blending rather than selection occurs. This can be amusingly demonstrated by asking someone to stand opposite you on the other side of a half-silvered observation window, and manipulating the room illuminations until the reflected and transmitted components are approximately equal. You will see a single face whose features are a startling blend of his and your own!

These prior findings suggested the use of videotape rather than mirrors, and of naturalistic episodes rather than stationary displays. Our principal hypothesis was that the subjects would easily be able to follow one episode and ignore the other. No special practice should be necessary, since everyone has had a lifetime of practice in observing events. Note, however, that few people have ever had to select among visually superimposed displays. If "selective attention" were a separate mechanism developed to "filter out" undesired signals, as some have supposed, long practice would be essential.

Although unrelated and optically superimposed displays do not occur in ordinary vision, they are produced by certain optical devices. Microscopes, telescopes, gun sights, television cameras, and similar instruments are usually constructed so that the critical field of view is presented only to one eye. Users are often advised to keep the other eye open as well, though everything in its field must be ignored. Such devices are therefore dichoptic: The two eyes are stimulated independently. Their very design seems to reflect the assumption that separate percepts require separate sensory channels. We believe, however, that selection is most naturally based on objects of attention rather than on channels of sense. To test this hypothesis, we performed our experiment in both a binocular and dichoptic mode. In the former case, both eyes saw the same optical display, on which two overlapping episodes were portrayed; in the latter, one episode was presented to the left eye and the other to the right.

Our experiment was specifically aimed at answering the following questions. First, how difficult (or how easy) is it to follow one episode and ignore another when both are presented at the same optical distance in the same binocular visual field? Second, does the substitution of dichoptic for binocular presentation change the difficulty of the task? Third, how complete is the selection? Will unusual events be noticed if they are not part of the episode being observed? Fourth, is it possible to follow two independent episodes at the same time if instructed to do so?

METHOD

The aims of our experiment imposed definite constraints on the choice of stimulus material. To determine whether subjects can attend to certain events without being distracted by others, one must have an operational measure of attention similar to that which overt "shadowing" provides in the selective listening studies. In other words, we needed an episodic framework in which a certain type of event could occur repeatedly but irregularly, so we could ask observers to respond to each occurrence. Moreover, we needed two such episodic frameworks, very different in kind and scale, so that subjects would not confuse them even when they were presented simultaneously. A further constraint was added by our desire to avoid artificial stimuli such as flashing lights and moving lines in favor of meaningful objects and events.

After considerable pilot work, we settled on two types of episodes. The first of these was the handgame. This familiar game requires two players. Player A holds his hands out, palms up, and B places his hands, palms down, over A's. Then, with a quick movement, to slap B's hands by bringing one or both of his own up, over, and down onto B's. B tries to elude him by jerking his own hands back or sideways. If A succeeds in his stroke they replace their hands in the initial position and A is free to try again. If he fails they exchange roles; B takes the bottom (aggressive) position and attempts to slap A. Successful aggressive play requires quick movements at unexpected intervals; successful defensive play requires that one jerks one's hands away quickly as the other begins his stroke. A skilled aggressive player will often make feints: small movements of his hands that induce the other to jerk his hands back for nothing, no actual attack having been made.

To make the handgame videotapes, the two authors played the game with each other. An associate operated the camera from a position somewhat higher than their hands; the field of view included only the hands and forearms of both players. The background was a chalk board, blank except for an index number in the top center of the screen, which was changed for each new episode. The videotape of each episode lasted about 74 sec. One of the players tapped the index number twice with one hand at 6 and again at 10 sec to provide synchronization signals; play began after 14 sec and lasted for 1 min. In a "slow" episode, the players made a total of about 20 (occasionally 19) attacking strokes irregularly during this min; in a "fast" episode, about 40 strokes. Except for their control of the average rate, (aided by the cameraman who called out the time at 15-sec intervals), they played naturally and attempted to score hits.

For an experimental subject who attended to the videotape of a handgame, each attacking stroke was a target event. He kept his right forefinger on a microswitch, which he was to press at each synchronization signal and then whenever attacking stroke occurred, whether or not a hit was scored. He had to be careful not to be deceived by feints. An event recorder connected to the microswitch provided a full record of his responses on paper tape.

The second type of episode was the ballgame. It was played by three men with a basketball, which they threw back and forth to one another
while moving around as irregularly as possible in the camera’s field of view. The light-colored walls of the room, including a door and floor-length drapes, provided the background. The actual playing area was some 12 × 15 ft (37 × 46 m) but it was not rectangular since the camera’s field is wider at greater distances. The distance of the camera was such that the image of a player took up about 80% of the height of the screen when he was in the foreground, and about 50% at the rear. A small index number, changed for each episode, was visible in the lower right corner of the screen.

