Eye Fixations Recorded on Changing Visual Scenes by the Television Eye-Marker

J. F. MACKWORTH AND N. H. MACKWORTH
Medical Research Council, Applied Psychology Research Unit, Cambridge, England

(Received February 4, 1958)

By the use of television techniques it is possible to record the position of a man's gaze upon a picture of the scene at which he is looking. The corneal reflection of a light is picked up by a television camera, which magnifies the movement of the spot about 100 times. The scene at which the man is gazing is provided by a second television camera, and the corneal reflection is superimposed upon another television monitor also showing the same scene. By suitable calibration, this spot can be made to lie upon that part of the scene being regarded and the accuracy with which this can be done is within one or two degrees.

Thus it is possible to see where a man is looking at any moment, and simultaneously to record the composite eye-scene picture by means of a motion picture camera. In addition to its practical uses, this method is particularly useful for answering such general questions as whether a subject always sees what he is directly looking at, what kind of cues catch his attention when he is searching, and what sort of changes in a display catch his attention. It is primarily intended for use with a moving display, as it records the point of view directly on to each state of the display. Successive glances therefore indicate the order in which particular parts of complex displays are selected, as well as the exact moment at which the gaze changes to a new position.

GENERAL PROBLEM

THERE is very often a need to know where a man is looking when he is watching a moving visual scene. Under laboratory conditions it is now possible to mark where the eyes are fixating by recording the successive positions directly on to a flat picture of the visual world. The problem has been to devise a way of detecting the regions of the scene that are being examined while the viewer scrutinizes a changing display.

Basically the idea is to do this by using two closed-circuit television cameras which together produce one composite picture. One television camera gives an image of the subject's visual world, and the other camera superimposes on this same display, but on a second screen, a bright patch from the corneal reflection which acts as an eye-marker to indicate the present region of regard. The key details being chosen by the eyes at any given moment from the complicated picture now on view are therefore constantly being identified by the eye-marker dancing across the visual scene.

Few people believe that visual perception depends only on the information received currently by the eyes; the brain of the beholder contributes just as much if not more. Nevertheless there are many forms of work which depend on a steady flow of exact information obtained visually from the environment by detailed signals that are rather small in size. The line of sight has to be within a degree or two of such stimuli for any very exact information about them to be available to the subject, even when fully alert.[1] Looking is not always seeing, but the eye-marker method will at least tell us which of the many visual cues are being discarded, even if it cannot indicate which signals are actually being perceived.

The following are some of the requirements met by this method.

1. The viewing of continuously changing scenes can be analyzed whether these are of real objects or images on motion picture or television screens.
2. The qualitative relationship between the line of regard and particular aspects of the external visual world are established at once in real time during the experiment, without subsequent tedious computation or special projection arrangements. This avoids the errors inherent in correlating eye and scene recordings by subsequent calculation.
3. An important advantage is that a constant watch can be kept on where the subject is looking during the actual visual searching, because the experimenter has this large view finder in the form of a recording television screen displaying the combined eye-scene picture which gives him immediate feedback by direct observation. This speeds up the detection and correction of any recording artifacts. It also enables any particularly interesting sequences in the composite eye-scene picture to be selected for permanent record by means of a 16-mm motion picture camera.
4. The exact quantitative analysis of the fixation records can therefore be undertaken later by accurate measurement within each frame of this eye-scene film record.
5. The corneal reflection method has been chosen because it is more suited to the testing of relatively large numbers of subjects than the much more accurate procedures based on the use of a corneal contact lens mirror.

SOME PREVIOUS CORNEAL REFLECTION STUDIES

Just after the fiftieth anniversary of the major publication of a new method is possibly an appropriate time to suggest yet another modification of a well-known basic technique—despite the fact that during

[1] S. B. Williams, J. Opt. Soc. Am. 39, 782 (1949).

the last half-century many investigators have already adapted the Dodge corneal reflection method of eye movement recording in so many different ways.[2–7] Dodge[2] mentioned that the corneal reflection technique measured visual fixations in the horizontal plane with errors of less than 10 minutes of arc, and even the simplified modern commercial ophthalmograph will register movements of 20–30 minutes of arc.[8]

In the recording of voluntary eye movements, the smallest movement that usually has to be detected is of the order of one degree of arc: ideally the system should be able to include a movement up to 35–40 degrees. Although a 90-degree horizontal arc is possible by eye movements alone, it is after this 35- to 40-degree total horizontal arc that head movements usually take over.[9] For example, the lines on a quarto-sized typescript are about 7 inches wide, and these are read by eye movements alone from a distance of 14 inches, since they subtend a total horizontal arc of only 28 degrees.

