spatial frequency gratings as fine as 30 cycles/°, almost 3 times as fine as could be studied with Kelly's best tracker subject (himself) and almost 4 times as fine as the other tracker subjects whose performance he and Arend and Timberlake (1986) described. Stage IV SRI tracker stabilization seems to be as good as optical lever stabilization with respect to the maximum possible reduction in contrast sensitivity, but the tracker restricts study to the lower part of the spatial frequency range the human eye is capable of resolving (at best only up to 12 cycles/° in the normal 60 cycle/° range). This spatial frequency restriction arises primarily from the noise of the stabilizing instrument, a fact that Arend and Timberlake (1986) go on to treat theoretically.

9.5. Calculation of stabilization accuracy required for disappearance

Arend and Timberlake point out that "if one's best-stabilized, 100% contrast gratings are detectable, there are two possible explanations: Either patterns are being detected by static pattern-detecting mechanisms [sustained or tonic neural elements] or one has not succeeded in eliminating all psychophysically important temporal change. To argue logically that static pattern-detecting mechanisms exist, one must (1) assume as a hypothesis that only dynamic mechanisms [transient or phasic neural elements] exist and then (2) somehow determine whether sufficient temporal change occurred to allow detection by dynamic mechanisms. Only if such change has not occurred is there need to postulate static mechanisms. The experimental problem is to determine how much temporal change is too much and whether one's stabilization technique produced that much. With any technique there is a formidable array of potential sources of stabilization failure. Whereas other sources of destabilization are potentially larger . . . [e.g., blinks, saccades, instrumentation delays], we have chosen to focus our argument conservatively on electronic noise in eyetracker position signals . . . One can directly measure the psychophysical consequences of small pattern movements only if the stabilization errors during measurements are small relative to the movements being evaluated. To evaluate directly movements as small as Purkinje image eyetracker noise (approximately 1 min of arc rms) one must be able to stabilize substantially better than that. To our knowledge no measurements meeting this requirement exist" (p. 238).

Arend and Timberlake (1986) pointed out that despite the fact that direct measurements were not available, with two simple assumptions it became possible to calculate the sensitivity of the visual system to small movements (imposed experimentally or by tracker noise) on the retinal image of a stabilized grating display. In other words, they proposed that it was possible to use available SRI tracker data obtained with counterphase flickering, stabilized gratings to evaluate Kelly's (or other) speculations about the neurological underpinnings of visual processing that had been inferred from a stabilization experiment with the tracker (or any other technique). This was possible because "a grating oscillating over a distance that is small relative to one grating period is closely approximated by the sum of two perfectly stabilized gratings" (p. 238). This can be accomplished by having one stabilized, relatively high contrast, grating remain invariant in time while a relatively much lower contrast, superimposed, as well as stabilized, grating undergoes counterphase-modulation. This kind of stimulus manipulation produces the same changes in the temporal variations of the illumination on a local retinal region as is produced when a relatively low spatial frequency grating wiggles with some temporal frequency (i.e., translates sinusoidally) through an amplitude small relative to the period of the spatial frequency of the moving grating. This can be visualized by imagining the very shallow slope of a low spatial frequency grating moving, ever so slightly, back and forth (try 1 cycle/° or shallower), over a very small portion of the retinal surface (perhaps 3 minutes of arc, which would contain a row of at most 9 cones at the center of the foveal bouquet). Readers with poor visual imagery might look at Fig. 5. Illumination varies very little over small distances in the situation illustrated in this figure.

The demonstration that there is a counterphase-modulated grating stimulus, which will have similar time-varying local illumination properties, required two equations and two assumptions, namely:

$$L(x,t) \approx L_M[1 + C_s \sin 2\pi fx + (C_s 2\pi fa)(\cos 2\pi vt)(\cos 2\pi fx)]$$

Here $L(x,t)$ is the luminance in space and time, L_M is the mean luminance, C_s is the contrast of the grating, a is the amplitude, f is the spatial frequency of the grating and v is the temporal frequency of the oscillatory movements of the grating. The left-hand term in the equation represents the stabilized, high-contrast, time-invariant grating and the right-hand expression represents the relatively low contrast, stabilized grating undergoing counterphase-modulation with contrast equal to $C_s 2\pi fa$. (See Appendix A in Arend and Timberlake, 1986, for the derivation of this equation.)

The two assumptions required to proceed are (a) dynamic (transient) variations of luminance are required to drive neural elements, i.e., there are no static (sustained) neural elements and the time-invariant high-contrast grating is not represented in any way in neural activity and cannot, therefore, be detected, and (b) the neural signals produced by this time-invariant high-contrast grating are not only imperceptible, they also do not affect the neural signals produced by temporal variations of the lower-contrast, counterphase-modulated, grating which can be detected, providing some threshold illumination change is exceeded. Accepting these assumptions allows us to ignore the time-invariant, high-contrast grating on the right-hand side of Equation 1 and the contrast required for threshold of the stabilized, low-contrast, counterphase-modulated grating is, therefore, represented by Equation 2, which assumes perfect, that is, noise-free, stabilization:

$$C_v f = C_s 2\pi fa$$

where C_v is the threshold for the counterphase-modulated grating of spatial frequency f, temporal frequency v. C_s is, as in Equation 1, the contrast of the counterphase-modulated grating. Equation 2 only works for perfectly stabilized gratings. So, how can we proceed from here?

Arend and Timberlake (1986) point out that "visual sensitivity to perfectly stabilized counterphase-modulated gratings will be closely approximated by actual eyetracker-stabilized sensitivity when the grating spatial period is large relative to eyetracker noise. Under this condition the temporal changes produced by the small oscillations of eyetracker noise are small relative to the temporal changes that are due to the counterphase modulation" (p. 239). This allowed them to use Equation 2 to estimate the threshold amplitude of grating displacement (a in Equation 2) that would produce a local variation of illuminance sufficient to just excite the dynamic neural elements. Their calculation was based on an extrapolation from existing measurements of contrast sensitivity thresholds to a sinusoidally flickered spatial frequency display, whose spatial frequency was sufficiently low to satisfy the requirement described just above. They calculated the hypothetical threshold displacement for a 4 Hz oscillatory sinusoidal movement of a perfectly stabilized, 100% contrast, 2 cycle/° grating. Their hypothetical displacement threshold calculation took contrast threshold data from Kelly's (1979b) measurements made with the same spatial and temporal frequencies. They extrapolated the peak-to-peak displacement threshold amplitude (2a) for a *hypothetical 100%* contrast, 2 cycle/° spatial frequency grating and found it to be only 0.0011° – only 8″! Arend and Timberlake (1986) suggested on the basis of their calculations that Kelly's assumed action of sustained neural elements at and below the noise level of his tracker was unwarranted. In their words, tracker "noise alone . . . prevents conclusions about dynamic detectors on the basis of data from electromechanical trackers" (p. 239).

Arend and Timberlake next attempted to test the equivalence of a moving and a flickering grating, whose parameters were chosen along the lines set forth in Equation 1. This led them to do 'a preliminary experiment' described, in their words, as "allowing direct comparison of thresholds for wiggling and counterphase-modulated gratings . . . In order to measure meaningful thresholds for sinusoidally translated [wiggling] gratings, the amplitude of movement must be small relative to one cycle of the grating and large relative to the noise of the stabilizing system. Both conditions can be met for a small range of spatial frequencies and movement amplitudes" (p. 240). A 1 cycle/° spatial frequency grating, oscillating through a peak-to-peak distance of 3.33′, was considered to meet this constraint sufficiently well to permit a direct comparison of the temporal CSF obtained by varying the temporal frequency of this stimulus (the contrast required to see the grating when it wiggled at various frequencies) with the temporal CSF obtained by counterphase-modulating a relatively low contrast grating of the same spatial frequency, superimposed on a similar, temporally invariant grating of much higher contrast. This second experimental condition is an example of the stimulating conditions, assumed to be equivalent to the stimulating conditions produced by a wiggling grating, in the calculations (described above) of the displacement threshold of the dynamic neural elements. In other words, the assumption that local temporal variation produced either by actually moving a grating (motion) or by counterphase modulation (flicker) has the same visual consequences was tested empirically. Temporal frequency was varied in octave-steps from 1.1 to 17.6 Hz. In the wiggle condition, these 5 frequencies moved the grating at average speeds ranging from 7.3′/s to 116′/s. The results of this experiment are summarized in Fig. 22.

