PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Analogue Models of Motion Perception [and Discussion]

M. J. Morgan and H. B. Barlow

Phil. Trans. R. Soc. Lond. B 1980 **290**, 117-135
doi: 10.1098/rstb.1980.0086

Phil. Trans. R. Soc. Lond. B **290**, 117–135 (1980)
Printed in Great Britain

Analogue models of motion perception

By M. J. Morgan
Department of Psychology, University of Durham, Science Laboratories, South Road, Durham, DH1 3LE, U.K.

An object moving in discrete spatial jumps is difficult to distinguish from a continuously moving object, provided the time between jumps is not too great. The extent of this perceived continuity may be measured by probing the perceived spatial location at times between the target jumps, by either a vernier alignment or a stereoscopic technique. As the time between jumps increases the accuracy of spatial interpolation falls, until finally the object is seen only at its actual spatial locations. These results can be analysed in the frequency domain by treating the signal for apparent motion as the analogue of a periodic waveform containing relatively low frequencies (the continuous motion) and higher frequencies giving rise to the discreteness of the motion. If such an input has the higher frequencies progressively removed by physical filtering, it is perceived as increasingly continuous. The fact that such filtering is not necessary for perceived continuity when the discrete jumps occur at rates greater than about 30 Hz suggests that frequencies greater than that limit are removed by the visual system itself.

1. Introduction

Vision is characterized by spatial representations, but we do not know what it is that gives these representations their tremendous power, coupled with their deceptive phenomenal simplicity. For example, a computer can easily be programmed to play the game of 'noughts and crosses' (tic-tac-toe) but in a numerical form that makes the strategy virtually inscrutable to the ordinary person. Even a child, however, can play the game skilfully when the problem is given a spatial form in which the underlying strategy of making the symbols form straight lines becomes intuitively evident. Vision is useful not just for seeing, but for many different kinds of problem solving. Very difficult problems can be solved by giving them a spatial form of representation, and allowing the innate skills of vision to do the rest.

Although we do not know exactly what characterizes spatial representations, certain obvious remarks can be made. Perhaps the most obvious is that spatial representation allows for the possibility of movement, and thereby links our perception of the world to the motility of our bodies. We can be confident that visual representation allows for movement because in the absence of the latter it would not possess its fundamentally metrical character. A metric is not defined until congruence definitions have been stated, and as Russell and others have pointed out, the notion of congruence depends upon operations such as translation and rotation, under which shapes are defined as remaining invariant. To establish congruence relations, we need a method with which to compare the sizes of lines and angles in different places, and it is difficult to see how we could give meaning to this unless we could in some way internalize the concept of moving an object around in space.

Motion, then, is basic to the representation of space. But what, in turn, characterizes movement in the phenomenal domain? This has been a philosophical puzzle ever since Zeno's terse remark to the effect that 'An arrow is where it is'. What Zeno claimed is that the distinctive attribute of motion is lost as soon as we try to pin down the spatial position of an object at an instant of time. The invention of the calculus seems to have resolved this problem to the satisfaction of the physicist by making motion mathematically tractable, but it is not obvious that it clears it up completely from the psychologist's point of view. The momentary position of a moving object can certainly be predicted by vision with sufficient accuracy for, say, a fielder to intercept a rapidly moving cricket ball. But what does this imply for the phenomenal location of a moving object at an instant of time? We know from experiments on flicker fusion threshold that the visual system is relatively sluggish in its response to temporal change. How is this related to the phenomenal continuity of movement; and how does the visual system cope, if it does in fact cope at all, with the problem of localizing a moving target in space?

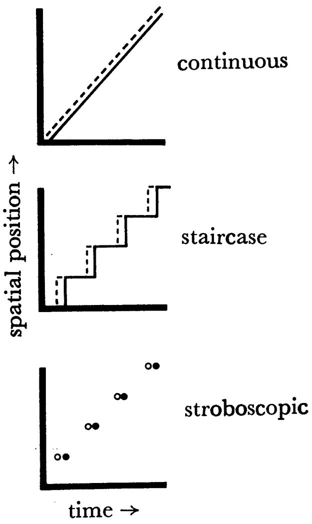


FIGURE 1. In each graph, the two lines or sets of dots represent the changes in spatial position over time of two targets in a dynamic vernier acuity task. Only the spatial positions in the direction of motion are plotted; in addition the targets are assumed to be slightly separated in the orthogonal spatial dimension as in a typical static vernier task. In the continuous case (top), both targets are in continuous motion with a constant spatial separation. In the two kinds of apparent motion (staircase and stroboscopic), the two targets are presented in exactly aligned spatial positions with a temporal displacement. Such targets appear to the observer as if they were spatially misaligned, provided that the motion is phenomenally continuous.

Several lines of evidence suggest that the phenomenal spatial location of a moving target is not determined solely by its instantaneous spatial position. In the phenomenon of 'apparent motion', for example, a target is seen as if it is changing its spatial location continuously, when in reality it is making discrete spatial jumps. One way of measuring this phenomenal continuity rather precisely is by attempting to make a vernier alignment of two moving targets. Consider the situations shown diagrammatically in figure 1. The graphs in this figure represent changes over time of the spatial location of the targets in the direction of motion. In continuous motion, the spatial location of one of the targets lags continuously behind that of the other. Not surprisingly, this is perceived by the subject as a spatial misalignment between the targets. But consider now the two kinds of apparent motion illustrated in figure 1. In 'staircase' motion the targets make discrete spatial jumps separated by periods of rest. Stroboscopic motion is similar except that the targets are flashed briefly after each jump. In both of these kinds of motion we can arrange that the targets occupy identical spatial locations, but that they occupy these positions at slightly different times, as shown in figure 1. It turns out that this temporal phase lag appears to the subject as a *spatial* offset between the stimuli, even though the spatial positions are actually aligned (Morgan 1976; Burr 1979). If the observer is asked to make a spatial adjustment to align the stimuli, he advances the spatial position of the temporally lagging stimulus in the direction of motion (figure 2).

We can refer to this descriptively as an 'interpolation' effect in apparent motion, since the observer is acting as if the phenomenal location of a discretely moving target were in between its actual spatial locations. The same effect is seen in a stereoscopic version of the alignment task. If the targets are presented to corresponding retinal points in the two eyes but with a small temporal delay, the phenomenal spatial misalignment is now detected as a depth shift as if the stimuli had a retinal disparity (Lee 1970; Morgan & Thompson 1975; Ross & Hogben 1975; Morgan 1979*a*; Burr & Ross 1979).

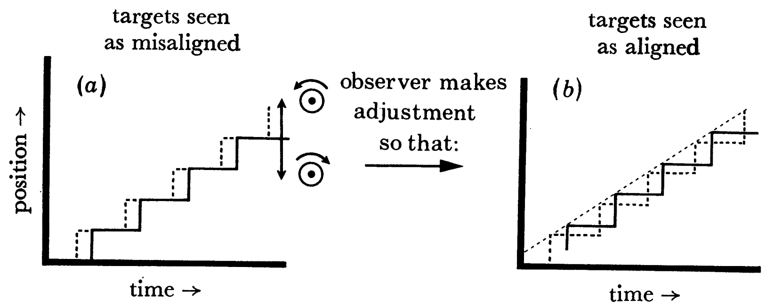


FIGURE 2. This figure explains how the observer can make a spatial adjustment so as to make temporally staggered 'staircase' targets (see figure 1) appear as if they were spatially aligned in a vernier task.

It is initially tempting to consider the interpolation effect as yet another example of the ability of the visual system to make intelligent deductions from limited sensory information. The naturally occurring analogue of the apparent movement situation would be one in which a target moves behind a picket fence (Burr & Ross 1979), or a prey is glimpsed fleetingly as it moves through a thicket. In the alignment situation, the visual system could be considered to be making a deduction in the following propositional form: 'target A arrives at the same place as target B, but at a later time; therefore, target B must be spatially ahead of target A'.

Although this idea is plausible, it is probably not the correct account of interpolation. The basic objection to the 'hypothesis formation' account is that it fails to explain the temporal constraints upon the effect. Both the vernier alignment interpolation and stroboscopic Pulfrich effect disappear when the interstimulus interval (i.s.i.) exceeds about 50 ms. At an i.s.i. of 200 ms, for example, the targets appear aligned when their actual physical positions are in alignment. This is despite the fact that such targets are still seen as moving. It would seem, then, that we must make a clear distinction between stimuli that produce an impression of movement, and stimuli that produce an impression of *continuous* movement.