Dearborn[4] soon introduced the technique of using a continuously moving strip of film as recording material instead of the original falling photographic plate; and this improvement gave either horizontal or vertical eye movements on separate film strips. Miles and Shen[10] managed to record horizontal and vertical eye movements successively on the same wide film strip by running their recording material twice through the camera. The simultaneous measurement of horizontal and vertical movement on two separate but synchronized moving film strips was first achieved by Jasper and Walker.[11] Buswell[12] also used two separate films continuously moving at right angles to each other. Then Brandt[13] published an account of his important advance. This was a 35-mm film camera which recorded both horizontal and vertical movements on the same film strip—although the recording of the horizontal was not on the same part of the film as the vertical record.

One major difficulty with all these methods has been to relate the physiological records of fixation to the visual scene, despite the fact that much of the work to date has been with the easier problem of static rather than moving scenes. To discover where the subject was looking at any given moment, it was necessary to re-project the X-axis record and the Y-axis record either on to a large reference grid for later transfer to the original scene after computation, or directly on to a facsimile of the original static picture. Although tedious, this method was however undoubtedly feasible; Buswell[12] for example, worked through 18 000 feet of recording film in this way during his classic analysis of how people look at static pictures. Again Brandt[14–16] used this approach to good effect in analyzing the attentional value of size and the preference for the upper left-hand part of the field of view in symmetrical visual displays.

Wendt[17] also adopted a similar plan for the viewing of continuously changing scenes and modified Brandt's camera for the analysis of eye movements during the viewing of motion pictures. The recordings were analyzed frame by frame by projection on to a large white screen; and each fixation point was marked by a pin, the distances being measured by ruler. These coordinates were written down and related to the moving scenes by counting individual frames along from the start of the 24-frame per second stimulus movie and the 6-frame per second record movie. Rather dissatisfied with this method, Wendt[17] called for an improved procedure. This required both the stimulus and recording motion pictures to be at the same speed of 24 frames per second. For analysis, both the films would then be projected simultaneously on the screen, "thus superimposing the actual fixation dot on the motion picture scene which caused this fixation. . . . The effect on the screen would be that of a moving fixation dot moving about on the screen exactly in the same way in which the subject viewed the teaching film." The present method meets these requirements.

METHOD

(a) Televising the Corneal Reflection

The corneal reflection itself can readily be displayed on the screen of a small closed-circuit television camera, the lens of which is close to the eye. Figures 1 and 2 show the general arrangements whereby a standard Pye industrial television camera (Type 2026/D) is provided with an unusually long telescopic lens tube which brings the $f/1.9$ 2-inch Watson lens to about 10 inches in front of the normal lens position, the lens being mounted in reverse.

The subject sits and looks straight forward; he bites on a firm support, consisting of plastic molded to his teeth and attached to an adjustable, rigid support. The back of his head is also well supported by a curved headrest fastened to the back of the chair. Cheek-bone supports (not shown) attached either to the table or chair prevent sideways movement in trained subjects. The

[2] R. Dodge, Psychol. Monogr. No. 35 (1907).
[3] R. S. Woodworth and H. Schlosberg, *Experimental Psychology* (Henry Holt and Company, New York, 1954).
[4] L. Carmichael and W. F. Dearborn, *Reading and Visual Fatigue* (Houghton Mifflin Company, Boston, 1947).
[5] M. D. Vernon, *The Experimental Study of Reading* (Cambridge University Press, New York, 1931).
[6] M. Cesa-Bianchi, Giorn. psichiat e neuropatol. 4, 1 (1955).
[7] L. A. Riggs and F. Ratliff, J. Exptl. Psychol. **40**, 687 (1950).
[8] Riggs, Armington, and Ratliff, J. Opt. Soc. Am. **44**, 315 (1954).
[9] H. C. Weston, *Sight, Light and Efficiency* (H. K. Lewis, London, 1949).
[10] W. R. Miles and E. Shen, J. Exptl. Psychol. **8**, 344 (1925).
[11] H. H. Jasper, and R. Y. Walker, Science **74**, 291 (1931).
[12] G. T. Buswell, *How People Look at Pictures* (University of Chicago Press, Chicago, 1935).
[13] H. F. Brandt, Am. J. Psychol. **49**, 666 (1937).
[14] H. F. Brandt, Am. J. Psychol. **53**, 564 (1940).
[15] H. F. Brandt, Am. J. Psychol. **53**, 260 (1940).
[16] H. F. Brandt, *The Psychology of Seeing* (Philosophical Library, Inc., New York, 1945).
[17] P. R. Wendt, Psychol. Monogr. **66**, No. 339 (1952).