Four of the 5 data points for flicker show greater sensitivity for counterphase modulation than for actual movement of the grating. In 3 of these 4 cases the standard deviations do not overlap, suggesting to us, at least, that Arend and Timberlake's description of "excellent agreement through 4.4. Hz" may be over-stated. Other features of potential significance in these results are the large and disorderly relationships at 8.8 and 17.6 Hz, which the authors suggest call into question their assumption that the time-invariant high-contrast grating is both undetectable and not interactive with the lower-contrast, counterphase-modulating grating. There is another disquieting feature in Fig. 22 that can be seen when it is compared with Fig. 21 (top), which reproduces the same subject's data for a drifting grating motion as a function of spatial frequency. Note that increasing the retinal image motion of the relatively low spatial frequency grating studied in Fig. 22 (1 cycle/°) led to a *reduction* in contrast sensitivity by a factor of almost 2 (0.3 log unit) whereas increasing retinal image speed with the same spatial frequency in Fig. 21 led to a modest *increase* in contrast sensitivity.* Furthermore, recall that when this subject's CSF was measured with the 'reduced-exposure technique' (described above), the difference between his SRI-tracker-stabilized contrast sensitivity and his contrast sensitivity when 'natural' biteboard fixational eye movements were permitted was as large as 0.8 log units (a factor of more than 6) with the same spatial frequency (his natural biteboard fixational drift eye movement speed, had it been recorded, would probably average about 9′/s). Ever since Van Nes's report (1968) we have come to expect an improvement in contrast sensitivity when motion is imposed on low spatial frequency gratings (see Fig. 11 for Van Nes's data, Fig. 15 for Arend's data and Fig. 20 for Steinman et al.'s data). The Arend and Timberlake (1986) result for a wiggling grating, which is reproduced in Fig. 22, is hard to understand when viewed in this historical perspective. Perhaps SRI tracker noise, which would have an average speed of 26′/s if it is assumed to be a 4 Hz sinusoid (a speed near the middle of the range shown in Fig. 22), had adverse effects on contrast sensitivity, adverse effects which would not be prominent with unidirectional drifting gratings or when tracker stabilization is not used.

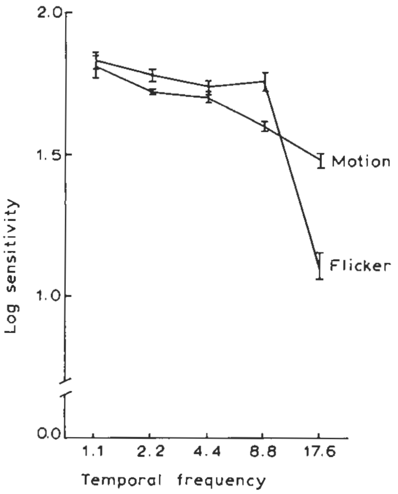

Fig. 22. Mean log sensitivities for stabilized 1 cycle/° gratings temporally modulated by two methods, i.e., by counterphase flicker (Flicker) or by moving the 1 cycle/° grating sinusoidally through 3.33 minutes of arc peak to peak (Motion). See the text for a discussion of the significance of this figure. (From Arend and Timberlake, 1986)

We believe that the role of retinal image motion in the detection of contrast has not yet been worked out. Studies during the past decade in which the SRI tracker has been used to impose motion on stabilized gratings are less than completely satisfying in a number of respects. In addition to the uncertainties described above, we are disappointed by the fact that this technique has, because of its noise, confined study of the effects of retinal image motion on contrast sensitivity to the lower portion of the spatial frequency range to which the human visual system responds. Furthermore, the shape of the CSFs measured below the 7–12 cycle/° 'high' frequency spatial cutoffs reported for eyes of different subjects, may not be free from important distortion as has been assumed (e.g., Kelly, 1979b). If the spectrum of tracker noise is not actually known because no artificial eye can actually reproduce the characteristics of an individual's natural eye, it will not be possible to correct the shape of

---

* We are assuming that the data point plotted for 3°/s motion at 1 cycle/° lies above the value for 0.15°/s motion in Fig. 21 (top) entirely on the basis of the fact that Arend and Timberlake described their functions as similar to Kelly's (bottom) in Fig. 21 where different plotting symbols were used and one can be sure the the sensitivity was greater when the grating moved faster. If we are wrong, the results summarized by Arend and Timberlake in Fig. 21 are as mysterious as the results summarized in Fig. 22 in the light of previous observations by Arend and others.

an individual's stabilized CSF for artifacts introduced by stabilization. Tracker noise may well, as has been assumed, become less important with lower spatial frequency gratings but its actual contribution below 12, or 8.8 or 7 cycles/° probably should not simply be ignored – the common current practice. The situation above 12 cycles/° is even more troublesome. There are no meaningful tracker-stabilized data, whatsoever, in this region. The spatio-temporal surface modeled by Kelly (1979b) only covers a restricted range of spatial frequencies. The effects of retinal image motion on spatial frequencies falling in the range that was the domain called 'visual acuity' (30–60 or more cycles/°), throughout the century of pre-Fourier Optics, has not been and cannot be explored with this methodology in its current stage of development.

In our opinion, the most valuable, and possibly lasting, contribution of contemporary SRI tracker research with motion imposed on stabilized images can be found in the Arend and Timberlake (1986) paper. Specifically, their attempt to provide a quantitative technique for working out the potential contribution of instrumentation noise to threshold measurements of contrast sensitivity is intriguing and seems to be theoretically sound as well. Even if their estimates of displacement threshold, which were based on a *hypothetical* 100% grating contrast, prove to be in error by a large factor, their insistence on the need to estimate these thresholds and a technique for doing so provided a refreshing approach in a very old, much studied and technically very difficult problem area. Anyone considering working further in this area would probably benefit by thinking hard about their concluding remarks.

They concluded their provocative paper by calling attention to the fact that their "calculation focussed on noise in the eye-position signal, there are much larger potential sources of destabilization. No practical ramping of onset and offset can eliminate important stimulus transients... Eyeblinks produce large illuminance changes over the entire retina, often more frequently than the 15–20 s required for the effects of smaller transients to subside [see Burbeck and Kelly, 1982, for effects of transient stimulation]. The time delay between eye motion and the tracker signal also produces substantial retinal-illuminance transients at edges in the image during saccades or high-velocity smooth movements... Even if these large sources could be eliminated, there are smaller destabilization sources that are difficult or impossible to remove. Small shifts of the natural pupil relative to the artificial pupil [in the stabilizing device] could be significant because of the Stiles-Crawford effect. At the high threshold contrasts in Fig. 21 the just detectable grating appeared and disappeared in synchrony with the subject's heartbeat... Given the extreme sensitivity indicated by the observations and slow-drift calculations, the remarkable observation is not that we failed to get full disappearance but that Yarbus and Gerrits et al. did" (pp. 240–241).

In this section, we have seen that research on retinal image motion, imposed on a stabilized display, is incomplete, in large part because of technical limitations of instrumentation available for stabilizing images. These limitations arise more from the exquisite sensitivity of the human visual system to moving contrasts than from any lack of sophistication on the part of contemporary instrument builders. Stabilizing instruments, used with grating targets whose retinal contrast is physically realizable (i.e., well under the 100% assumed in Arend and Timberlake's calculations), need position accuracy better than 20″ and velocity accuracy better than 40″/s to be better than the visual system under study. Even such accuracy is very difficult to achieve in intact living organisms. Our current knowledge is incomplete at both ends of the spatial frequency spectrum. At the low end, we have not yet determined, after almost 40 years of stabilized image research, whether there are functionally significant sustained neural elements (called 'static' elements by Arend and Timberlake) which could provide vague visual outlines of patterned stimulation (the kind of percepts Barlow, 1963, reported), if prolonged, functionally 'perfect' stabilization could be achieved. It remains quite plausible to continue to subscribe to Yarbus's, Gerrits's, Arend's and their coworkers's claim that transient stimulation is a *necessary* condition for vision. The upper part of the normal spatial frequency range (the region traditionally studied under the rubric 'visual acuity') has also been inaccessible to the best current non-invasive stabilizing instrumentation. We have no direct, meaningful, measurements made with an SRI tracker of the effects of motion imposed on sinusoidal gratings whose spatial frequencies are greater than 7–12 cycles/° (depending on the subject). Available data only cover a rather modest portion of the range of discriminable spatial frequencies – a range which extends as high as 60 cycles/° in normal observers.