An alternative to the idea that interpolation is a kind of visual 'hypothesis' is that it is a fact forced upon the visual system at a relatively peripheral level by physiological constraints. We can begin the search for these constriants by asking two theoretical questions. First, how would one set about building a physical device to assign a momentary position to a moving target? Secondly, why does a continuously moving target not appear to the visual system as a shapeless blur? I shall argue that answers to these two questions are closely related.

Any device that is attempting to represent the spatial position of a moving target will have to cope with the problem that worried Zeno, namely, that the position is constantly changing, and is thus indefinite inside any finite period of time. Only if the system could change its own state in no time at all could it respond to an instantaneous target position. Any physically realizable system will have a finite speed of response that will limit the precision with which it can assign a location to the moving target. It will be an inevitable property of any such system that is unable to distinguish between continuous and discontinuous motion when the jumps of the latter occur sufficiently rapidly in time. The critical i.s.i. at which this occurs can be used to measure the time constants of the system under investigation.

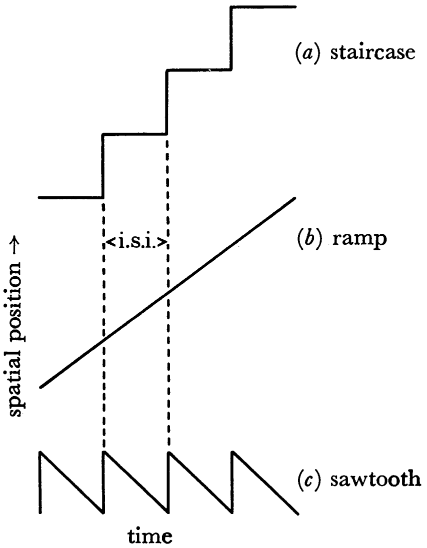


FIGURE 3. This figure shows how 'staircase' motion (*a*) can be synthesized from continuous motion (ramp) (*b*) and a higher frequency modulation of spatial position (sawtooth) (*c*). For further explanation see text.

Before describing experiments based upon this idea, it is necessary to show in slightly greater detail how a discretely moving target can be considered to contain within it different spatio-temporal frequency components. Consider the staircase representation of apparent motion shown in figure 3*a*. As well as representing the trajectory of an apparently moving target, this staircase could be taken to represent a signal used to produce a discretely moving target, such as a voltage applied to a spot on an oscilloscope screen. If we wished to synthesize such a signal, we could do so by a combination of the signals shown in figure 3*b* and *c*. The first of these is a ramp, which by itself would correspond to a target in continuous movement. The second is a high frequency modulation of the target in a sawtooth pattern.

We can therefore characterize a staircase signal in the spatio-temporal frequency domain as a mixture of different components, with lower frequencies corresponding exactly to a continuously moving target, and higher frequencies being responsible for the discontinuous features of the motion. This is further illustrated in figure 4 which shows the superimposed power spectra of three different staircases, of differing i.s.is. Each of these staircases has the same low frequency components, corresponding to a continuous triangular wave. The higher frequency components differ between the staircases. For example, a staircase of i.s.i. 128 ms has a fundamental at

7.8 Hz and harmonics at 15.6 Hz, 23 Hz, and so on. The 32 ms staircase, on the other hand, has no frequency component below 31 Hz to distinguish it from continuous motion.

Given this way of describing discontinuously moving targets, the suggestion can now be made that the phenomenal continuity of a staircase may be due to the inability of the visual system to represent its higher frequency components. For example, it will be shown below that a staircase of i.s.i. 32 ms has a very high degree of phenomenal continuity. This can be explained by supposing that the frequency components distinguishing it from a continuously moving

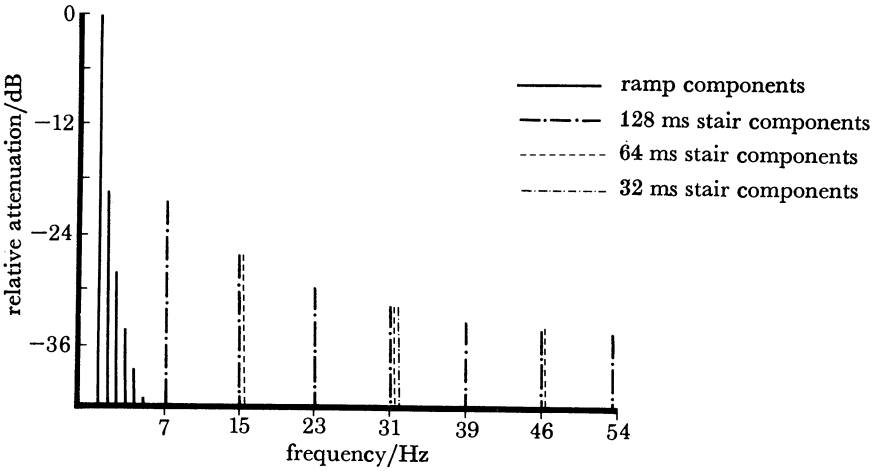


FIGURE 4. The figure shows how 'staircase' motion (see figure 3) can be broken down into different frequency components. The power spectra of three different staircases with differing i.s.is (see figure 3 for explanation) have been superimposed to illustrate their different frequency composition. The three staircases have the same ramp components, which define a low frequency triangular wave, but differ in the higher frequency components making up the added sawtooth modulation (see figure 3). The data were obtained from a fast Fourier transform of actual staircase stimuli.

target are effectively not represented at the level of the visual system where the analysis of visual direction takes place. According to this argument, the 'interpolation effect' ceases to be a paradox and becomes a simple consequence of the relatively sluggish response of the visual system. The seeming paradox was that temporally staggered targets could be seen as spatially out-of-phase even though their physical positions were the same. However, we now see that it is in an important sense misleading to say that their 'physical positions are in alignment'. This statement refers to the *instantaneous* physical positions of the targets: for example, the discrete space–time coordinates of a stroboscopic target. But if the visual system cannot respond quickly enough to analyse these discrete positions independently, it will necessarily be averaging the target position over finite periods of time. In these circumstances, a temporal delay between the targets would have an identical result to a spatial offset.

2. Frequency components above 25 Hz have little effect upon phenomenal continuity

In this experiment the observers carried out a vernier alignment between two moving stimuli. The experimental arrangement is illustrated in figure 5 for the case where the two bars were in staircase motion with an i.s.i. of 128 ms. The display was generated on an oscilloscope screen with fast decay P15 phosphor. There was a temporal delay in the plotting of the two bars such that one was lagging by a constant fraction (0.25) of the i.s.i. The observer's task was to adjust the spatial position of the lagging stimulus so that it appeared to be in alignment with the leading stimulus (figure 2), and from the chosen setting a motion continuity index was calculated as explained in figure 6. A central fixation point was provided and a closed-circuit television picture of the eye was examined to check that no tracking took place. Further details of the

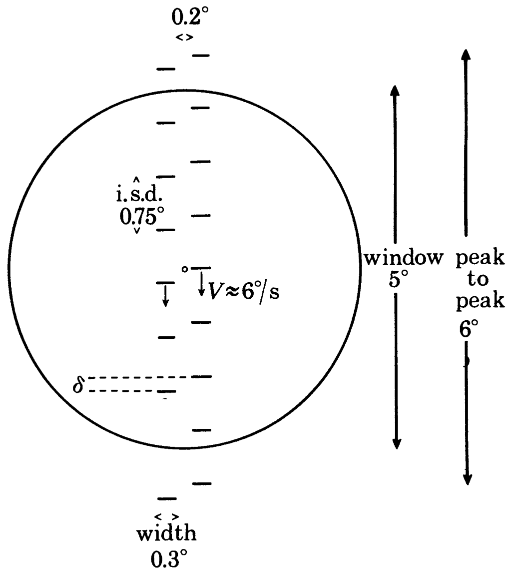
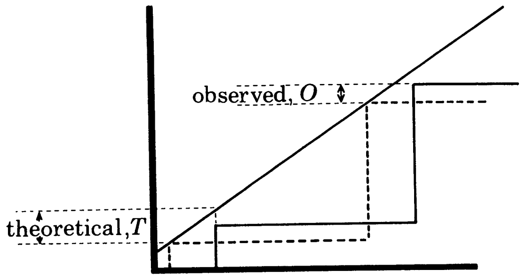


Figure 5. Schematic representation of the display used for experiments on vernier alignment of stimuli in staircase motion. The two bars were in vertical motion, and the observer had to adjust the spatial separation (δ) so that they appeared to be in alignment. In this illustration the i.s.i. was 128 ms, there were therefore eight steps from peak to peak with an interstep distance of 0.75°. With an i.s.i. of 64 ms the number of steps would be doubled and the inter-step distance halved, and so on for other i.s.is.



motion continuity index, M:
$M = 100\, O/(O+T)$
$M = 0$ if no interpolation
$M = 50$ if complete interpolation

Figure 6. This figure explains the calculation of the motion continuity index, which measures the interpolation effect in vernier or stereoscopic alignment of apparently moving targets. The observer sets the spatial position ordinate of the temporally lagging target (lag $= 0.25 \times$ i.s.i.) to make it appear in alignment with the leading target. The theoretical value for complete interpolation assumes that the presentations of the lagging target will be placed on the continuous trajectory joining the discrete space–time positions of the leading target.

experimental procedure are available (Morgan 1979*b*). The temporal frequency composition of the stimuli was manipulated in two ways. First, three different staircases were used, with i.s.is of 128, 64 and 32 ms respectively. As shown in figure 4, these staircases have different frequency components distinguishing them from a continuously moving target. Secondly, before being used to move the targets on the display, the various staircase signals were passed through an analogue filter to remove frequency components above a specified frequency. In other words, they were subjected to low-pass filtering. For example, with the low-pass corner frequency set at 12 Hz, a staircase of i.s.i. 128 ms would be physically distinguished from a continuously moving target only by the presence of an added sinusoidal modulation of frequency 7.8 Hz.