importance of this rigid head fixation is seen in the fact that the required minimum registerable eye movement of one degree of arc could easily be confused with an artifact from a head movement; such artifacts can arise from head movements as small as 0.075 millimeter.[18]

The television camera, mentioned above, is placed at 45 degrees to the right of the line of regard and its lens is at eye level, two inches away from the right eye (Fig. 1). At an angle of 45 degrees to the left of the line of regard, also at eye level, is the light source which is 20 inches from the right eye. This is a Pointolite 150 candle-power Ediswan bulb, close behind an iris aperture opened to a diameter of about three-quarters of an inch (Fig. 1). Although this light falls on the right eye, neither the light source nor the lens of the eye camera is distracting to the subject during visual search because both these objects are well to one side of his line of regard. It is however necessary to have the important items at which the subject is looking sufficiently brightly lit.

The magnification required is great, and is of the order of X 100 because a horizontal movement of the eye through 20 degrees gives an apparent lateral movement of about 2 millimeters of the corneal reflection; this becomes approximately 20 centimeters movement across the television monitor screen. Such considerable enlargement demands particular care in the sighting of the television camera, which must be focused accurately on the corneal reflection. Campbell[19] has indicated a way in which this can be done. After switching off the television camera, a small lamp is placed inside the long lens tube of the eye camera. The camera is more or less correctly aligned when this extra light is sharply focused to cover the same apparent position on the right eye as that of the corneal reflection. This aiming light is then removed from the camera tube, and after switching on the television camera, the corneal reflection will either appear on the television screen, or can be made to do so if the subject moves his eyes around a

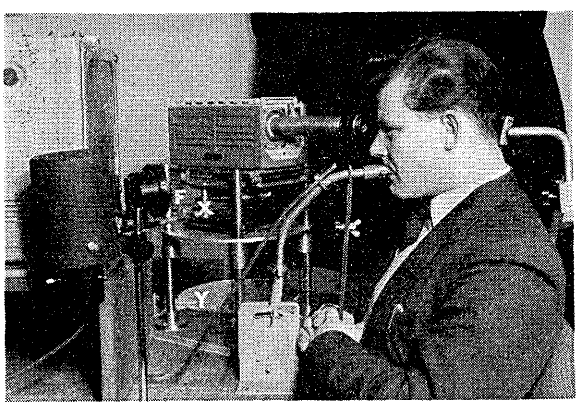

FIG. 1. General arrangement for eye camera.

[18] R. W. Ditchburn and B. L. Ginsborg, J. Physiol. **119**, 1 (1953).
[19] F. W. Campbell (personal communication, 1957).

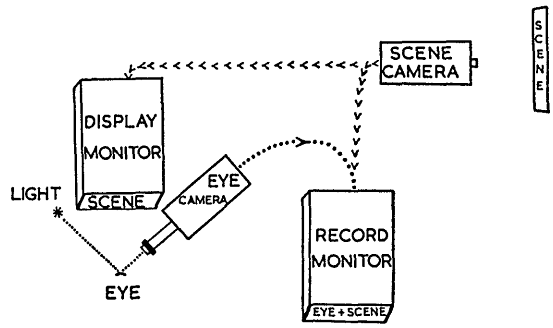

FIG. 2. Schematic vertical view.

little. The iris on the camera lens is then closed until the eye-maker is about one centimeter or less in diameter on the television screen. The camera is now ready for its final alignment which is described below. This third stage entails a most exact superimposition of the two pictures, one representing the eye-marker position in a blank frame, the other the visual scene itself.