We also still do not know whether unidirectional drifts imposed on stabilized gratings have the same effects on contrast sensitivity as imposed sinusoidal oscillations. This is important because a sinusoidal oscillation, rather than a unidirectional drift, is more like the image motion produced by the natural fixational eye movement pattern of a human being who is sitting without artificial head supports, standing or moving about. Arend and Timberlake (1986) have taken a small step towards answering this question. They have shown that very small amplitude sinusoidal oscillations (3.3′ p-p) and unidirectional drifts of a 1 cycle/° grating which have about the same average speed have somewhat similar effects on contrast sensitivity. Larger-amplitude oscillations (the kind observed when the head is not supported artificially) and higher spatial frequencies have not yet been compared.

In summary, detrimental effects of image motion on contrast sensitivity in the upper, 'acuity' range of discriminable spatial frequencies, and beneficial effects of image motion at the lower end of this range, have not been described adequately in the currently available literature which has studied effects of motion imposed on a 'stabilized' grating display. The most recent work by Arend and Timberlake (1986) on the significance of instrumentation noise and local retinal adaptation in stabilized image experiments suggests that it will be exceedingly difficult, perhaps even impossible, to use this kind of technique to eliminate all influences, other than image motion introduced by the experimenter, on measurements of contrast sensitivity across the entire functional range of the human visual system.

Is there an alternative way to find out how retinal image motion affects contrast sensitivity? Could it be that CSFs measured in experiments done with what we have called 'natural retinal image motion' (i.e., retinal image motion resulting from incomplete oculomotor compensation during oscillations of the head) provide a useful beginning of a description of the effects of retinal image motion on contrast sensitivity? These CSFs could provide the initial basis for developing models of visual information processing in the presence of image motion (and also suggest a technique for future work) if we could be sure that a subject making contrast threshold judgements while he shook his head actually followed instructions and made all of his observations near the center of his head trajectory, where retinal image speed would be close to its highest average value. In other words, the CSFs described in section 8.2.3 and illustrated in Fig. 20 might provide a valid, albeit very approximate, description of human contrast sensitivity across almost the entire range of discriminable spatial frequencies if we could be sure that the subjects made all of their threshold settings on the basis of information obtained when the image of the grating display was moving rapidly on their retinas. Said more succinctly, it would be easier to take the 'natural' CSFs, shown in Fig. 20, seriously if we could be sure that the subject did not sneak a peek when he turned about. In the next section, we will show that similar CSFs can be obtained when sneaking a peek is impossible.*

---

* We thank Dr. Z. Pizlo for making valuable suggestions about how we might treat the material reviewed in this and previous sections on research with the SRI tracker. Dr. Pizlo brought high-level engineering skills to an unfamiliar problem area and was prepared, therefore, to raise very cogent questions and demand justifications for assumptions and analyses that might be ignored or overlooked by visual scientists, such as ourselves, who have been measuring visual CSFs for many years.

10. Effect of motion on the CSF when natural image speed controls contrast

10.1. Instrumentation for monitoring head and retinal image speed

We used the revolving magnetic cube-surface field-sensor coil technique to monitor the absolute horizontal position of the head or eye in space while psychophysical measurements of contrast sensitivity were made, both while the subject (RMS) sat still and while he oscillated his head. The principle of this technique for measuring the orientation of the head or eyes in space was described originally by Collewijn (1977), who used it for the first time in his study of the eye movements of freely moving rabbits. The development of this type of instrumentation for use with human beings (first reported by Steinman and Collewijn, 1980), was described recently by Steinman (1986). The properties of the particular revolving field monitor (RFM) developed at the University of Maryland and used for the new research on contrast sensitivity, which will be reported for the first time in subsequent sections, were first described in Collewijn et al. (1981). Examples of its effectiveness in studying the full range of human binocular eye and head movement with exceptional accuracy and precision can be found in Steinman (1982) and, most recently, in Collewijn et al. (1988a,b) and Erkelens et al. (1989a,b).

Here, we will only describe the main features of the Maryland RFM and then describe a few specific elaborations of this instrumentation, which made it impossible for the subject to 'sneak-a-peek' while his thresholds were measured as he oscillated his head (the reader should consult the references cited above for additional details). A sensor coil, mounted either on a head-band or on a silicone annulus which was sucked on to the eye, was located in a large (3.6 m diameter) homogeneous magnetic field generated around the subject. This homogeneous magnetic field was rotating in the horizontal plane at 976 Hz. The phase of the alternating electric potential, which is induced by the rotating magnetic field in a sensor coil located within it, is linearly related to the angular orientation of the rotating magnetic field. It follows that the phase of the signal induced in the sensor coil indicates the orientation of the sensor coil with respect to an earth-fixed coordinate system (the coordinates are provided by the large stationary field coils that generate the revolving magnetic field).

In the Maryland RFM, the generation of the rotating magnetic field and the measurement of the phase of the signal induced in the sensor coil is accomplished primarily with digital instrumentation. This allows the measurement of very small changes in the angular orientation of the sensor coil within a very large range of possible orientations. Specifically, RFM rms noise with output at 488 Hz (the output sample rate used in the experiment described in the following sections) is less than 40″ and the linearity of RFM output is better than 0.01% within the instrument's 360° range of operation. The angular position output of the Maryland RFM is digital (16 bits) and its slewing speed is equivalent to angular velocities of 12000°/s. Drift (measured from a stationary sensor coil) is less than 6 seconds of arc for periods ranging from 1 second up to 24 hours. Measures of angular position (orientation) are not measurably sensitive to translations (linear displacements) of the sensor coil within a region of about 50 cm near the center of the rotating magnetic field where the subject sits. The stability of the self-adhering silicone annulus sensor coil, when it is inserted properly, was demonstrated by its inventors for saccadic eye movements as large as 20° (Collewijn et al., 1975) and reconfirmed in the Maryland RFM for much larger eye and head movements (Collewijn et al., 1981), and also during blinks and prolonged closures of the eyelids (Collewijn et al., 1985).

Digital indications of the orientation of the sensor coil (position output in angular units) were fed to a dedicated microprocessor running at 12 MHz, which stored them in a FIFO buffer (first in, first out), and did a running calculation of velocity as they passed through a programmable digital filter. These digital velocities were converted to speeds (absolute velocities) and then fed to a digital thresholding device, which permitted the output to be limited to some maximum value selected by the experimenter. The digital output of the thresholding device was then converted to a voltage analogue of speed, which was fed to the Z-axis of a CRT display, located 6 meters from the subject, where the display subtended 1.5° horizontally and 1.2° vertically (a Tektronix Model 604 display with a P-4 phosphor was used). The noise (rms) in the voltage proportional to speed was 0.15°/s (9′/s) with the parameters employed (viz., position signals fed at 488 Hz into a sliding window, whose width was 33 ms). This measurement of noise in the speed voltage output was made when the input to the microprocessor was provided by a stationary sensor coil.

Sinusoidal spatial frequency displays were generated and varied by conventional analogue techniques, which meant that, when the voltage derived from the rotational speed of the sensor coil was fed to the Z-axis of the display, the contrast of the grating was proportional to the speed with which the sensor coil was oscillating. The maximum contrast available to the subject was limited to a value chosen by the experimenter. This limit, coupled with the control of contrast by the speed of the sensor coil, made it possible to measure thresholds in, and always in, the presence of motion of a subject's head or eye, depending on the placement of the sensor coil. This was possible because when the coil was not moving, contrast was zero. In other words, there was no grating pattern – the display was homogeneous. A grating pattern was present in the stimulus when, and only when, the coil was moving. The contrast produced by motion might, or might not, allow the grating pattern to be above threshold. Exceeding the psychophysical contrast threshold required that the coil was moving fast enough, and also that the experimenter had set the maximum permissible contrast level above the threshold value needed for the particular spatial frequency under study. The Z-axis amplifier was adjusted so that a voltage proportional to a speed of either 50 or 5°/s produced a contrast of about 75%, depending on whether the sensor coil was mounted on the head or on the eye, respectively (these values were determined empirically in pilot work).

*10.2. Procedures in the 'sneak-a-peek' control experiments**

Two experiments were performed with this new instrumentation. Both were controls for the earlier work (see section 8 and Steinman et al., 1985), which had reported rather modest detrimental effects of natural retinal image motion on high spatial frequencies (that is, frequencies above 8 cycles/°) as compared with the detrimental effects of motion that had been imposed on stabilized displays (Kelly, 1979a,b; Arend and Timberlake, 1986). The previous experiments with natural retinal image motion had also shown beneficial effects of image motion on a wider range of 'low' spatial frequencies (up to about 6 cycles/°), whereas benefits of image motion were not observed for spatial frequencies above 2 cycles/° when motion was imposed on a stabilized display.