One player bounced the ball off the floor twice at 8 and again at 12 sec to provide synchronization signals; play started at 14 sec and continued for 1 min. In a “slow” episode the ball was thrown about 20 times; in a “fast” episode, about 40 times. Except for their control of the average rate, the players tried to throw and move irregularly and unpredictably. They also occasionally dribbled the ball and made fake throws. (On a few occasions when two players were in a direct line with the camera, a real throw could not be seen on the videotape. The occurrence of such a throw immediately became evident, however, because a new player came into possession of the ball.) An experimental subject, who attended at the videotape of a ballgame, kept his left forefinger on another microswitch. He was to press it at the synchronization signals, and thereafter whenever the ball was thrown from one player to another.

It should be noted that the “action” of a ballgame occurred more or less all over the screen. It consisted of the exchange of the ball, whose video image was small, between players whose images were large and unpredictably located. The “action” of the handgame, in contrast, always took place near the center of the screen, where the images of the players’ hands loomed rather large. Figures 1a and 1b, traced from momentary video images, suggest these sizes and positions.

In addition to the standard and fast episodes of both games, we videotaped a number of “odd” episodes, to determine whether unusual events in an unattended episode would be noticed. The odd events used in this experiment included: (a) handgame-handshake: after 15 sec of normal handgame play at the standard rate, the players stopped and shook hands; they resumed play for 30 sec, shook hands again, and resumed play once more; (b) handgame-disappearance: after 20 sec of standard play, one of the players threw the ball out of the field of view; all three players remained, moving as before and making fake throwing movements with an imaginary ball for 20 sec until the real ball was thrown back in and normal play resumed; (c) handgame-throw: after fifteen seconds of standard handgame play, one of the players stopped, picked up a small ball and threw it to the other, who quickly threw it back to the first; they resumed handgame play for 30 sec, repeated the ball-throwing sequence, and once again resumed play; (d) ballgame-exchange: after 20 sec of normal play, one of the (male) players slipped offscreen and was replaced by a woman, and after a few sec this happened to the second man, then the third player as well; the three women continued to play for about 20 sec, after which the men slipped back in one at a time to replace them and play normally.

All episodes were videotaped, without sound, using Sony CV-series half-inch equipment. All the handgame episodes were recorded on one videotape, and all ballgame episodes on another; the videotapes were marked so that the onset of each episode was easy to find. They were played back with two separate videotape recorders, each connected to its own monitor. The monitor screens were 5½ in. high by 7½ in. wide (140 × 184 mm). In one condition, the two monitors were optically superimposed with the aid of a half-silvered mirror set at a 45° angle to the subject’s line of sight. One screen, displaying the handgame, was directly ahead of the subject and was viewed through the mirror. The other, displaying the ballgame, was at his left and was reflected in the mirror (Fig. 2a). The subject’s head position was fixed with a chin rest.
Fig. 2. Schematic top view of the optical arrangement. A, binocular condition. B, dichoptic condition. (Abbreviations: S = Subject; TV = Television monitor; 1 for ballgame; 2 for handgun; M = half-silvered mirror; OS = opaque screen; AP = apparent position of superimposed images.)

Both monitors were at an optical distance of 35 in. (889 mm) from his eyes subtending a visual angle of 12° horizontally and 7° vertically. With this arrangement he saw what appeared to be a single television screen. There was no way to tell on which monitor an episode was actually being shown, except that the small index number on the ballgame videotape was reversed left-for-right by the mirror. Since the subject saw the same (superimposed) episodes with both eyes in this arrangement, we termed it the binocular condition. A traced outline of a typically overlapped pair of episodes is shown in Figure 1c.

In the dichoptic condition, two mirrors were placed in a V-shaped arrangement (the apex toward the subject) so that each one was at 45° to the line of sight (see Fig. 2b). When his head was properly positioned, the subject could see a reflection of the monitor which displayed the ballgame with his left eye and a reflection of the handgun with his right eye. As in the binocular case, the images of the two monitors were perfectly aligned so that one had the impression of looking at a single screen. Again each monitor was at an optical distance of 35 in. (889 mm). (We made no attempt to measure or control for eye dominance, but every subject was asked whether two episodes seemed about equal in clarity and brightness. All replied affirmatively.) In both conditions, the actual optical arrangement was fully open to view. All the observers were quite aware that there were really two video screens which were being made to appear as one.