These fine adjustments of the eye camera position clearly demand a heavy mounting, and accurate but light screw-thread controls, e.g., in Fig. 1 the Controls X and Y alter the horizontal and vertical positions of the eye-marker on the television screen and Control F adjusts the optical focus.

(b) Televising the Visual Scene

A second television camera is used to provide the visual scene both to the television screen in front of the subject and simultaneously to the other monitor displaying the eye-scene picture for record purposes. In practice, a Pye Image Orthicon camera (Type 2011/A/8) was used for this, with a 4-inch Dallmeyer Super-Six $f/1.9$ lens. This double-monitor plan, one for the subject's display without the eye-marker and the other for the record, is a convenient way of standardizing the subject's viewing conditions for a wide range of visual material. The viewing distance from the eyes to the center of this vertical display screen is 20 inches, and the effective width of the screen is 10 inches. All seeing can therefore be done by horizontal movements within a total arc of 28 degrees. The vertical height of the display is eight inches which represents a vertical arc of 23 degrees. The screen is placed directly in front of the subject and the center of the screen is at eye level.

(c) Superimposing the Eye-Marker on the Visual Scene

The best way of combining the picture of the eye-marker with that of the visual scene is by means of an electronic mixing unit so that the images from each of the two television cameras are fed to the same recording screen. This combined eye-scene picture can then be directly viewed by the experimenter and recorded by a 16-millimeter motion picture camera. If

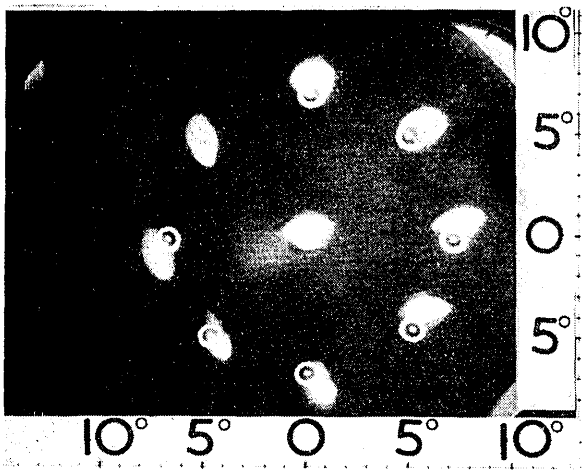

FIG. 3. Calibration scene with eye-markers.

run at 6 frames a second, this film should register nearly all the voluntary fixations, but it will not record the involuntary eye tremors during fixation. (These tremors convert the point of view into an area of regard, but this circle of uncertainty is not likely to be more than half a degree in diameter.[20])

The movements of the eye-marker are adjusted so that when the subject looks at known points in the visual scene, the marker comes to lie over these points in the combined eye-scene picture. This can be done, as in Fig. 3, by using a scene showing circles at the eight compass points about a center circle, or alternatively by a square of 25 letters. The extent of the movement of the eye-marker for a given arc of eye movement may have to be altered for different subjects by adjusting the length of the telescopic lens tube on the eye camera. It is possible to compensate for lack of linearity in the movement of the eye-marker to some extent by adjusting the controls of the monitor at which the subject is looking. It is usually found to be necessary

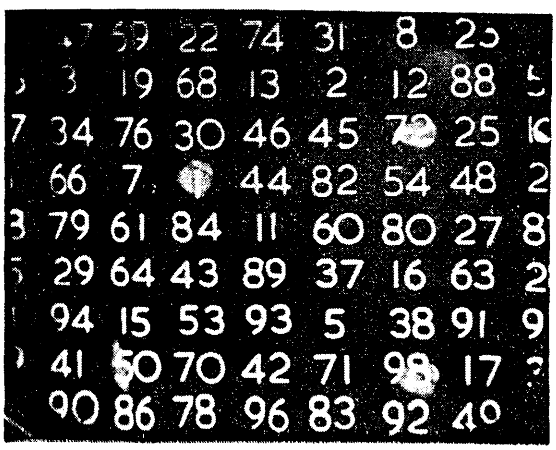

FIG. 4. Looking at four successive numbers.

[20] Ditchburn, Fender, and Mayne (personal communication, 1957).

to turn the monitor slightly so that the left side of the monitor is an inch or so nearer to the subject than the right side. Once these adjustments have been made for any one subject, no further change should be necessary for that subject, whatever display is used. Given the plastic mouth support, the eye-marker for any subject can be located on the test scene in only a minute or two.