The sensor coil was mounted on the head in the first experiment. This permitted long psychophysical sessions in which contrast thresholds could be measured, and replicated carefully, both when the subject (RMS) sat still and when he oscillated his head. When he sat still, the contrast available depended entirely on the value set by the experimenter (JZL) before each trial. When he oscillated his head, the contrast available depended *both* on the speed of the head and on the value set by the experimenter before each trial. In this condition, contrast varied appreciably throughout the trial, depending on how fast the subject moved his head. The maximum available contrast, which had been set by the experimenter before the trial, was only available while the head was moving fast. Contrast was reduced by the microprocessor as the head slowed down and dropped to zero whenever the head stopped moving. This technique made it difficult, probably impossible, for the subject to set contrast

---

* We thank Dr. T. Park for his help in running these experiments.

thresholds during head movement to be similar to the contrast thresholds set when he sat still simply by basing his judgements on what could be seen when his head slowed down.

In both conditions, monocular contrast thresholds were measured with a double staircase procedure (Cornsweet's, 1962, method, which uses two randomly interleaved staircases). In both conditions, the subject continued on the pair of staircases until the experimenter was satisfied that steady-state performance had been achieved (see Nachmias and Steinman, 1965, for a comparison of the Double Random Staircase Procedure with the Method of Limits). A single spatial frequency grating was studied in each experimental session. Threshold measurements, made when the subject sat still or when he oscillated his head, were interleaved in alternating blocks of 10 trials each. The CRT display was surrounded by a homogeneous 'white' baffle which subtended about 8° horizontally and about 5° vertically. The light reflected from this baffle was adjusted to the same space average luminance as the CRT display, which could be seen through a rectangular cut-out at the center of the baffle. Trials, in both conditions, began with the subject fixating a black cross located 1.5° to the left of the center of the grating display (the cross was drawn on the left side of the baffle).

The subject started each 8-s trial when he felt ready after the experimenter had indicated that the contrast that would be available on the next trial had been prepared. When the subject started a trial during which he would keep his head still, he made a single saccade to the center of the display and continued to fixate at this position until either he reported that he could see the grating pattern or the trial ended, which was recorded as a failure to see the pattern. The test contrast was switched-on during the time the saccade was made to the center of the CRT display. This strategy was used to prevent the subject from basing his judgements either (1) on the appearance of the contrast in the eccentric display before the trial had started, or (2) on switching-transients in the CRT display, occurring after he had shifted fixation from the eccentric pre-trial fixation cross to the center of the CRT. In the second condition, the subject also began each 8-s trial when the experimenter indicated that contrast had been set but, in this condition, contrast in the display was only available while the head was oscillating. Its maximum was limited to a value selected by the experimenter on the basis of the report made by the subject on the preceding trial on the same staircase (the response on the preceding trial on the same staircase also determined the test contrast presented when the subject sat still – it is in this sense that staircase procedures are described as 'interactive'). The subject began to oscillate his head as he started the trial and shifted fixation to the center of the display while his head was in motion. The frequency and amplitude of his head oscillations were self-selected on the basis of what had been helpful during previous trials for obtaining the contrast required to see the grating (both the frequency and the amplitude of head oscillations were influenced by the particular spatial frequency under test). Here, as in the condition during which the subject sat still while his CSF was measured, either the subject reported seeing the grating pattern at some time during the trial or the trial was counted as below threshold, which meant that the next test on the same staircase would be made with a higher level of maximum contrast.

10.3. Results when head speed controlled contrast

Fig. 23 summarizes the psychophysical results obtained. The functions plotted in this figure should be compared with the functions for the same subject (RMS) in Fig. 20. The earlier data were obtained when this subject *could have* sneaked peeks near head-turnabouts as he judged high spatial frequency gratings and *could have* made threshold judgements near the center of head swings as he judged low spatial frequencies. The results of the new control experiment do not support these suggestions, particularly at the high spatial frequency end (above 8 cycles/° where the functions cross). Here, the detrimental effects of motion on contrast sensitivity are modest throughout the range of higher spatial frequencies (about 25% at and above 12 cycles/°). The extrapolated high-frequency cutoffs of both functions are also not very different, both falling well above 30 cycles/°. The beneficial effects of motion on low spatial frequency gratings were observed in a smaller range of spatial frequencies in the new control experiment than they were originally. In the original experiment, which is summarized in Fig. 20, this subject showed beneficial effects of motion at and below 8 cycles/°, whereas in the new experiment the benefits of motion only appeared at and below 4 cycles/°. The magnitude of the maximum beneficial effect of motion at the low frequency end was, however, quite comparable, namely a factor of about 2.5 at 2 cycles/°. The functions for both 'moving' and 'sitting still' were indistinguishable between 3 and 8 cycles/°, where they cross over as motion began to reduce contrast sensitivity.

It does not seem useful to try to explain the differences between the original and the sneak-a-peek control experiments, which are summarized in Figs. 20 and 23, because the new control experiment utilized a novel and very strange coupling of head motion and grating contrast, a coupling that could never occur under normal viewing. In other words, shaking the head to obtain contrast and shaking it faster to increase contrast precluded sneaking peeks but it does not provide a useful method for studying the normal relationship of head-movement-produced retinal-image slip to contrast sensitivity. We believe that the technique used in this control experiment seems no more likely to produce useful information about this relationship than imposing a unidirectional drift on a more-or-less stabilized grating display. The sneak-a-peek experiment did, however, serve its purpose inasmuch as it showed that the modest deleterious effects of image motion on high spatial frequencies observed in the previous experiment were not caused by the subject's carelessness, dereliction of duty or dishonesty.

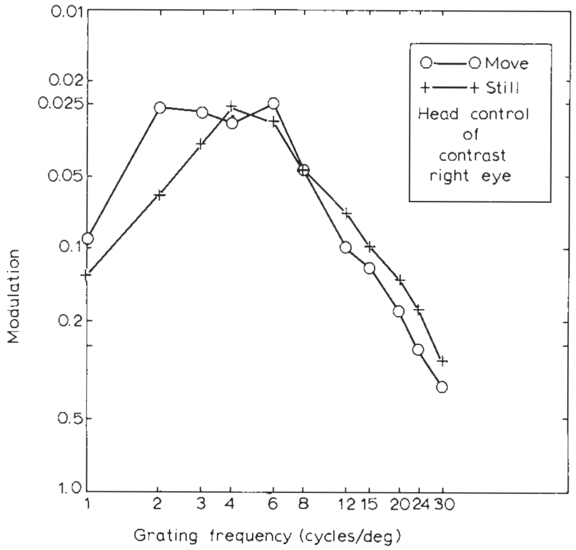

Fig. 23. Threshold grating contrast-modulation settings as fractions of 100% contrast at eleven spatial frequencies. These contrast sensitivity functions were obtained either while the subject (author RMS) sat still (Still), shown by the filled circles, or while he oscillated his head about its vertical axis (Move), shown by the open circles. When he oscillated his head, the speed with which the head was moving determined the maximum available contrast. See the text for a discussion of these functions.

10.4. Results when eye speed controlled contrast

The basic experiment was repeated with the sensor coil mounted on the eye, rather than on the head; a change which guaranteed that contrast sufficient to cause the perception of a grating would never be presented when the display was stationary on the retina. This change in the placement of the sensor coil only allowed a partial replication of the sneak-a-peek experiment because only relatively short psychophysical sessions were possible. Wearing the silicone annulus sensor coil restricted psychophysical sessions to 20 minutes.* Two-hour sessions had

---

* RMS is only comfortable for about 20 min when he wears the silicone annulus. Many other subjects are comfortable for much longer periods, up to 50 minutes, which is considered by many users to be the upper limit for the safe use of this kind of attachment to the eye (this limit is imposed by intraocular pressure, which increases slowly as the annulus is worn for prolonged periods).

been possible when the sensor coil was mounted on the head. This limitation made it relatively difficult to collect even the very modest set of threshold data that will be described below because it was deemed prudent to skip at least one day between sessions in addition to keeping the sessions very short. This restriction was partially overcome by using both of the eyes for making observations, one eye at a time in a single session and alternating eyes in successive sessions. The experiment was also simplified by reducing the number of spatial frequencies studied from 11 to 8 and by using a very modest number of reversals on the staircases to estimate thresholds, rather than continuing measurements until completely satisfied that the best possible estimate had been obtained in a given session. To sum up, using a silicone annulus sensor coil sucked on to the eye precluded the careful, well-replicated psychophysical measurements that were possible when the sensor coil was attached to the head.