The mean data collected from four observers are illustrated in figure 7. There are three main points to note. First, there is much more of an interpolation effect with the 32 ms staircase than with one of i.s.i. 128 ms. The 64 ms staircase is intermediate. The 32 ms staircase produced

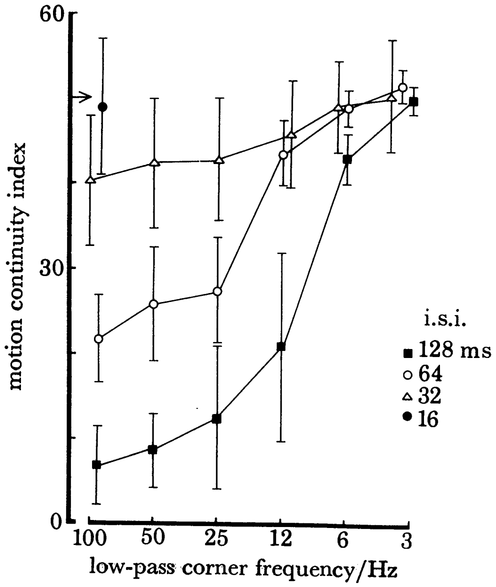


FIGURE 7. Results from an experiment in which observers carried out vernier alignment between staircase-moving targets as illustrated in figure 5. The staircase signals were subjected to various degrees of low-pass analogue filtering (abscissa) before being used to move the targets on the c.r.o. screen. For explanation of the motion continuity index (ordinate) see figures 2 and 6. The points are means over four observers and the vertical bars represent standard deviations. For further explanation see the text.

almost complete interpolation even when it was virtually unfiltered (low-pass frequency 100 Hz), whereas in the same circumstances the 128 ms staircase produced no interpolation effect. The second obvious point is that as the staircases had more of their higher frequency components removed by filtering, the motion continuity index rose. Even the 128 ms staircase produced a large interpolation effect when filtered to remove frequencies higher than 6 Hz. Finally, the results suggest that there is relatively little effect of filtering until frequencies below 25 Hz are removed.

These results can all be put together if we assume that frequencies in the signal higher than about 25 Hz are not represented as spatial movements at the level of the visual system where the visual directions of the targets are compared. The staircase of i.s.i. 32 ms contains no component below 31 Hz to distinguish it from continuous motion, and will thus have a high degree of motion continuity. A 128 ms staircase filtered at 12 Hz will still have a component at 7.8 Hz to distinguish it from a continuous target, and thus will not show a high degree of motion continuity. There is a steep rise in the continuity of the 64 ms staircase when the low pass frequency changes from 25 Hz to 12 Hz because this staircase has a fundamental at 15 Hz (figure 4).

The results therefore suggest that modulations of the target position at frequencies greater than about 25 Hz have little influence upon its phenomenal location, as measured by the alignment technique. A prediction from this is that the phase of such modulations relative to the continuous low frequency components will be of little importance. This was verified in the next experiment.

3. Mirror image equivalence between high frequency components

As explained diagrammatically in figure 3, a staircase signal can by synthesized by addition of a ramp and a sawtooth in which the local velocity is equal and opposite to the ramp. If the sawtooth is mirror-imaged and then added to the ramp, we have the signal illustrated in figure 8*a*. The local velocity of this stimulus is now twice that of the ramp, and the instantaneous jumps are in the reverse direction of the ramp. Despite the physical differences between this signal and the staircase, we can predict that they will have equivalent perceptual effects if the fundamental frequency of the sawtooth exceeds about 25 Hz. This was verified in the same vernier alignment situation that was used in the previous experiment. The two stimuli to be aligned were presented either as staircases with an i.s.i. of 32 ms, or as the mirror-imaged versions shown in figure 8. The results for two observers conformed to expectations in that both kinds of stimulus produced the same degree of interpolation, as measured by the motion continuity index. Phenomenally, the two kinds of stimulus were difficult to distinguish. In particular, the 'reverse jumps' of the mirror-imaged stimulus were not perceptible.

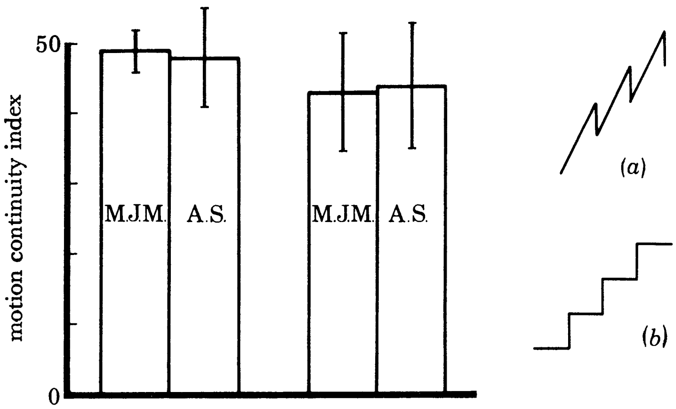


Figure 8. The figure shows the results of an experiment comparing the motion continuity of two kinds of moving stimuli. One of these (*b*) was a staircase, synthesized from a ramp and a sawtooth modulation as in figure 3. The second (*a*) was identical except that the sawtooth was mirror-imaged. Motion continuity was assessed by the interpolation effect in vernier alignment with a delay of 0.25 × i.s.i. Other details as in figure 5. Results are shown separately for two observers (M.J.M. and A.S.). For further explanation see the text (§3).

4. Effects of light adaptation state upon interpolation

It has been argued above that the interpolation effect depends upon the relatively sluggish nature of the visual response, which forces an averaging of visual direction over time, and thus a confusion between a temporal and a spatial offset. If this is so, it might be possible to manipulate the extent of the interpolation by controlling the speed of the visual response. A possible means that suggests itself is the control of light adaptation level. It is a well established finding that as the eye becomes more dark adapted, its time constants are increased. Thus, visual persistence increases with dark adaptation (Mollon 1969) and the flicker fusion threshold is reduced at lower frequencies (Kelly 1961). It is therefore possible that in the apparent movement situation, an effect of adaptation state upon interpolation might be found. This result has been reported in the stroboscopic stereophenomenon (Morgan 1979*a*). This experiment first of all verified the influence of the i.s.i. upon the interpolation effect. It was found in this case that interpolation was complete with an i.s.i. of 20 ms, but absent with an i.s.i. of 50 ms. Results with an i.s.i. of 30 ms showed an intermediate degree of interpolation. The luminance of the display and the surround was then decreased by 2 logarithmic units. This had no effect upon the extent of interpolation with the two extreme i.s.is (20 and 50 ms) but interpolation with the intermediate i.s.i. of 30 ms became complete. The prediction was therefore verified.

5. Demonstrations of spatial averaging with moving targets

The experiments reported so far suggest that there is some form of filtering involved in the interpolation phenomenon, but they have had little to say about the mechanism of such filtering. One hypothesis is as follows. We know that even very brief visual inputs can give rise to a considerably more protracted visual response; this is the phenomenon of visual persistence. It follows that a target sweeping along a retinal array is going to provide information not only about its current position, but about previous positions as well, depending upon target velocity and the length of persistence. Thus at any one time there will not be a single retinal image corresponding to the target, but rather a range of different images. If this is so, how could two such targets be compared, either in the vernier or in the stereo situation? One possibility would be to compare just the leading edges, that is, the most recently available information about direction. However, this would not explain the interpolation effect, since it postulates a comparison between targets based upon their momentary actual physical position. An alternative is that the distribution of visual directions from persisting signals is sharpened in some way, perhaps by an inhibitory process, to assign a mean position to the target. This averaging process would have the required property of acting as a low-pass filter for spatio-temporal modulation, and of explaining the interpolation effect. Some further experiments will now be described supporting the view that there is such an averaging process.