As all investigators have found, very small head movements can introduce serious errors into corneal reflection recordings. Fortunately, this present method enables the position of the head to be checked, as the subject can be asked to fixate a central spot at intervals and any movement of the eye-marker in relation to this central spot indicates a head movement.

RESULTS AND ILLUSTRATIVE USES

In general, the technique appears to be useful in three broad categories of human performance: (a) visual aiming; (b) visual search; and (c) thinking.

(a) Visual Aiming

An analysis can be made of tasks consisting largely of visual aiming or the following of moving objects with the eye. Figure 3 is an unretouched photograph indicating the accuracy with such aiming can be achieved. The visual scene consisted of these eight circular targets arranged round the central target—and the eye-markers represent successive fixations on these targets. It is clear that the accuracy with a trained subject is such that the marker is within one to two degrees of the target. This can be achieved over a test period of 15 to 20 minutes without any recentering of the eye-marker on the central calibration target, i.e., after an initial settling spell of two minutes. This small discrepancy remaining between eye-marker and target, seen in Fig. 3, is biological in origin; there is no detectable drift in the physical system when a fixed artificial eye is reflecting the light spot into the camera. This residual biological error arises partly from a nonlinearity in the optical system of the cornea and partly from very slight head movements.

(b) Visual Search

Figure 4 shows the possible use of the eye-marker in the experimental analysis of visual search tasks, especially in complex visual presentations with highly detailed numerical displays. (This illustration is on the same scale as that indicated in Fig. 3.)

This example can be used to demonstrate a common difficulty in using such displays that the one wanted signal resembles many other unwanted signals. (In this case the subject was looking for a given number on four successive occasions.) Such situations sometimes arise on the monitor screens displaying the data held in store by an automatic computer. These details could therefore refer to anything from a filtered combat situation (expressed either in tabular form or as a tactical display)

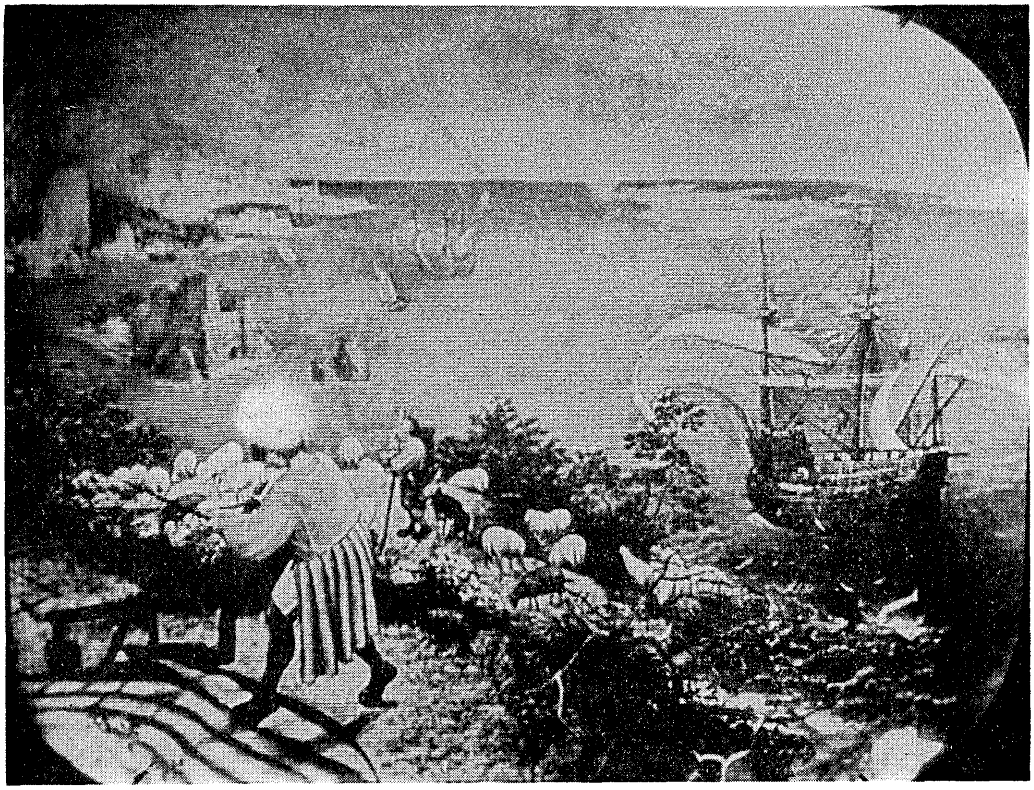

FIG. 5. The eye-marker on a static picture.