The results of this control experiment are summarized in Fig. 24. The functions plotted are based on contrast sensitivity thresholds measured in 18 short sessions spread out over a period of more than 2 months (the data points are the means of RMS's right and left eye thresholds; his eyes had similar sensitivities). Measurements could not be extended above 20 cycles/° because the variability associated with contrast threshold measurements on the rapidly falling upper limb of the CSF precluded reasonable estimates with spatial frequencies above this value. In other words, sessions were too short to permit an acceptable estimate of sensitivity with the double random staircase we employed. Nevertheless, the, admittedly noisy, data summarized in Fig. 24 are, in our opinion, sufficiently clear to allow us to suggest that the main effects of motion on contrast sensitivity, which were measured when contrast was controlled by the speed of the head, can also be found in measurements made when contrast was controlled by the speed of retinal image (eye speed in space and retinal image speed are almost the same with distant targets; see Steinman et al., 1982; or Ferman et al., 1987). That is, the visibility of spatial frequency gratings below 3 cycles/° was enhanced by retinal image motion and retinal image motion had only relatively modest adverse effects on contrast sensitivity with higher spatial frequencies (above 6 cycles/° these modest effects of motion on contrast sensitivity were obscured by the noise of the psychophysical measurements). The basis upon which the visual system maintains contrast sensitivity in the presence of appreciable retinal image motion must still be worked out (see section 8.2.3 for our speculations about the way in which signals from the vestibular system might aid vision in the presence of retinal image motion).

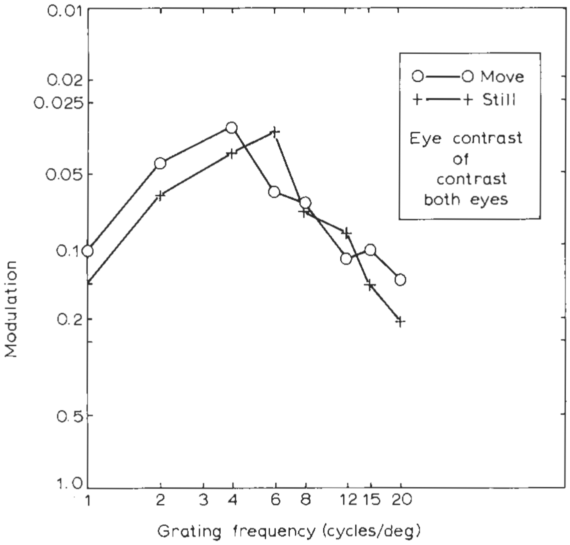

Fig. 24. Threshold grating contrast-modulation settings as fractions of 100% contrast at eight spatial frequencies. These contrast sensitivity functions were obtained either while the subject (author RMS) sat still (Still), shown by the filled circles, or while he oscillated his head about its vertical axis (Move), shown by the open circles. When he oscillated his head, the speed with which the retinal image was moving in either the right or the left eye determined the maximum available contrast. The functions during movement are the means of the threshold settings made with each eye at a separate session. See the text for a discussion of the significance of these functions and details about how these observations were made.

11. Conclusions

We began this chapter claiming that the role of eye movements in basic visual information processing has a long history. Their role has varied from virtually complete neglect in 'static' theories to fundamental significance in 'dynamic' theories. A variety of intermediate roles have also been proposed. The reader will probably agree that we have justified this claim. We also began by claiming that current emphasis favors an important role for eye movement in basic visual information processing. At this time it is generally accepted that transient stimulation, which may be produced by movements of the eye as well as by movements of objects in the visual field, can drive visual neurons whose consequent action allows us to see. Exactly what would be seen with high-contrast targets in the central fovea if all retinal image motion were to be eliminated is still unknown and alternative expectations are plausible, i.e., either a completely homogeneous visual field or a low-contrast visual field with only low or medium spatial frequency content. It looks as though it may be exceedingly difficult, probably actually impossible, to resolve this issue because of the exquisite sensitivity of the visual neurons to motion of any appreciable extent across the retinal surface and the incessant motion of visual neurons embedded in pulsating retinal tissues.

Fortunately, it has also become clear in recent years that the unsuccessful attempt to resolve this classical question during the past 40 years may represent an unnecessary and misdirected effort. The retinal image is in no danger, whatsoever, of even approaching the level of stability observed when the head is artificially supported once the subject gets off the biteboard. The head resting only on its natural biological platform moves quite a bit and this motion is not compensated completely by eye movements. Natural retinal image motion, arising from incomplete oculomotor compensation of bodily motions, guarantees appreciable transient stimulation of visual neurons (several orders of magnitude greater than the minimum retinal image displacement required to provide effective transient stimulation). This is true even in the limiting case in which the human being attempts to sit or stand as still as possible while fixating a stationary target located at optical infinity. Here, only about 50% of the movements of his head are compensated by eye movements, and retinal image motions are 2–4 times faster than retinal image motions present when the head is on a biteboard. Once the subject relaxes or looks at near targets, retinal image motion more than doubles. As soon as he talks or moves or chews or does anything at all, the problem for the visual scientist changes the traditional question of how eye movements are used to provide transient stimulation into a complementary question. Namely, how do we see a clear and stable world in the presence of such turbulence. Actually, the oculomotor system does its job very well; on average, about 97.5% of bodily motion is likely to be compensated for by eye movements when the head is free and the subject makes rather natural movements. But 2.5% of uncompensated bodily motion allows the retinal image to move through several degrees of visual angle at velocities of several degrees per second during most non-violent natural activities. Also, remember that the motions in each of the eyes are different. This allows vergence (and absolute retinal disparity) to vary continually by like amounts whenever natural bodily motion is permitted.

Some of this turbulence has been shown to be beneficial for vision. In essence, the incomplete nature of oculomotor compensation is an inescapable but useful characteristic. Furthermore, there are suggestions that the degree of oculomotor compensation seems to be tuned at both ends of its effective functional range so as to facilitate some properties of visual information processing. Specifically, oculomotor compensation of bodily motion is far from complete when the subject sits or stands still. This allows effective transient stimulation. Also, even when compensation increases as the subject starts to move enough turbulence remains to facilitate perception of the global features of visual forms because the low spatial frequency content of the visual world benefits from increasing motion of the visual scene on the retina. Even the relatively high spatial frequency content required to make out fine details has been found to suffer only moderately when images move as a consequence of incomplete oculomotor compensation of bodily motion. How vision of high spatial frequencies manages to resist degradation in the presence of these natural retinal image motions is the major problem remaining.

It seems likely that the relationship between characteristics of oculomotor compensation and requirements for the detection of contrast and spatial detail is built-in and is maintained automatically without any voluntary oculomotor acts. It would be very awkward if the human being had to intentionally move her eyes so as to generate transient neural signals required to detect contrast or to achieve resolution of fine details. Simply living without artificial bodily restraint guarantees retinal image motion sufficient to activate all basic visual processing mechanisms. In essence, living and seeing go together naturally. It is only when stimuli are very faint or when contrast is extemely low that active looking for seeing comes into play. These conditions are likely to be encountered at night, in dense fog, with objects far in the periphery or in the visual science laboratory. In other situations, the interplay of oculomotor control and the processing of basic visual information are tuned to each other reflexively and voluntary oculomotor actions are unnecessary. This arrangement, from a teleological perspective, is a simple and useful way for the visual and oculomotor systems to cooperate. Significant features of visual stimuli appear effortlessly. Voluntary action is required only when we wish to search the visual array to find or to contemplate conspicuous features contained within it. More basic visual processing works well without explicit, willful oculomotor intervention.

Acknowledgement

The preparation of this chapter was supported, in part, by Grant EY 04647 from the National Eye Institute of the National Institutes of Health (U.S.A.).

References

Adler, F.H. and Fliegelman, M. (1934) Influence of fixation on the visual acuity. Arch. Ophthalmol. 12, 475–483.

Arend, L.E. (1973) Spatial differential and integral operations in human vision: implications of stabilized image fading. Psychol. Rev. 80, 374–395.

Arend, L.E. (1976) Temporal determinants of the form of the spatial contrast threshold MTF. Vision Res. 16, 1035–1042.

Arend, L.E. and Timberlake, G.T. (1986) What is perfect stabilization? Do perfectly stabilized images always disappear? J. Opt. Soc. Am. A 3, 235–241.

Attneave, F. (1957) Physical determinants of the judged complexity of shapes. J. Exp. Psychol. 53, 221–227.

Averill, H.L. and Weymouth, F.W. (1925) Visual perception and the retinal mosaic. II. The influence of eye movements on the displacement threshold. J. Comp. Psychol. 5, 147–176.

Barlow, H.B. (1952) Eye movements during fixation. J. Physiol. 116, 290–306.