Suppose a horizontally moving target is viewed binocularly. If the signal could be made to persist for longer in one eye than the other, the account just given would predict that there would be a depth shift, as if the more persistent signal were spatially lagging. Greater persistence in one eye could be arranged by covering that eye with a neutral density filter, it being well established that *decreases* in intensity of the target cause *increases* in persistence (Mollon 1969). The depth shift with one eye covered by a filter is just the Pulfrich effect (Pulfrich 1922), which is hardly a novel prediction. But we now see that the traditional 'transmission lag' interpretation of the Pulfrich stereophenomenon has an alternative, which stresses the effects of the light reduction upon persistence, rather than latency, of the visual response. One way to test the persistence hypothesis is to cause physically greater persistence of the signal in one eye, and to pit this against the effects of a filter over the other eye. In one experiment (Morgan 1975) a stroboscopically moving target was presented separately to the two eyes. Every time the target flashed in a certain place in one eye it was flashed not only in the exactly corresponding place in the other eye, but was also presented at its previous position. Thus, at every time the target was presented at the *n*th position in one eye, it was flashed simultaneously at the *n*th and the

$(n-1)$th position in the other eye. As predicted this produced a depth shift, as if the signal in the 'more persistent' eye were lagging. Moreover, this effect could be cancelled out by placing a suitable neutral density filter over the 'less persistent' eye. The second experiment (Morgan 1977) was a continuous version of the previous one. One eye viewed a horizontally moving vertical bar target, and the other eye viewed an identically moving but slightly wider bar. When the leading edges of the two bars were in physical alignment, the observer saw a depth effect as if the wider bar were spatially lagging. This depth could be cancelled by placing a filter over the eye seeing the thinner bar. When the thinner bar was physically aligned with the exact geometrical centre of the wider bar, there was no depth effect. This result suggests that the averaging process was remarkably linear.

6. Similarities between staircase and stroboscopic motion

One advantage of relating the interpolation effect to visual persistence is that we are now able to explain why the effect applies to both staircase and to stroboscopic motion. In staircase motion the stimulus remains physically present on the retina in the intervals between its jumps, whereas in stroboscopic motion it is flashed briefly at each position. However, if each presentation of the target gives rise to a persisting impression, the physical difference between the stimuli may not entail an important phenomenal difference. To check on this point, an experiment was carried out in which the duration of the target at each of its positions was systematically varied, and apparent alignment was measured as in the first experiment. At one extreme condition, only a single 2 ms flash of the target was presented, corresponding to stroboscopic motion. At

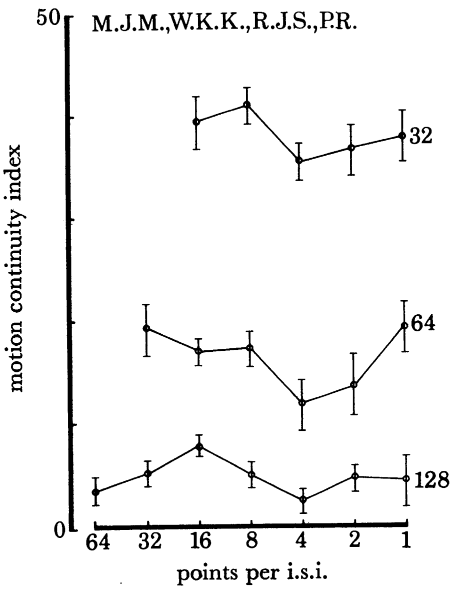


Figure 9. Results of an experiment investigating effects upon motion continuity of the length of time for which the target was physically present during the i.s.i. The three sets of points show results for different i.s.is (values in milliseconds). Target duration was manipulated by the number of times it was plotted during the i.s.i. (abscissa), each plot taking 2 ms. The points plotted during the i.s.i. were in identical spatial locations. Motion continuity was assessed by the interpolation effect in vernier alignment with a delay of $0.25 \times$ i.s.i. The points are means over four observers and the vertical lines represent standard deviations. For further explanation see the text (§6).

the other extreme, the target was flashed every 2 ms until the i.s.i. was complete. Between these two extremes, an intermediate number of equally spaced 2 ms flashes were plotted. The experiment was carried out with three different i.s.is and the results are plotted in figure 9. First of all, the experiment confirms the effect of i.s.i. found in previous experiments. Secondly, it is apparent that there is little systematic effect of the target duration (number of points per i.s.i.) upon the motion continuity index. In particular, there is little difference between the limiting stroboscopic case of a single flash and the staircase situation in which the target duration fills the i.s.i. There are some interesting aspects of the data that require further investigation. With all three i.s.is there is a small but systematic tendency for the continuity index to reach a minimum when 4 points were plotted in each i.s.i., that is, when the cumulative target duration was 8 ms. There is no obvious explanation of this phenomenon.

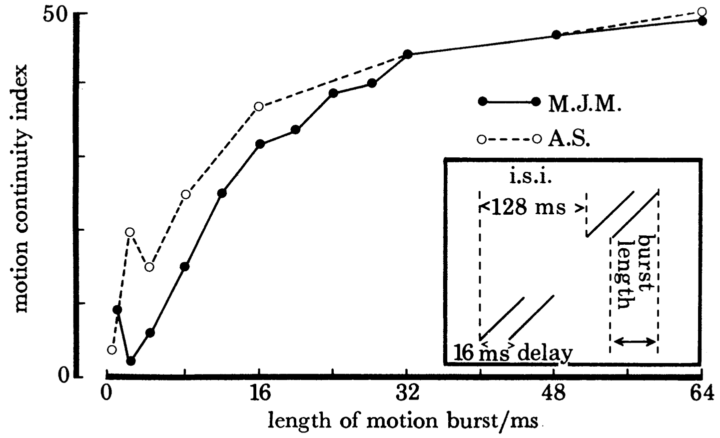


FIGURE 10. Results of an experiment investigating effects upon motion continiuty of the length of time for which the targets were continuously moved during the i.s.i. The observer performed a vernier alignment between two targets with the spatial and temporal relations shown in the inset panel. Other details as in figure 5. For further explanation see the text (§7).

7. A HYBRID BETWEEN CONTINUOUS AND STROBOSCOPIC MOTION

We have seen that with an i.s.i. of 128 ms, targets are aligned by their physical positions, irrespective of the temporal delay between them. The interpretation of this finding in terms of the spatial averaging hypothesis is that the temporal period over which integration occurs must be less than 128 ms. Inside any shorter period, the mean spatial position of the temporally staggered staircases will generally be zero. By the same argument, the integration period must be greater than 32 ms, or we should not get such a high interpolation figure when the staircases have an i.s.i. of 32 ms. Suppose we now manipulate the average spatial position of the targets within the i.s.i. by subjecting them to continuous movement. This gives rise to a hybrid between continuous and stroboscopic motion as illustrated in figure 10. The two temporally staggered targets are presented for the duration of a 'motion burst' and then disappear until the next burst of motion. We now ask how long the burst must be for the observer to cancel out the temporal delay by an appropriate spatial offset. The experiment was carried out with two observers, using a vernier alignment task as in experiment 1. The results (figure 10) show that a motion burst of 64 ms was sufficient to produce completely accurate alignment. We can therefore conclude that assignment of target position takes place within an integration period no greater than 64 ms, and probably not much greater than 32 ms. This agrees well with the finding of the first experiment, that spatial modulations of frequency greater than 25 Hz are phenomenally attenuated.

It is also interesting to compare the present result with Ross & Hogben's (1974) finding of a 50 ms 'short-term' memory in stereopsis, which suggested that inputs to the two eyes are integrated provided they do not occur more than 50 ms apart.

8. Threshold for the interpolation effect

Supposing the visual direction of a moving target to be determined by some averaging process, how precise are the judgements involved? If a moving target simply becomes a shapeless blur, we should not expect judgements of vernier alignment and stereoscopic depth to be carried out with any precision. However, work by Westheimer & McKee (1975, 1978) has shown that vernier acuity and stereoscopic depth judgements for continuously moving targets remain accurate, even when the presentation time is too short to permit pursuit eye movements. In the following experiment, Walia Kani and I compared vernier and stereoscopic acuity in a situation where the targets were in staircase motion with an i.s.i. of 20 ms.