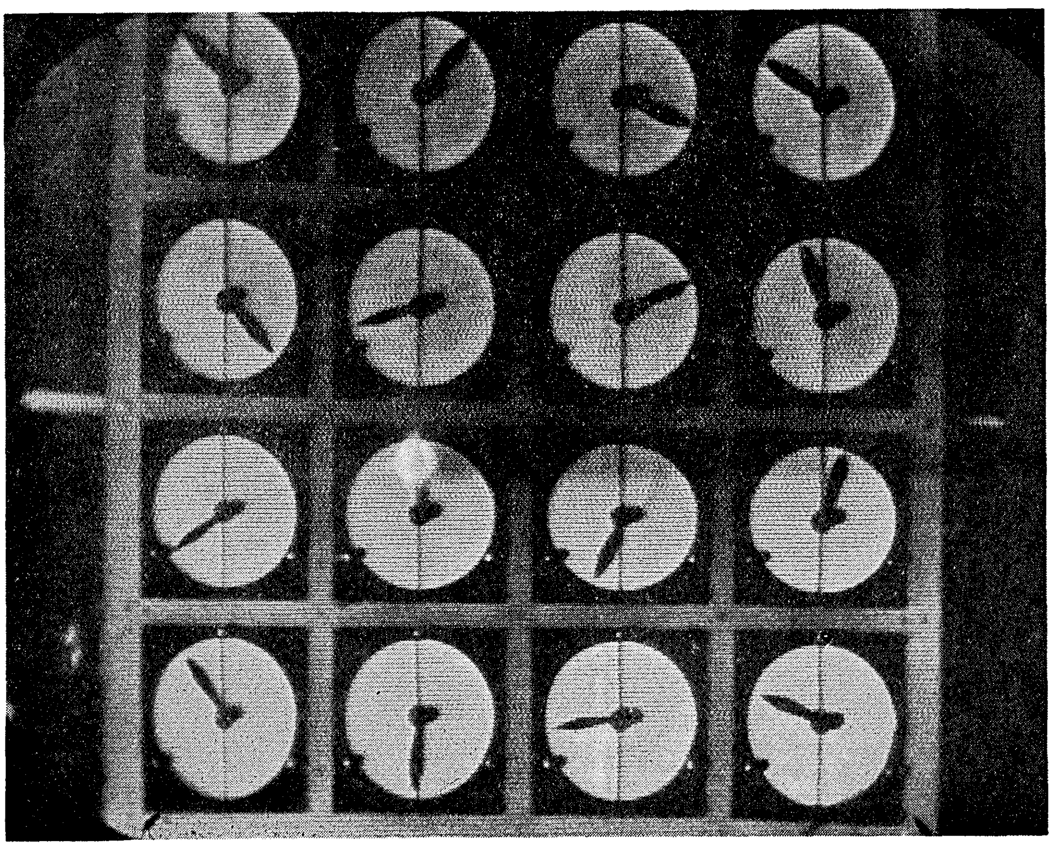

FIG. 6. The eye-marker on a changing scene.