Barlow, H.B. (1963) Slippage of contact lenses and other artefacts in relation to fading and regeneration of supposedly stabilized images. Q. J. Exp. Psychol. 15, 36–51.

Barlow, H.B., Fitzhugh, R. and Kuffler, S.W. (1957) Change in organization in the cat's retina during dark adaptation. J. Physiol. 137, 338–354.

Blakemore, C. and Campbell, F.W. (1969) On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. J. Physiol. 203, 237–260.

Brindley, G.S. (1960) Physiology of the Retina and the Visual Pathway, Edward Arnold, London.

Brown, C.R. and Forsyth, D.M. (1959) Fusion contour for intermittent photic stimuli of alternating duration. Science 129, 390–391.

Bryngdahl, O. (1961) Effect of retinal image motion on visual acuity. Optica Acta 8, 1–16.

Burbeck, C.A. and Kelly, D.H. (1982) Eliminating transient artifacts in stabilized-image contrast thresholds. J. Opt. Soc. Am. 72, 1238–1243.

Burbeck, C.A. and Kelly, D.H. (1984) Role of local adaptation in the fading of stabilized images. J. Opt. Soc. Am. A 1, 216–220.

Byram, G.M. (1944) The physical and photochemical basis of visual resolving power. I. The distribution of illumination in retinal images. J. Opt. Soc. Am. 34, 571–591.

Campbell, F.W. and Robson, J. (1964) Application of Fourier analysis to the modulation response of the eye. J. Opt. Soc. Am. 54, 581.

Campbell, F.W., Kulikowski, J.J. and Levinson, J.Z. (1966) The effect of orientation on the visual resolution of gratings. J. Physiol. 187, 427–436.

Campbell, F.W., Carpenter, R.H.S., Levinson, J.Z. (1968) Visibility of aperiodic patterns compared with that of sinusoidal gratings. J. Physiol. 204, 283–298.

Clowes, M.B. (1962) A note on colour discrimination under conditions of retinal image constraint. Optica Acta 9, 65–68.

Collewijn, H. (1977) Eye and head movements in freely-moving rabbits. J. Physiol. 266, 471–498.

Collewijn, H. (1989) The vestibulo-ocular reflex: an outdated concept? In: J.H.J. Allum and M. Hulliger (Eds.), Afferent Control of Posture and Locomotion, Progress in Brain Research, Vol. 80, Elsevier, Amsterdam, pp. 197–209.

Collewijn, H. and Tamminga, E.P. (1984) Human smooth and saccadic eye movements during voluntary pursuit of different targets on different backgrounds. J. Physiol. 351, 21–250.

Collewijn, H., van der Mark, F. and Jansen, T.C. (1975) Precise recording of human eye movements. Vision Res. 15, 447–450.

Collewijn, H., Martins, A.J. and Steinman, R.M. (1981) Natural retinal image motion: origin and change. Ann. N. Y. Acad. Sci. 374, 312–329.

Collewijn, H., Martins, A.J. and Steinman, R.M. (1983) Compensatory eye movements during active and passive head movements: fast adaptation to changes in visual magnification. J. Physiol. 259–286.

Collewijn, H., Van der Steen, J. and Steinman, R.M. (1985) Human eye movements associated with blinks and prolonged eye-lid closures. J. Neurophysiol. 54, 11–27.

Collewijn, H., Erkelens, C.E. and Steinman, R.M. (1988) Binocular co-ordination of human vertical eye movements. J. Physiol. 404, 183–197.

Collewijn, H., Steinman, R.M., Erkelens, C.E., Pizlo, Z. and Van der Steen (1990a) The effect of freeing the head on eye movement characteristics during 3-D shifts of gaze. In: A. Berthoz, W. Graf and P.P. Vidal (Eds), The Head-Neck Sensory-Motor System, Oxford University Press, Wiley, New York, in press.

Collewijn, H., Steinman, R.M., Erkelens, C.E. and Regan, D. (1990b) Binocular fusion, steropsis and stereoacuity with a moving head. In: D. Regan (Ed.), Binocular Vision and Visual Dysfunction, Vol. 10A, Macmillan, London, in press.

Cornsweet, T.N. (1956) Determination of stimuli for involuntary drifts and saccadic eye movements. J. Opt. Soc. Am. 46, 987–993.

Cornsweet, T.N. (1962) The staircase method in psychophysics. Am. J. Psychol. 75, 485–491.

Cornsweet, T.N. and Crane, H.D. (1973) An accurate eye tracker using first and fourth Purkinje images. J. Opt. Soc. Am. 63, 921–928.

Craik, K.J.W. (1938) The effect of adaptation on differential brightness discrimination. J. Physiol. 92, 406–421.

Craik, K.J.W. (1939) The effect of adaptation upon visual acuity. Br. J. Psychol. 29, 252–266.

Crane, H.D. and Steele, C.M. (1978) Accurate three-dimensional eyetracker. Appl. Opt. 17, 691–705.

Crook, M.N. (1937) Visual discrimination of movement. J. Psychol. 3, 541–588.

Crozier, W.J., Wolf, E. and Zerrahn-Wolf, G. (1937) Critical illumination and critical flicker frequency for response to flickered light in dragonfly larvae. J. Gen. Physiol. 20, 363–392.

Cushman, W.B., Tangney, J.F., Steinman, R.M. and Ferguson, J.L. (1984) Characteristics of smooth eye movements with stabilized targets. Vision Res. 24, 1003–1009.

de Lange Dzn., H. (1957) Attenuation characteristics and phase-shift characteristics of the human fovea – cortex systems in relation to flicker-fusion phenomena. Thesis, Technical University, Delft.

de Lange Dzn., H. (1958) Research into the dynamic nature of the human fovea – cortex systems with intermittent and modulated light. J. Opt. Soc. Am. 48, 777–784.

Ditchburn, R.W. (1957) Report of N.P.L. symposium on colour vision. p. 415.

Ditchburn, R.W. (1973) Eye-Movements and Visual Perception, Clarendon Press, Oxford.

Ditchburn, R.W. (1980) The function of small saccades. Vision Res. 20, 271–272.

Ditchburn, R.W. and Drysdale, A.E. (1977a) The effect of retinal image motion on vision: I. Step-movements and pulse-movements. Proc. R. Soc. Lond. B 197, 131–144.

Ditchburn, R.W. and Drysdale, A.E. (1977b) The effect of retinal image motion on vision: II. Oscillatory movements. Pr. Roy. Soc. Lond. B 197, 385–406.

Ditchburn, R.W. and Ginsborg, B.L. (1953) Involuntary eye movements during fixation. J. Physiol. 119, 1–17.

Ditchburn, R.W. and Pritchard, R.M. (1956) Stabilized interference fringes on the retina. Nature 177, 434.

Dodge, R. (1903) Five types of eye movement in the horizontal meridian plane of the field of regard. Am. J. Physiol. 8, 307–329.

Dodge, R. and Cline, T.S. (1901) The angle velocity of eye movements. Psychol. Rev. 8, 145–157.

Duwaer, A.L. (1982) Assessment of retinal image displacement during head movement using an afterimage method. Vision Res. 22, 1379–1388.

Ercoles, A.M. and Zoli, M.T. (1968) Contrast thresholds for moving Landolt rings. Atti. Fond. G. Ronchi 23, 515–525.

Erkelens, C.E., Van der Steen, J., Steinman, R.M. and Collewijn, H. (1989a) Ocular vergence under natural conditions. I. Continuous changes of target distance along the median plane. Proc. R. Soc. Lond. B 236, 417–440.

Erkelens, C.E., Steinman, R.M. and Collewijn, H. (1989b) Ocular vergence under natural conditions. II. Gaze-shifts between real targets differing in distance and direction. Proc. R. Soc. Lond. B 236, 441–465.

Erkelens, C.E., Collewijn, H. and Steinman, R.M. (1989c) Asymmetrical adaptation of human saccades to to anisometropic spectacles. Invest. Ophthalmol. Vis. Sci. 30, 110–123.

Estevez, O. and Cavonius, C.R. (1976) Low-frequency attentuation in the detection of gratings: sorting out the artefacts.

Vision Res. 16, 497–500.

Falk, J.L. (1956) Theories of visual acuity and their physiological bases. Psychol. Bull. 53, 109–133.

Fechner, G.T. (1860) Elemente der Psychophysik, Breitkopf and Harterl, Leipzig.

Ferman, L., Collewijn, H., Jansen, T.C. and Van der Berg, A.V. (1987) Human gaze stability in horizontal, vertical and torsional direction during voluntary head movements, evaluated with a three-dimensional scleral induction coil technique. Vision Res. 27, 811–828.