The two targets were presented on separate oscilloscope screens (Hewlett-Packard 1333A with P15 phosphor). In the stereo task the two screens were fused with a mirror stereoscope, and in the vernier task they were optically superimposed with a half-silvered mirror. The targets were bars whose width subtended 0.25° of visual angle. In the vernier alignment task the separation between the ends of the bars was 1′. The velocity of the ramp component in the staircase was 4.8°/s. The temporal delay in the plotting of the two staircases was varied over trials in a randomized psychophysical procedure to determine the minimum temporal separation that could be detected by the observer. In the stereoscopic experiment, either the right or the left eye was delayed, and the observer had to report the direction of target motion in depth ('clockwise' or 'anticlockwise'). In the vernier alignment task, either the top or the bottom target was delayed, and the observer reported which target appeared to be spatially lagging ('top' or 'bottom'). The display was generated at a rate of 1 frame every 290 μs.

Results are shown in figure 11. In presenting these data, the temporal delays have been converted into a notional spatial equivalent, by calculating the distance the target would have moved during the delay if had been in actual continuous motion. For example, at a velocity of 4.8°/s a temporal delay of 1 ms corresponds to a notional spatial offset of 17″. It should be remembered, however, that the targets were actually in staircase rather than continuous motion, so that the calculated spatial offset refers to the *mean* separation between the targets rather than their instantaneous offset, which is at most times zero.

If we take the threshold as the point of 75% detection, the results for two observers in figure 12 show that stereoacuity for the staircase targets was in the region 10–20″. This corresponds to a temporal delay of approximately 1 ms. Thresholds for the vernier alignment situation were twice as great as this for one subject (M.J.M.) and one and a half times as great for the other (W.K.K.). To see if vernier acuity would be improved by allowing the observers to track the target, the experiment was repeated with instructions to follow the target with the eyes rather than fixating on a stationary mark as in the previous condition. Results (figure 11) showed an improvement in both observers to an acuity of about 10″.

Under ideal conditions, the threshold for stereoacuity and vernier acuity is as little as 2″. With stationary targets and a 1 min gap, McKee & Westheimer (1978) have recently reported improvements with practice from 10 to 5″. With targets moving up to 3.5°/s, Westheimer & McKee (1975) found vernier thresholds of between 6 and 10″. It is apparent that the acuity in the present situation is quite a bit worse than these values, although it is still impressive when it is considered that the stimuli were in discontinuous motion, and that the temporal separation was only 1 or 2 ms. It is not clear whether the difference in acuity between our study and that of Westheimer & McKee (1975) was due to the difference in the kind of motion, or to the higher velocity used in our experiment. It is of interest that the motion used by Westheimer & McKee was in fact a staircase with an i.s.i. of as long as 10 ms, but the offset between the stimuli was spatial rather than temporal.

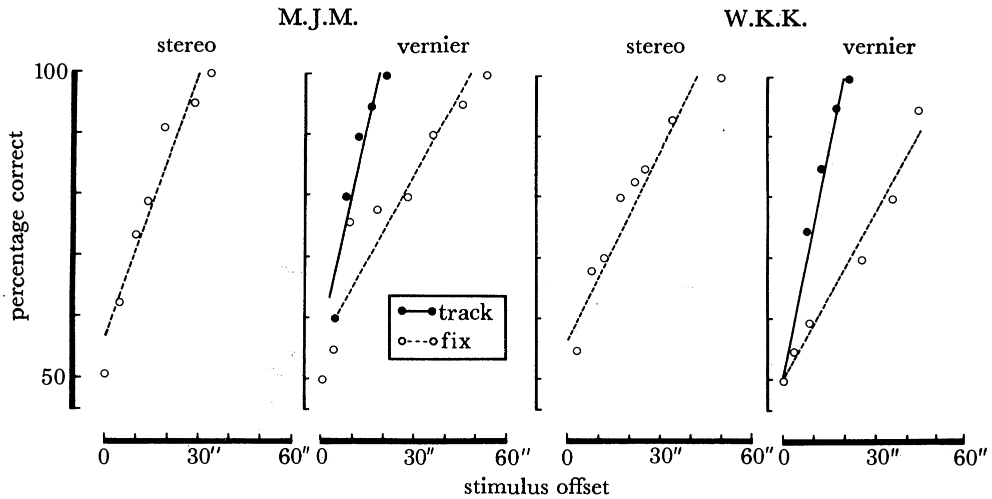


FIGURE 11. Results of an experiment on vernier and stereoscopic acuity with targets in staircase motion (i.s.i. = 20 ms). The number of correct identifications of depth (stereo) or left–right alignment (vernier) is shown as a function of the virtual spatial separation between the staircases. A virtual separation of 17″ corresponds to a delay between the staircases of 1 ms. Results are shown separately for two observers (M.J.M., emmetropic; W.K.K., myopic corrected to 6/3). For further explanation see the text (§8).

9. The role of eye movements in interpolation

The finding in the previous experiment, that vernier acuity was improved by tracking the target, raises the question of whether the interpolation effect might depend entirely upon eye movements. If the eyes track the apparently moving targets, any temporal delay between them will be automatically converted into a spatial offset on the retina, and there would be no need to invoke an interpolation process. A very similar question about the role of eye movements has been extensively discussed in relation to a phenomenon first described by Zöllner (1862). A shape such as a triangle is moved behind a very narrow slit, so that only a very small fraction of the shape is visible at any time. If the observer is instructed to maintain careful fixation on the slit, all the parts of the shape ought to fall on the same part of the retina, yet despite this, a spatially extended image of the shape can sometimes be seen. The shape is frequently compressed in the direction of its motion. There is an intriguing analogy here with the interpolation effect in apparent motion, in that spatial perception is apparently resulting from purely temporal information. But as Helmholtz pointed out (see Southall 1962), Zöllner's effect is very easily explained if the observer makes tracking eye movements so as to 'paint' the temporally successive parts of the shape on to different retinal points. Despite much argument in the literature (reviewed by Anstis & Atkinson 1967), it seems that Helmholtz was probably right about the role of eye movements. For example, if the display is switched off when the subject starts to track, or if tracking is prevented by the barbiturate drug sodium amytal, then the subject's ability to identify the moving shapes is considerably impaired (McManus & Morgan 1978). The results of this experiment are illustrated in figure 12.

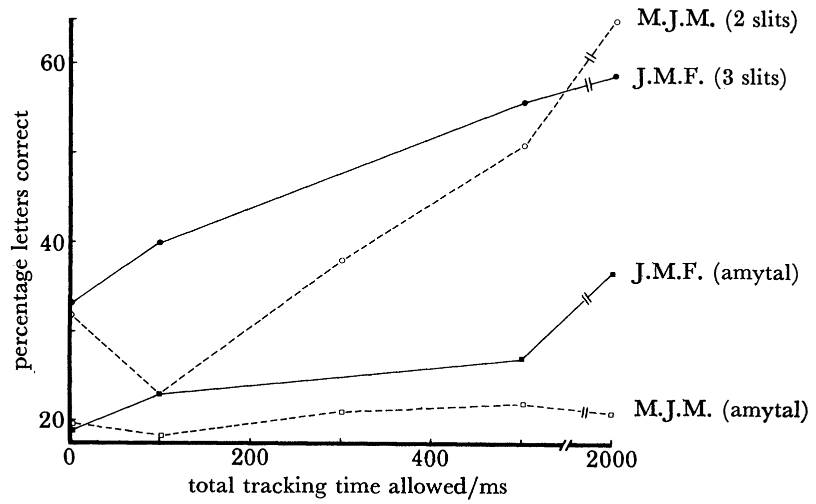


FIGURE 12. Results of an experiment in which the observer attempted to identify alphanumeric characters as they were moved rapidly behind narrow slits. Eye movements were recorded by a computer, and the stimulus could be switched off after a controlled duration of smooth tracking had occurred. In the zero tracking condition (abscissa) the observer was instructed to fixate and the display would be switched off immediately if tracking was detected by the computer. In the amytal condition the two observers (M.J.M. and J.M.F.) took an oral dose of 240 mg sodium amytal before the experiment, to block eye tracking. For further details see text (§9).