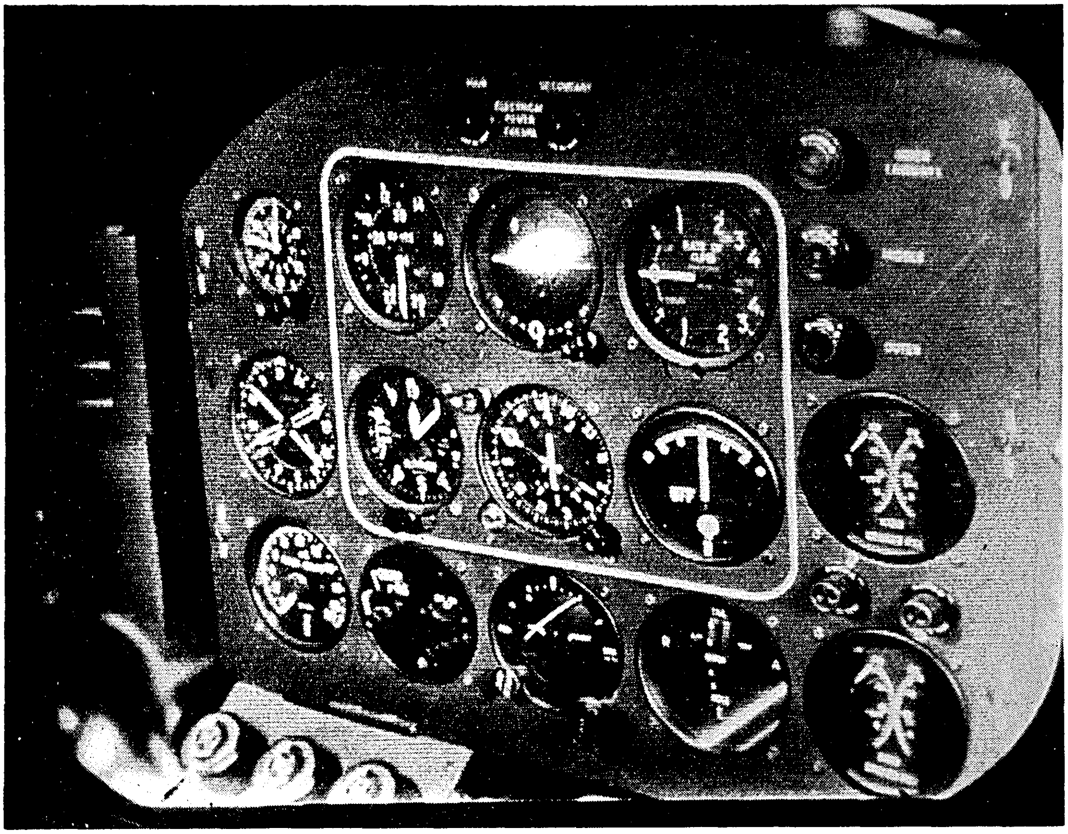

FIG. 7. Selecting from a moving display—Britannia instrument panel.

to the present state of the production line in an automatic factory.[21]

Visual search studies could also analyze the search patterns used by skilled industrial inspectors, and people extremely able at detecting flaws. Analysis has shown that in search situations such as Fig. 4 the subject can look directly at the number for which he is searching without seeing it. Indeed, this looking without seeing can be found even in circumstances when the subject actually pauses on the required number without reporting its presence by pointing to it.

One particular advantage of the method is that it reveals the sequence in which the subject scans the separate symbols in the display, and also the time relationships between these shifts in the gaze. This is perhaps less useful in a static picture such as Fig. 5 than in rapidly changing scenes such as Fig. 6. (This multidial display in Fig. 6 needs fast scanning if the subject is to press the correct switch just before one of these pointers crosses the vertical line.) Figures 5 and 6 also suggest the general point that perhaps eye-marker records are particularly useful in analyzing the viewing of the highly symbolized and disconnected moving displays of modern technology,[22] although there are thought to be some applications in aiding the understanding of seeing when the subject is viewing rather more natural scenes.

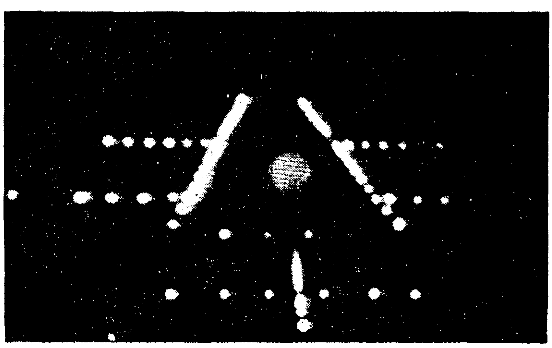

FIG. 8. Selecting from a changing pattern—night landing movie.

(c) Thinking

It is possible to record the input selected by the subject from the vast complexity of his visual world when he is thinking out the answer (or series of answers) to certain intellectual situations. Carmichael[23] has

[21] N. H. Mackworth and J. F. Mackworth, Brit. J. Psychol. 49, 89 (1958).