Festinger, L. (1971) Eye movements and perception. In: P. Bach-y-Rita and C.C. Collins (Eds.), The Control of Eye Movements, Academic Press, New York.

Fiorentini, A. (1961) Dynamic characteristics of visual processes. In: E. Wolf (Ed.), Progress in Optics, North-Holland Press, Amsterdam.

Fiorentini, A. and Ercoles, A.H. (1957) Vision of oscillating non-uniform fields. Optica Acta 4, 150–157.

Graham, N. (1972) Spatial frequency channels in the human visual system: effects of luminance and pattern drift rate. Vision Res. 12, 53–69.

Gerrits, H.J.M. (1978) Differences in foveal and peripheral effects observed in stabilized vision. Exp. Brain Res. 32, 225–244.

Gerrits, H.J.M. and Vendrik, A.J.H. (1970) Artificial movements of a stabilized image. Vision Res. 10, 1443–1456.

Gerrits, H.J.M. and Vendrik, A.J.H. (1972) Eye movements necessary for continuous perception during stabilization of retinal images. Bibl. Ophthalmol. 82, 339–347.

Gerrits, H.J.M. and Vendrik, A.J.H. (1974) The influence of stimulus movements on perception in parafoveal stabilized vision. Vision Res. 14, 175–180.

Gerrits, H.J.M., Haan, B. de, Vendrik, A.J.H. (1966) Experiments with retinal stabilized images. Relations between the observations and neural data. Vision Res. 6, 427–440.

Gerrits, H.J.M., Stassen, H.P.W. and Erning, L.J. (1984) The role of drifts and saccades for the preservation of brightness perception. In: L. Spillmann and B.R. Wooten (Eds.), Sensory Experience, Adaptation, and Perception, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 439–459.

Gilbert, D.S. and Fender, D.H. (1969) Contrast thresholds measured with stabilized and non-stabilized sine-wave gratings. Optica Acta 16, 191–204.

Hartline, H.K. (1938) The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. Am. J. Physiol. 121, 400–415.

Hartline, H.K. (1940) The receptive field of the optic nerve fibers. Am. J. Physiol. 130, 690–699.

Hartridge, H. (1922) Visual acuity and the resolving power of the eye. J. Physiol. 57, 52–67.

Hebb, D.O. (1949) The Organization of Behavior, John Wiley and Sons, New York.

Hecht, S. (1927) A quantitative basis for the relation between visual acuity and illumination. Proc. Nat. Acad. Sci. USA 13, 569–574.

Hecht, S. (1928) The relation between visual acuity and illumination. J. Gen. Physiol. 11, 255–281.

Hecht, S. and Mintz, E. (1939) The visibility of single lines at various illuminations and the retinal basis of visual resolution. J. Gen. Physiol. 22, 593–612.

Hecht, S., Shlaer, S. and Pirenne, M.H. (1942) Energy, quanta, and vision. J. Gen. Physiol. 25, 819–840.

Heckenmueller, E.G. (1965) Stabilization of the retinal image: a review of method, effects, and theory. Psychol. Bull. 63, 157–169.

Helmholtz, H.L.F. von (1863) Die Lehre von den Tonempfindungen ais physiologische Grundlage fur die Theorie der Musik, First Edition, Braunschweig.

Helmholtz, H.L.F. von (1866) Handbuch der Physiologischen Optik, Vol. 2, Voss, Hamburg-Leipzig.

Hering, E. (1899) Über die Grenzen der Schschärfe. Ber. d. math.-phys. Kl. d. Königl. Geo. d. Wissensch. zu Leipzig, 16–24.

Hering, E. (1920) Grundzuge der Lehre vom Lichtsinn, Springer, Berlin.

Hirsch, J. and Hylton, R. (1984) Quality of the primate photoreceptor lattice and limits of spatial vision. Vision Res. 24, 347–355.

Hofmann, F.B. (1920) Die Lehre vom Raumsinn des Auges, Springer, Berlin.

Jones, L.A. and Higgins, G.C. (1947) Photographic granularity and graininess III. Some characteristics of the visual system of some importance in the evaluation of graininess and granularity. J. Opt. Soc. Am. 37, 217–263.

Keesey, U.T. (1960) Effects of involuntary eye movements on visual acuity. J. Opt. Soc. Am. 50, 769–774.

Kelly, D.H. (1979a) Motion and vision I. Stabilized images of stationary targets. J. Opt. Soc. Am. 69, 1266–1274.

Kelly, D.H. (1979b) Motion and vision II. Stabilized spatio-temporal transfer surface. J. Opt. Soc. Am. 69, 1340–1349.

Köhler, W. and Wallach, H. (1944) Figural after-effects: an investigation of visual processes. Proc. Am. Philos. Soc. 88, 269–357.

Khurana, B. and Kowler, E. (1987) Shared attentional control of smooth eye movement and perception. Vision Res. 27, 1603–1618.

Kowler, E. and Steinman, R.M. (1979) The effect of expectations on slow oculomotor control: I. Periodic target steps. Vision Res. 19, 619–632.

Kowler, E. and Steinman, R.M. (1980) Small saccades serve no useful purpose: reply to a letter by R.W. Ditchburn. Vision Res. 20, 273–276.

Kowler, E., Martins, A.J. and Pavel, M. (1984a) The effect of expectations on slow oculomotor control – IV. Anticipatory smooth eye movements depend on prior target motions. Vision Res. 24, 197–210.

Kowler, E., Van der Steen, J. and Collewijn, H. (1984b) Voluntary selection of the target for smooth eye movement in the presence of superimposed, full-field stationary and moving stimuli. Vision Res. 24, 1789–198.

Krauskopf, J. (1957) Effect of image motion on contrast thresholds for maintained vision. J. Opt. Soc. Am. 47, 740–745.

Krauskopf, J. (1962) Effect of target oscillation on contrast resolution. J. Opt. Soc. Am. 52, 1306.

Krauskopf, J. (1963) Effect of retinal image stabilization on the appearance of heterochromatic targets. J. Opt. Soc. Am. 53, 741–744.

Kuffler, S.W. (1953) Discharge patterns and functional organization of the mammalian retina. J. Neurophysiol. 16, 37–68.

Le Grand, Y. (1967) Form and Space Vision. Revised Edition translated by M. Millodot and G. Heath, Indiana University Press, Bloomington and London.

Levinson, J.Z. (1959) Fusion of complex flicker. Science 130, 919–921.

Levinson, J.Z. (1960) Fusion of complex flicker. II. Science 131, 1438.

Levinson, J.Z. (1966) One-stage model for visual temporal integration. J. Opt. Soc. Am. 56, 95–97.

Levinson, J.Z. (1968) Flicker fusion phenomena. Science 160, 21–28.

Lit, A. (1968) Visual acuity. Annu. Rev. Psychol. 19, 27–54.

Ludvigh, E. (1948) The visibility of moving objects. Science 108, 63–64.

McCree, K.J. (1960) Colour confusion produced by voluntary fixation. Optica Acta 7, 281–290.

Marshall, W.H. and Talbot, S.A. (1942) Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. Biol. Symp. 7, 117–164.

Matin, E. (1974) Saccadic suppression: a review and an analysis. Psychol. Bull. 81, 451–461.

Müller, J. (1826) Zur vergleichenden Physiologie der Geisichtssinnes der Menchen und Thiere, Leipzig.

Murphy, B.J. (1978) Pattern thresholds for moving and stationary gratings during smooth eye movements. Vision Res. 18, 521–530.

Murphy, B.J., Kowler, E. and Steinman, R.M. (1975) Slow oculomotor control in the presence of moving backgrounds. Vision. Res. 15, 1263–1268.

Nachmias, J. (1958) Brightness and visual acuity with intermittent illumination. J. Opt. Soc. Am. 48, 726–730.

Nachmias, J. (1959) Two-dimensional motion of the retinal image during monocular fixation. J. Opt. Soc. Am. 49, 901–908.

Nachmias, J. (1961) Determiners of drift of the eye during monocular fixation. J. Opt. Soc. Am. 51, 761–766.

Nachmias, J. and Steinman, R.M. (1965) An experimental comparison of the method of limits and the double-staircase method. Am. J. Psychol. 78, 112–115.

Osgood, C.E. (1953) Method and Theory in Experimental Psychology, Oxford University Press, New York.

Osgood, C.E. and Hyer, A.W. (1952) A new interpretation of figural aftereffects. Psychol. Rev. 59, 98–118.