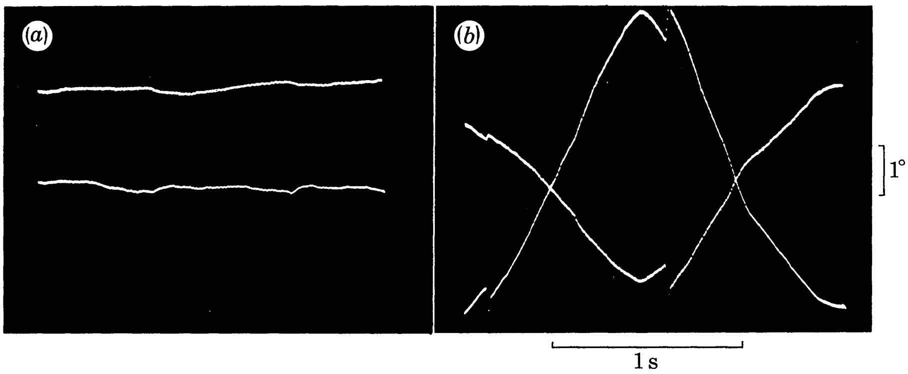


FIGURE 13. Binocular eye movement recordings from a subject (M.J.M.) engaged in a stereoacuity task (see figure 11). One of each pair of records represents movements of the left eye and the other the right eye. Record (*a*) was taken when the subject was instructed to fixate; record (*b*) when he tracked the moving targets. Records from the left eye are phase-inverted relative to the right because for both eyes records were taken from the temporal margin of the limbus.

However, it seems unlikely that tracking eye movements are responsible for the interpolation effect in vernier and stereoscopic alignment. In an elegant experiment, Burr (1979) has shown that the interpolation effect is still observed when the display is presented within the latency period for instituting smooth pursuit movements. To investigate this matter further, binocular eye movement records were taken from a subject engaged in the acuity task previously described. The recorder consisted of a fibre optic Y guide for locating the limbus in each eye (Findlay 1974). The limit of resolution of this device was 2′, and it was sufficiently sensitive to detect microsaccades. A typical example of a record when the subject was attempting to fixate is shown in figure 13*a*, and it reveals no systematic pursuit movements. In contrast, figure 13*b* shows a record when the subject was instructed to track the moving target, which was a staircase with an i.s.i. of 20 ms and an interocular delay of 4 ms.

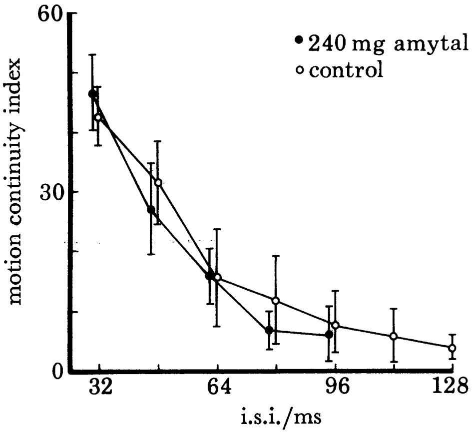


FIGURE 14. The figure shows results of an experiment on effects of sodium amytal (240 mg) on the vernier interpolation effect in one subject (J. M. F.). The severely disruptive effects of the drug upon smooth pursuit movements are shown in figure 15. Despite this disruption, the drug has no effect on the motion continuity index at any of the i.s.is.

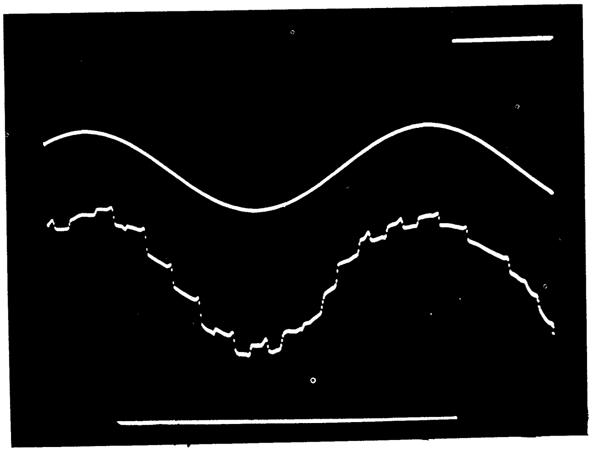


FIGURE 15. The record shows an example of the way in which sodium amytal (240 mg) disrupts smooth pursuit movements. The top record shows movement of the target (1 Hz), the bottom movement of the eye. Note the continual interruption of pursuit by saccadic eye movements. (Subject J. M. F.)

In another experiment on this problem, a subject carried out the vernier task before and after a dose of sodium amytal sufficient (240 mg) to disrupt smooth tracking. The results showed no effect of the drug upon the motion continuity index over a large range of i.s.is (figure 14). Eye movement records confirmed that the ability to track moving targets was severely impaired (figure 15) (Rashbass 1961).

These experiments do not prove that interpolation would occur in the absence of small involuntary drifts and tremors of the eye, but they have shown that large-scale pursuit movements are not necessary. The problem of the role of very small eye movements is one shared with studies of 'static' acuity (Westheimer & McKee 1975).

10. The dynamic visual noise stereo effect

If dynamic visual noise (d.v.n.), which resembles the electronic snowstorm on a detuned television, is viewed with a filter over one eye, an effect similar to the Pulfrich stereophenomenon is seen: the noise takes on a coherent rather than a random movement, and rotates in depth. (The motion is clockwise with a filter over the left eye and anticlockwise with a filter over the right eye). Tyler (1974; 1977) has put forward a 'random spatial disparity hypothesis', according to which the filter introduces a delay so as to cause fusion of random dots in different frames, thereby giving rise to a retinal disparity. Only horizontally separated dots in different frames will give rise to depth shifts, and these give rise to a monocular movement cue, thus explaining the correlation between movement and depth. A problem with this explanation is that the effect can be seen even when the filter introduces a far smaller delay than the television inter-frame interval. We therefore have the same problem as in explaining the stroboscopic Pulfrich effect (Lee 1970; Morgan & Thompson 1975). Tyler (1977) suggests an explanation based upon visual persistence and averaging. His account is similar to the one suggested here, except that the averaging involves disparity detectors rather than spatial position and movement.

The d.v.n.-Pulfrich effect is often considered to be the same as an effect described earlier by Ross (1974) in which a square of d.v.n. stood out in depth from a d.v.n. background if the latter were presented with an inter-ocular delay. However, this effect was only seen if the delay was greater than 70 ms, so it is rather hard to relate it to effects produced by neutral density filters. According to the findings of Ross & Hogben (1974), stereo fusion breaks down with delays greater than 50 ms, so perhaps the Ross (1974) effect was due to a depth difference between monocularly and binocularly seen parts of the noise field (Wist, 1969). The observation that d.v.n. can sustain smooth tracking eye movements (Ward & Morgan 1978) may also be relevant to depth effects in d.v.n.

11. Discussion

There has been considerable discussion recently in the cognitive psychology literature of a role for an 'analogue' process in visual perception (Shepard 1978). The idea of a more holistic or analogical representation of visual events contrasts strongly with prevailing neurophysiological and computational approaches, both of which favour more discrete, symbolic or propositional representations. One of the key notions in the analogue approach is that when a subject attempts to imagine a visual transformation such as a rotation in space, he constructs a continuous, or almost continuous, series of intermediate spatial representations. Thus the time taken to imagine the transformation will rise linearly with the spatial separation between the initial and final figures. (Shepard (1978) may be consulted for a review of the experimental evidence.)

Robins & Shepard (1977) have specifically attempted to implicate a form of analogue processing in the perception of apparent motion. They claim that in the apparent movement situation, the subject constructs intermediate spatial representations corresponding to a target in continuous or near-continuous motion. They attempted to verify this claim by asking subjects to decide whether a dot flashed at a point on the phenomenal trajectory of the apparently moving target occurred before or after the target reached that position in space. The further along the phenomenal trajectory the point was flashed, the longer it had to be delayed relative to the start of the i.s.i. to be judged 'after'. Robins & Shepard concluded that their results are 'in agreement with the idea that subjects based their judgements in the experimental condition on the comparison between the probe and an internal representation of a bar rotating back and forth'.

At first sight, this finding is compatible with the 'interpolation' effect in apparent motion. However, further examination shows that the relation is superficial. Robins & Shepard used a display consisting of two bars only, with an onset–onset interval of 380 ms and an offset–onset interval of 180 ms. Since these figures greatly exceed even the longest i.s.i. used in the 'interpolation' experiments it is very unlikely indeed that true interpolation occurred in the Robins & Shepard study. Their findings can perhaps be explained in a completely different way. Robins & Shepard were apparently unaware of earlier studies of interpolation in vernier alignment and stereopsis, so they took no account of the importance of eye movements. Their subjects were not instructed to fixate, so they may have tracked the movement, probably by a saccade rather than smooth pursuit (Morgan & Turnbull 1978). This means that the 'before–after' judgement could have been made entirely by retinal location of the probe flash, the subjects responding 'before' if the flash was to one side of the fovea and 'after' if it was to the other side. Since relative visual direction of flashes presented before and during a saccade depend primarily on their relative retinal locations, with little compensation for eye movement (Matin & Matin 1969), this interpretation would explain the data.