The subjects were 24 Cornell undergraduate students, 15 males and nine females, who volunteered for the experiment. Half were assigned at random to the binocular, and half to the dichoptic condition. Each subject was presented with a series of 10 trials in a fixed order. In the first two trials, a fast (40/min) episode was presented on one monitor while the other remained blank. The ballgame was shown on trial 1, the handgun on trial 2. These two trials served to establish baselines of performance. On trial 3, two episodes were presented simultaneously, but the subject was instructed to respond to the ballgame and ignore the handgun; on trial 4 he had the opposite task. Fast episodes were used in both of these trials as well. On trials 5 and 6, the subject was to respond to target events in both episodes, ball-throws with his left hand and hand-slashes with his right. Slow (20/min) episodes were used, so that the combined response rate required was no higher than in trials 1–4. On the last four trials the subject was again instructed to follow one game and ignore the other. Slow episodes were used. The unattended episodes actually included the “odd” events described above, in the order indicated, but no mention of them was made. After each of the 10 trials, the subject was asked questions about the difficulty of the task, about his strategies, and about anything else he might want to report. In a more extensive inquiry at the end of the session, the experimenter specifically mentioned the four “odd” events and asked if the subject had noticed them.

Each videotaped episode has its own particular characteristics and its own pattern of difficulty. To counterbalance this variable, episodes used in one condition for half the subjects were used in a different condition for the remainder. Thus, although the order of conditions was the same for everyone, the individual episodes were shown in two different orders. Each order was used with six binocular and six dichoptic subjects. Episodes used in trials 1 and 2 (without accompaniment) in one order were used in trials 3 and 4 (with a distracting episode of the other game) in the other. Episodes used in the divided-attention trials 5 and 6 in one order were used as “covers” for the odd episodes, trials 7–10, in
the other order. Although some individual episodes did indeed produce more errors than others, the pattern of results was the same for the two orders, and they have been combined in the presentation of data.

The event-recorder tapes produced by subjects responding to a given episode were scored by aligning them against a standard tape. In preparing this tape, each of us independently viewed the episode in the undistracted condition (i.e., with the other screen blank), responding to the synchronization signals and the target events just as the subjects were to do. If we felt unsatisfied with our performance, we repeated it. This task is very easy, and so the two independent records we prepared for each episode were in close coincidence. When discrepancies occurred (none was over 0.25 sec) the earlier response (i.e., the shorter reaction time) was selected. Thus one standard tape was prepared on the basis of our two independent records. In scoring a subject's tape, his responses to the synchronization signals were placed in coincidence with the corresponding responses on the standard tape (this had the effect of correcting for systematic differences in reaction time). Each subsequent response was scored as a "hit" if it occurred in a critical interval, defined as extending from 0.25 sec before to 1.50 sec after each response on the standard tape; otherwise it was a "false alarm." A "miss" was scored if the subject failed to respond during a critical interval.

In order to determine whether the length of the critical interval affected the pattern of results, the tapes were also scored with a shorter one, extending from 0.25 sec before the standard to 0.75 sec after it. Any response with a measured latency between 0.75 and 1.50 sec, which had been scored as a "hit" with the longer interval, had to be counted as both a "miss" and a "false alarm" when this more stringent criterion was applied. Altogether 73 such slow responses occurred in the ballgame, and 29 in the handgame. Since the pattern of results was essentially unchanged by these additional "misses" and "false alarms," they are not included in Table 1. The analyses below are based on the longer critical interval.

**RESULTS**

Table 1 shows how the various conditions affected the subjects' ability to follow the episodes. Results for viewing the handgame and the ballgame are presented separately. We have not combined misses (M) and false alarms (FA) into a single index of sensitivity, because the false alarms do not occur randomly. Many of them were produced by specific events, such as feints executed by the handgame players or aborted throws in the ballgame. Since the players had more time to make these misleading gestures in the slow episodes (5-10) than the fast ones (1-4), more false alarms resulted. This should not be interpreted as an effect of rate or condition per se.

The first row of Table 1 shows that one can easily follow episodes of

---

**Table 1**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Ballgame</th>
<th>Handgame</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>M</td>
<td>FA</td>
</tr>
<tr>
<td>Attended to one 40 min episode shown alone (trial 1 or 2)</td>
<td>480</td>
<td>1</td>
</tr>
<tr>
<td>Attended to one 40 min episode, ignore the other (trial 3 or 4)</td>
<td>480</td>
<td>12</td>
</tr>
<tr>
<td>Attended to both 40 min episodes at once (trials 5 and 6)</td>
<td>479</td>
<td>12</td>
</tr>
<tr>
<td>Attended to one 40 min episode, ignoring the other (trials 7, 8, 9, 10)</td>
<td>479</td>
<td>12</td>
</tr>
</tbody>
</table>

Note: F is the number of events presented at one time (40 min or two 20 min episodes, summed over 12 subjects). A few trials were sometimes lost because only 19 events were shown or because the onset of the two videos were not properly synchronized. M is the total number of misses, FA of false alarms (see text for definition).
the types used here when they are presented alone. Binocular subjects missed less than 1% of the target events; for the 80 targets in trials 1 and 2, the median number of misses/subject was zero. Dichoptic subjects were nearly as good, with a combined miss rate of about 1.5%. There were few false alarms.