Penner, M.J. (1978) Psychophysical methods and the minicomputer. In: M. Mayzner and T. Dolan (Eds.), Minicomputers in Sensory and Information Processing Research, Lawrence Erlbaum Associates, Hillsdale, NJ.

Polyak, S.L. (1941) The Retina, University of Chicago Press, Chicago.

Puckett, J. De W. and Steinman, R.M. (1969) Tracking eye movements with and without saccadic correction. Vision Res. 9, 695–703.

Ratliff, F. (1952) The role of physiological nystagmus in monocular acuity. J. Exp. Psychol., 43, 163–172.

Ratliff, F. (1965) Mach Bands: Quantitative Studies on Neural Networks in the Retina, Holden-Day, San Francisco.

Ratliff, F. (1984) Why Mach bands are not seen at the edges of a step. Vision Res. 24, 163–166.

Ratliff, F. and Riggs, L.A. (1950) Involuntary motions of the eye during monocular fixation. J. Exp. Psychol. 40, 687–701.

Riggs, L.A. (1965) Visual acuity. In: C.H. Graham (Ed.), Vision and Visual Perception, John Wiley and Sons, New York, pp. 321–349.

Riggs, L.A. and Armington, J.C. (1952) Angular displacements of the eye during prolonged fixation. Am. Psychol. 7, 252.

Riggs, L.A. and Ratliff, L.A. (1951) Visual acuity and the normal tremor of the eye. Science 106, 107–108.

Riggs, L.A. and Schick, A.M.L. (1968) Accuracy of retinal image stabilization achieved with a plane mirror on a tightly fitting contact lens. Vision Res. 8, 159–169.

Riggs, L.A., Ratliff, F., Cornsweet, J.C. and Cornsweet, T.N. (1953) The disappearance of steadily fixated visual test objects. J. Opt. Soc. Am. 43, 495–501.

Riggs, L.A., Armington, J.C. and Ratliff, F. (1954) Motions of the retinal image during fixation. J. Opt. Soc. Am. 44, 315–321.

Sachs, M., Nachmias, J. and Robson, J. (1971) Spatial frequency channels in human vision. J. Opt. Soc. Am. 61, 1176–1186.

Schade, O.H. (1956) Optical and photoelectric analog of the eye. J. Opt. Soc. Am. 46, 721–739.

Senders, V.L. (1948) The physiological basis of visual acuity. Psychol. Bull. 45, 465–490.

Senders, V.L. (1949) Visual resolution with periodically interrupted light. J. Exp. Psychol. 40, 453–465.

Shannon, C.E. (1948) A mathematical theory of communication. Bell System Tech. J. 27, 379–423; 623–656.

Skavenski, A.A., Hansen, R., Steinman, R.M. and Winterson, B.J. (1979) Quality of retinal image stabilization during small natural and artificial body rotations in man. Vision Res. 19, 365–375.

Stark, L. (1983) Normal and abnormal vergence. In: C.M. Schor and K.J. Ciuffreda (Eds.), Vergence Eye Movements: Basic and Clinical Aspects, Butterworths, Boston, pp. 3–13.

St.-Cyr, G.J. and Fender, D.H. (1969) Nonlinearities of the human oculomotor system: gain. Vision Res. 9, 135–1246.

Steinman, R.M. (1975) Oculomotor effects on vision. In: P. Bach-y-Rita and G. Lennerstrand (Eds.), Basic Mechanisms of Ocular Motility and Their Clinical Implications, Pergamon Press, Oxford, pp. 395–416.

Steinman, R.M. (1976) Role of eye movements in maintaining a phenomenally clear and stable world. In: R.A. Monty and J.W. Senders (Eds.), Eye Movements and Psychological Processes, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 121–149.

Steinman, R.M. (1986a) The need for an eclectic, rather than a systems, approach to the study of the primate oculomotor system. Vision Res. 26, 101–112.

Steinman, R.M. (1986b) Eye movement. Vision Res. 26, 1389–1400.

Steinman, R.M. and Collewijn, H. (1978) How our two eyes are held steady. J. Opt. Soc. Am. 68, 1359.

Steinman, R.M. and Collewijn, H. (1980) Binocular retinal image motion during natural active head rotation. Vision Res. 20, 415–429.

Steinman, R.M., Cunitz, R.J., Timberlake, G.T. and Herman, M. (1967) Voluntary control of microsaccades during maintained monocular fixation. Science 155, 1577–1579.

Steinman, R.M., Haddad, G.M., Skavenski, A.A. and Wyman, D. (1973) Miniature eye movement. Science 181, 810–819.

Steinman, R.M., Cushman, W.B. and Martins, A.J. (1982) The precision of gaze. Hum. Neurobiol. 1, 97–109.

Steinman, R.M., Levinson, J.Z., Collewijn, H. and Van der Steen, J. (1983) Vision in the presence of known natural retinal image motion. J. Opt. Soc. Am. 73, 1856.

Steinman, R.M., Levinson, J.Z., Collewijn, H. and Van der Steen, J. (1985) Vision in the presence of known natural retinal image motion. J. Opt. Soc. Am. A 2, 226–233.

Stork, D.G., Falk, D.S. and Levinson, J.Z. (1985) Receptive field asymmetry probed using converging gratings. J. Opt. Soc. Am. A 2, 275–279.

Talbot, S.A. and Marshall, W.H. (1941) Physiological studies of neural mechanisms of visual localization and discrimination. Am. J. Ophthal. 24, 1255–1264.

Tulunay-Keesey, U. and Jones, R.M. (1976) The effect of micromovements of the eye and exposure duration on contrast sensitivity. Vision Res. 16, 481–488.

Van Nes, F.L. (1968) Enhanced visibility by regular motion of the retinal image. Am. J. Psychol. 81, 367–374.

Van Nes, F.L. and Bouman, M.A. (1967) Spatial modulation transfer of the human eye. J. Opt. Soc. Am. 57, 401–406.

Van Nes, F.L., Koenderink, J.J., Nas, H. and Bouman, M.A. (1967) Spatiotemporal modulation transfer in the human eye. J. Opt. Soc. Am. 57, 1082–1088.

Volkmann, A.W. (1863) Physiologische Untersuchungen im Gebiete der Optik, Breitkopt and Hartel, Leipzig.

Volkmann, P.C. (1986) Human visual suppression. Vision Res. 26, 1401–1416.

Wald, G. (1948) Selig Hecht (1892–1947). J. Gen. Physiol. 32, 1–16.

Walls, G.L. (1943) Factors in human visual resolution. J. Opt. Soc. Am. 33, 487–505.

Walls, G.L. (1962) The evolutionary history of eye movements. Vision Res. 2, 69–80.

Westheimer, G. (1960) Modulation thresholds for sinusoidal light distributions on the retina. J. Physiol. 152, 67–74.

Westheimer, G. (1965) Visual acuity. Annu. Rev. Psychol. 16, 359–380.

Westheimer, G. (1981) Visual hyperacuity. Prog. Sensory Physiol. 1, 1–30.

Westheimer, G. and McKee, S. (1975) Visual acuity in the presence of retinal image motion. J. Opt. Soc. Am. 65, 847–850.

Weymouth, F.W., Andersen, E.E. and Averill, H.L. (1923) Retinal mean local sign; a new view of the relation of the retinal mosaic to visual perception. A.J. Physiol. 63, 410–411.

Wilcox, W.W. and Purdy, D. McL. (1933) Visual acuity and its physiological basis. Br. J. Psychol. 23, 233–261.

Williams, D.R. (1985) Aliasing in human foveal vision. Vision Res. 25, 195–206.

Wilson, V.J. and Melvill Jones, G. (1979) Mammalian Vestibular Physiology, Plenum Press, New York.

Winterson, B.J. and Collewijn, B.J. (1976) Microsaccades during finely-guided visuomotor tasks. Vision Res. 16, 1387–1390.

Woodworth, R.S. (1938) Experimental Psychology, Holt, New York.

Wülfing, E.A. (1892) Über den kleinsten Gesichtwinkel. Z. Biol. 29, 199–202.

Yarbus, A.L. (1957a) A new method for studying the activity of various parts of the retina. Biophysics 2, 165–167.

Yarbus, A.L. (1957b) The perception of an image fixed with respect to the retina. Biophysics 2, 683–690.

Yarbus, A.L. (1967) Eye Movements and Vision. Translated by B. Haigh and L.A. Riggs, Plenum Press, New York.

Yellott, J.I. Jr. (1983) Spectral consequences of photoreceptor sampling in the rhesus retina. Science 221, 382–385.