[22] J. F. Mackworth and N. H. Mackworth, film entitled "The Eye-Marker," Symposium on Automation at XVth International Congress of Psychology (1957).

[23] L. Carmichael, "Comprehension time, cybernetics and regressive eye movements in reading," XVth International Congress of Psychology (1957).

recently drawn attention to the importance of eye movement recordings in relation to comprehension during the reading of static printed material.

Again, however, with this eye-marker technique the display for these thinking problems can be changing all the time—as, for example, in the Britannia instrument panel shown in Fig. 7.

These eye-marker records on changing displays of real situations have mostly been obtained from motion picture films such as Fig. 8. This shows a night landing illustration which is a single shot from a movie, representing the pilot's view when he is 180 feet above ground, about 25 seconds from touch-down.

In these studies the sequence of signals unrolls according to a set program regardless of the subject's actions, but it is already possible under laboratory conditions to obtain eye-marker recordings from laboratory tasks in which the scene is altering all the time in response to previous decisions taken by the subject.

ACKNOWLEDGMENTS

Dr. F. W. Campbell of the Department of Physiology, University of Cambridge, assisted with general advice on the technique, as also did Professor R. W. Ditchburn and Dr. D. H. Fender of the Department of Physics, University of Reading.

Mr. I. Waters of Pye Ltd., Cambridge, has also advised on the necessary alterations to standard industrial television equipment. Mr. A. Davidson, Mr. C. J. Deane, and Mr. R. G. Olsson of this Unit have constructed the rest of the equipment. Mr. D. C. V. Simmonds of this Unit undertook the photographic work.

The equipment costs have been mainly provided by the Flying Personnel Research Committee of the Air Ministry. Some apparatus has been lent from the Admiralty through the Royal Naval Personnel Research Committee of the Medical Research Council. The Ministry of Supply and British Overseas Airways Corporation have also given advice and facilities.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 48, NUMBER 7 JULY, 1958

Basis for Judgments of Relative Brightness*

RICHARD M. WARREN AND ROSLYN P. WARREN
Psychology Department, Brown University, Providence, Rhode Island

(Received July 17, 1957)

The hypothesis is offered that relative brightness judgments are based upon experience with the way the amount of light reflected by objects changes with their distance from the light source.

The experimental results support this hypothesis. One group of 40 subjects judged half brightness, and another 40 estimated the change in luminance corresponding to moving a hidden point light source to twice the distance from an illuminated standard field. Judgments of the two groups were equivalent.

Under stimulus conditions designed to represent the common visual situation (stimuli subtending wide visual angles, adaptation approximating stimulus levels), one quarter the standard luminance was correctly chosen for the effect of doubling distance from the source, and the same fraction was chosen for half brightness for all standard intensities (0.00086 to 87 millilamberts).

Under less familiar conditions similar to those employed for the *bril* scale (small stimuli with black backgrounds, indeterminate levels of adaptation) half brightness judgments were again equivalent to estimates of the effect of doubling distance from object to light source. These estimates were less than one-quarter standard luminance.

The hypothesis is discussed in terms of sensory scaling in general, and the neutral *value* and *bril* scales in particular.

INTRODUCTION

NEARLY 100 years ago Fechner[1] postulated that just noticeable differences (j.n.d.'s) in stimulus intensity represented subjectively equal increments in sensation for all modalities. Although experimental evidence has limited the applicability of this postulate, there has been a reluctance to abandon it completely.

An analysis of the literature has suggested to the authors a simple hypothesis unrelated to Fechner's law, concerning the basis for judgments of the relative intensity of sensations. Predictions based upon this general hypothesis have been verified for judgments of heaviness[2] and loudness,[3] and the present experiment is designed to test its validity when applied to brightness judgments. The hypothesis considers that when some physical characteristic regularly associated with the stimulus in the everyday life of individuals changes in

* Supported in part by a grant from General Foods Corporation. The manuscript was prepared while the first author held a Senior Postdoctoral Fellowship in Physiological Psychology from the National Research Council, National Academy of Sciences, New York University College of Medicine.

[1] G. F. Fechner, *Elemente der Psychophysik* (Breitkopf und Härtel, Leipzig, 1860), Erster Theil, p. 65.

[2] R. M. Warren and R. P. Warren, Am. J. Psychol. 69, 640 (1956).

[3] Warren, Sersen, and Pores (in preparation).