The idea that apparent motion involves the active construction of intermediate representation is somewhat implausible given the very small time intervals involved. It is relevant here to point out that the times taken for an imaginary 'mental rotation' are at least an order of magnitude greater than the critical i.s.is for apparent movement (Shepard & Judd 1976). It begs the question to put this difference down to 'the generally greater speed of imagery when it is driven externally than when it has to be programmed and generated internally' (Shepard & Podgorny 1978, p. 224), since this is a description of the phenomenon rather than an explanation of it. Another problem with the 'intermediate representation' idea in apparent motion is that it appears to ignore the existence of phenomenal persistence. If intermediates are constructed, how are they combined with the persisting signals from earlier physical presentations? The argument that I have attempted to put forward is that the 'intermediates' *are* in fact the set of persisting images, with a phenomenal location depending upon a spatial averaging process. If this is accepted, there is no need for an active process to construct intermediates.

There are at least two senses that one could give to the idea of an analogue process in perception. The first is that in respect to the 'primary qualities' of space, time and motion we generally ascribe the same set of properties both to physical events and to perception. Thus it is meaningful to speak of the duration of an experience, and of the physical event that caused that experience. Although the durations may differ numerically, they are both durations. This contrasts sharply with 'secondary qualities' like colours and smells, where we draw a rigid distinction between the *conceptual* properties of the physical event (e.g. wavelength) and the *perceptual* properties of the phenomenal event (e.g. colour).

There is thus a certain sense in which we treat spatial and temporal aspects of our experience as being 'analogues' of the physical world, rather than being merely arbitrary representations. This often seems to have been what is meant by calling perception a 'model' of the external world. For example, Hertz spoke of perception as giving us 'the actual dynamical relations' between things, and this would hardly make sense unless we could apply the concepts of dynamics both to the phenomenal and physical domain.

Spatial and temporal representations (and therefore movement) are thus analogical representations simply in the Kantian sense that, with respect to them, our concepts and our sensory intuitions are in some sort of agreement. This point is not particularly controversial.

The other meaning of 'analogue' processing is much more controversial. It implies that the neural code for an event must bear a resemblance, either to the event itself, or to the phenomenal event, or to both. Again, if this merely implies that neural events occur in time and space, it is not controversial, but usually more than this is implied; the idea is that the brain contains something like moving pictures of the outside world. A perception of continuous movement, for example, would involve something moving continuously in the brain, as in the Gestalt isomorphist theory. Shepard makes it quite clear that this is not what he has in mind when he invokes 'analogical' representations. The claim is rather of the first kind, namely, that perceived motion resembles motion (for example, in being continuous). There is little reason to dispute this very weak version of the claim that the perception of motion is an analogue process.

References (Morgan)

Anstis, S. M. & Atkinson, J. 1967 Distortions in moving figures viewed through a stationary slit. *Am. J. Psychol.* **80**, 572–585.

Burr, D. C. 1979 Acuity for apparent vernier offset. *Vision Res.* **19**, 835–838.

Burr, D. C. & Ross, J. 1979 How binocular delay gives information about depth. *Vision Res.* **19**, 523–532.

Findlay, J. M. 1974 A simple apparatus for recording microsaccades during visual fixation. *Q. Jl. exp. Psychol.* **26**, 167–170.

Kelly, D. H. 1961 Visual responses to time-dependent stimuli. 1. amplitude sensitivity measurements. *J. opt. Soc. Am.* **51**, 422–429.

Lee, D. N. 1970 A stroboscopic stereophenomenon. *Vision Res.* **10**, 587–593.

Matin, L. & Matin, E. 1969 Visual perception of direction when voluntary saccades occur. 1. Relation of visual direction of a fixation target extinguished before a saccade to a flash presented during a saccade. *Percept. Psychophys.* **5**, 65–80.

McKee, S. P. & Westheimer, G. 1978 Improvement of vernier acuity with practice. *Percept. Psychophys.* **24**, 258–262.

McManus, I. C. and Morgan, M. J. 1978 The effect of eye movements on the perception of form and depth of moving targets. Paper presented to meeting of the Experimental Psychology Society, July 1978.

Mollon, J. D. 1969 Temporal factors in perception. D.Phil. thesis, University of Oxford.

Morgan, M. J. 1975 Stereoillusion based on visual persistence. *Nature, Lond.* **256**, 639–40.

Morgan, M. J. 1976 Pulfrich effect and the filling in of apparent motion. *Perception* **5**, 187–195.

Morgan, M. J. 1977 Differential visual persistence between the two eyes: A model of the Fertsch–Pulfrich effect. *J. exp. Psychol.: hum. Percept. and Perform.* **3**, 484–495.

Morgan, M. J. 1979*a* Perception of continuity in stroboscopic motion: a temporal frequency analysis. *Vision Res.* **19**, 491–500.

Morgan, M. J. 1979*b* Spatio-temporal filtering and the interpolation effect in apparent motion. *Perception*. (In the press.)

Morgan, M. J. & Thompson, P. 1975 Apparent motion and the Pulfrich effect. *Perception* **4**, 3–18.

Morgan, M. J. & Turnbull, D. F. 1978 Smooth eye tracking and the perception of motion in the absence of real movement. *Vision Res.* **18**, 1053–1059.

Pulfrich, C. 1922 Die Stereoscopie im Dienst der isochromen und heterochromen Photometrie. *Naturwissenschaften* **10**, 553–564.

Rashbass, C. 1961 The relationship between saccadic and smooth tracking eye movements. *J. Physiol., Lond.* **159**, 326–338.

Robins, C. & Shepard, R. N. 1977 Spatio-temporal probing of apparent rotational movement. *Percept. Psychophys.* **22**, 12–18.
Ross, J. 1974 Stereopsis by binocular delay. *Nature, Lond.* **248**, 363–364.
Ross, J. and Hogben, J. H. 1974 Short term memory in stereopsis. *Vision Res.* **14**, 1195–1201.
Ross, J. & Hogben, J. H. 1975 Pulfrich effect and short-term memory in stereopsis. *Vision Res.* **15**, 1289–1290.
Shepard, R. N. 1978 The mental image. *Am. Psychol.* **33**, 125–137.
Shepard, R. N. & Judd, S. A. 1976 Perceptual illusion of rotation of three dimensional objects. *Science, N.Y.* **191**, 952–954.
Shepard, R. N. & Podgorny, P. 1978 Cognitive processes that resemble perceptual processes. In *Handbook of learning and cognitive processes* (ed. W. K. Estes), pp. 189–237. Hillsdale, N.J.: L. Erlbaum.
Southall, J. P. C. 1962 *Helmholtz's treatise on physiological optics*, vol. 3, p. 251. New York: Dover.
Tyler, C. W. 1974 Stereopsis in dynamic visual noise. *Nature, Lond.* **250**, 781–782.
Tyler, C. W. 1977 Stereomovement from interocular delay in dynamic visual noise: A random spatial disparity hypothesis. *Am. J. Opt. Physiol. Opt.* **54**, 374–386.
Ward, R. & Morgan, M. J. 1978 Perceptual effect of pursuit eye movements in the absence of a target. *Nature, Lond.* **274**, 158–159.
Westheimer, G. & McKee, S. P. 1975 Visual acuity in the presence of retinal image motion. *J. opt. Soc. Am.* **65**, 847–50.
Westheimer, G. & McKee, S. 1978 Stereoscopic acuity for moving retinal images. *J. opt. Soc. Am.* **68**, 450–55.
Wist, E. R. 1970 Do depth shifts resulting from an interocular delay result from a breakdown of binocular fusion? *Percept. Psychophys.* **81**, 15–19.
Zöllner, F. 1862 Über eine neue Art anorthoskopischer Zerrbilder. *Annln Phys.* **117**, 477–484.

Discussion

H. B. Barlow, F.R.S. (*Physiological Laboratory, Cambridge CB2 3EG, U.K.*). I think that these results are fascinating because they seem to imply something that I did not previously suspect, namely that the visual system performs some kind of integration along a constant velocity path in space and time. That seems to me what is implied by a spatial shift and a temporal delay producing exactly interchangeable effects with regard to the perception of a vernier or stereo misalignment. But to reach this conclusion it is crucial to know that resolution is unimpaired by movement. It would be relatively simple for the visual system to *ignore* evidence of the blurring of moving objects, but the ability to extract high resolution information in spite of movement is quite another matter and a much more interesting one. So my question is: What was the resolution attainable in the task that Professor Morgan showed in figures 7 and 9, and was it affected by the variable filtering imposed; also, what was the velocity of movement?

M. J. Morgan. The question of acuity is indeed a central one, and I have added some relevant data in the written version of the paper to try to answer this point (§8 and figure 11). Briefly, the acuity for a temporal delay was about 1 ms in the stereoacuity task and 2 ms in the vernier offset task. Target velocity was 4.8°/s. These times would correspond to notional distances of 17 and 34″ respectively if the targets had been moving continuously. This acuity is not as good as that reported by Westheimer & McKee for a real spatial offset: they found a threshold of 6–10″ both for stereo and vernier acuity. Therefore, one cannot conclude that resolution is completely unimpaired by discontinuous motion. Moreover, vernier acuity was considerably improved if the observer tracked the target (figure 11). There is therefore loss of acuity, but one has to remember that the real offset is temporal, and that the spatial offset is only an average value. These studies of acuity are not very extensive, and clearly much more has to be done, particularly in comparing continuous and discontinuous movement. I have not yet examined effects of filtering upon acuity.

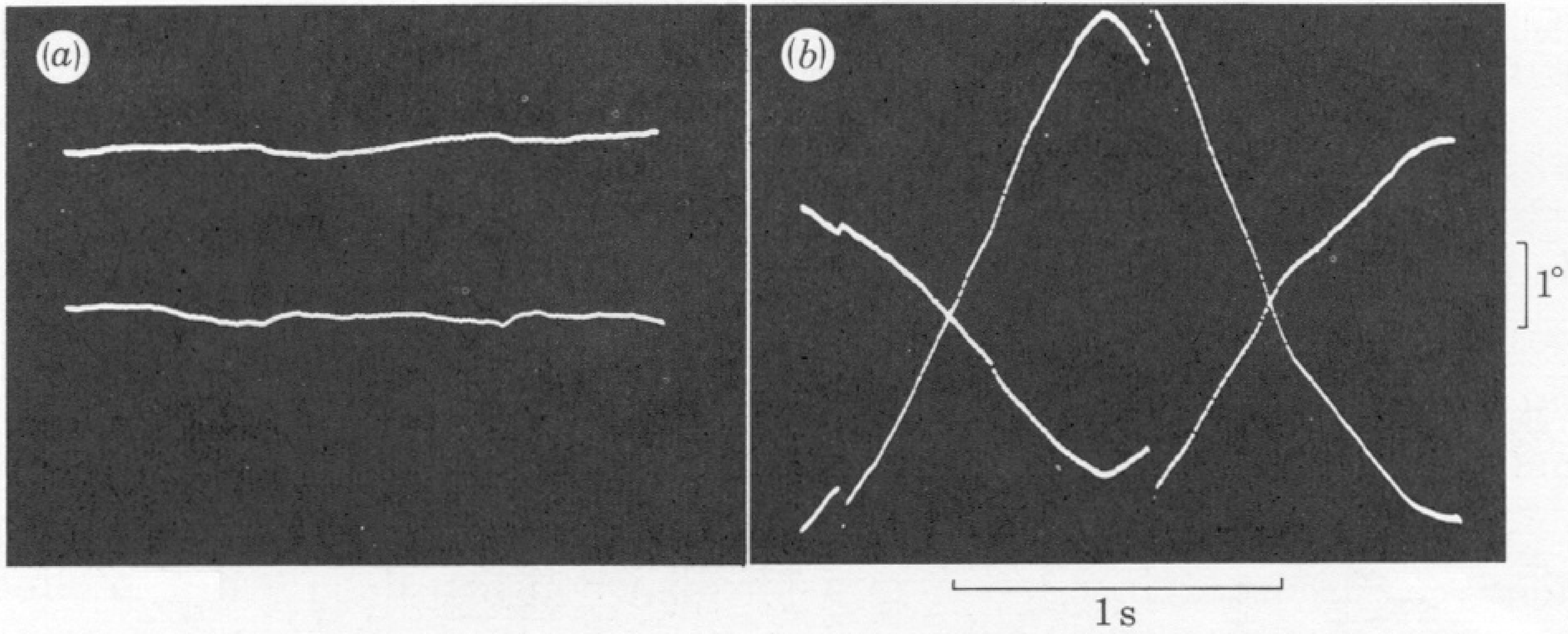


FIGURE 13. Binocular eye movement recordings from a subject (M.J.M.) engaged in a stereoacuity task (see figure 11). One of each pair of records represents movements of the left eye and the other the right eye. Record (*a*) was taken when the subject was instructed to fixate; record (*b*) when he tracked the moving targets. Records from the left eye are phase-inverted relative to the right because for both eyes records were taken from the temporal margin of the limbus.

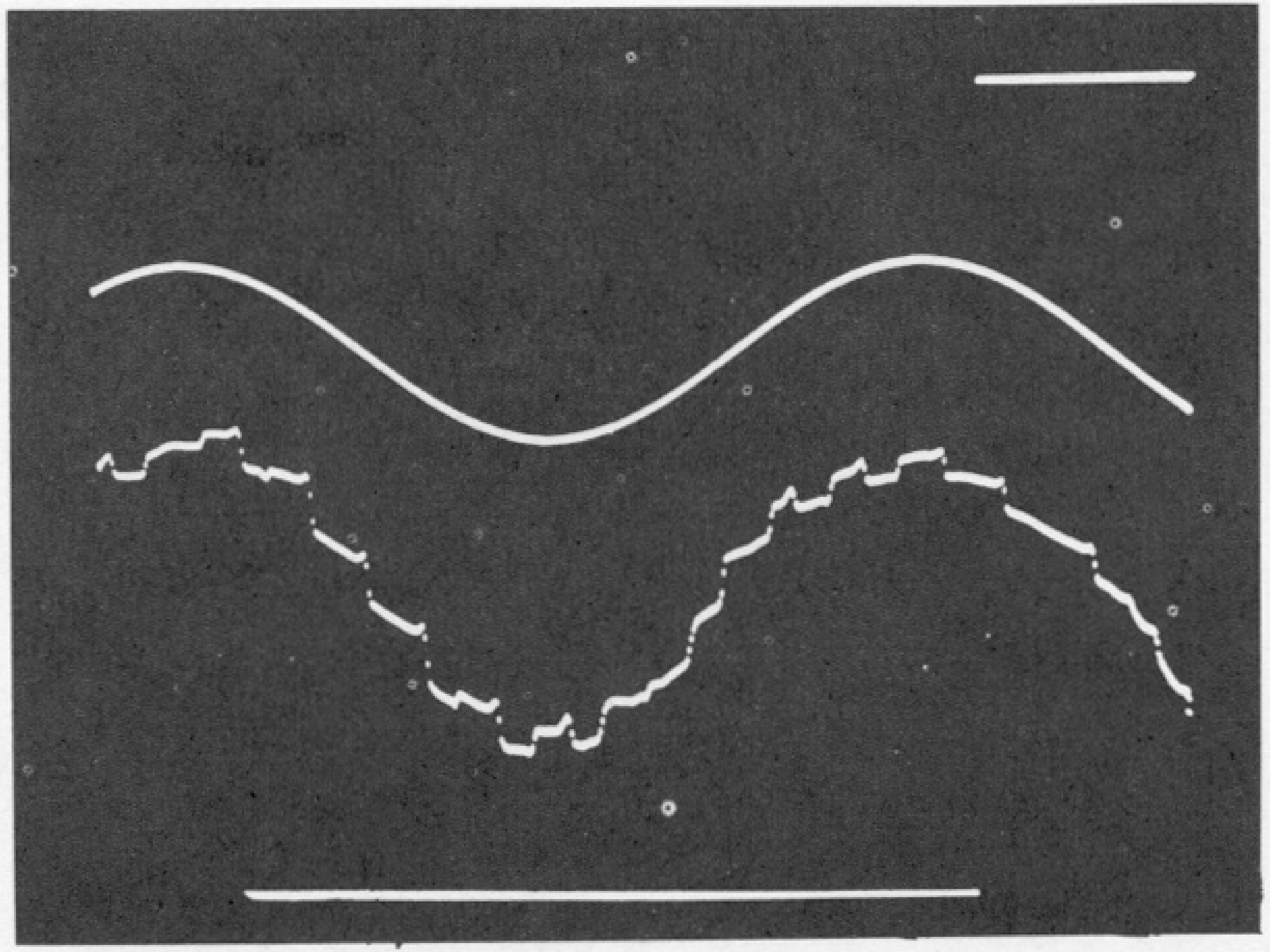


FIGURE 15. The record shows an example of the way in which sodium amytal (240 mg) disrupts smooth pursuit movements. The top record shows movement of the target (1 Hz), the bottom movement of the eye. Note the continual interruption of pursuit by saccadic eye movements. (Subject J. M. F.)