The second row of the table gives a clear-cut answer to our principal experimental question. Subjects had little difficulty in following the action of a given episode, even when another irrelevant one was superimposed on it. To be sure, there was a slight drop in accuracy compared to the control conditions. The median number of errors/subject rose to 1 (of 80 targets). Fourteen subjects missed more targets when following one episode and ignoring another (trials 3 + 4) than when following an episode shown alone (trials 1 + 2); six missed equally often in these two conditions; only four showed the opposite tendency. A binomial test indicates that this trend is significant at the 0.01 level. Nevertheless, miss rates in the superimposed condition remained under 3% for both games and both types of viewing. Further confirmation of this finding appears in the last line of the table, which gives the results of four additional trials under the same instruction, but with an event rate of 20/min. Again only about 3% of the targets were missed. Binomial tests show that the miss rate on these trials was significantly ($p < .01$) higher than on the control trials (1 + 2), but not significantly different from the earlier trials with the same instruction trials (3 + 4). The false alarm rate was higher, however, for reasons considered above.

The third line of the table shows a very drastic deterioration of performance when subjects were asked to monitor both episodes simultaneously. Twenty to forty percent of the target events were missed; the miss rate increased by a factor of eight for binocularly-viewed ballgame targets and by larger factors in the other conditions. Every subject missed more ballgame targets in the double task (trials 5 + 6) than in the single one (trial 3); $p = 0.03$. This was also true for the handgame. A great many false alarms occurred. The subjects’ own descriptions of the double task ranged from “demanding” to “impossible.”

Comparison of the left and right halves of Table 1 indicates that the subjects who viewed the episode dichoptically—one to each eye—missed more targets than those for whom the episodes were superimposed and presented to both eyes together. The binocular condition is superior in all four conditions of viewing the handgame, and in three of the four conditions for the ballgame. If this difference in miss rates reflects a real effect, it is probably due to spontaneous binocular rivalry. In any case, it is apparent that selection among visually-given episodes is made by presenting them to different eyes.

The “odd” events in the unattended episodes were rarely noticed, and then only in a fragmentary way. While watching the ballgame on trial 7, a single subject (of 24) spontaneously reported seeing a handshake in the handgame; only three others mentioned it in the postexperimental inquiry. No subject reported the disappearance of the ball which took place while he was watching the handgame in trial 8, although one later remembered “thinking that something was wrong in the ballgame.” In trial 9, the handgame players twice threw a ball back and forth while subjects watched the ballgame. Three of the viewers spontaneously reported seeing such a throw, and three more mentioned it in the inquiry, four of these six were among those who had noticed something odd on an earlier trial. Two other subjects remarked spontaneously that they had seen an unusual motion of the to-be-ignored hands—one even responded to the motion as if it were a throw in the ballgame, though he realized his mistake—but insisted that they had not seen a ball being thrown by the hands. In trial 10 three subjects, including two who had seen the ball-throw on the previous trial, remarked spontaneously on some aspect of the exchange of women for men in the unattended ballgame, and two others described some aspect of the exchange in the later inquiry.

All the comments were incomplete, in that they specified only partial aspects of the odd events. No subject saw more than one handshake or ball-throw, for example, although each occurred twice on its respective trial. Although a few subjects noticed something like “a girl’s legs and skirt” on trial 10, none could describe the whole exchange. Those who did notice anything were quite vague about what they had seen; one said “I thought I saw a different person, but I thought it was my imagination.” Half the subjects (12) gave no indication that they had observed or been influenced by any of the odd events at all. Indeed, the most common response to the inquiry was incredulity. Nearly everyone, including those who had picked up one or more fragments, expressed astonishment that the odd events described to them (or shown in a replay) had actually been presented. These events are normally quite obvious, and pilot subjects monitoring the odd episodes have always reported them. Nevertheless, they were rarely noticed in the experimental conditions.

CONCLUSIONS

Without any prior practice, it is easy to attend to one sequential episode and ignore another, even when both are presented in haphazard overlap. What defines one episode and distinguishes it from the other is not its distance or its clarity, or the sense organ involved, but its intrinsic properties and structure. Once picked up, the continuous and coherent motions of the ballgame (or of the handgame) guide further pickup; what is seen guides further seeing. It is implausible to suppose that special “filters” or “gates,” designed on the spot for this novel situation, block the irrelevant material from penetrating deeply into the “processing system.” The ordinary perceptual skills of following visually-given events, which develop in the first year of life (Bower,
are simply applied to the attended episode and not to the other.

One of the most important of these skills, of course, is the ability to follow the flow of events with one's eyes; to insure that the critical information falls on the foveae. Though we have no eye-movement records for the subjects of the study reported here, we can assume that they generally fixated the critical events of the episode which they were attending, and made pursuit movements appropriate to it. Thus much of the action of the to-be-ignored episode must have been imaged on the retinal periphery, and blurred by motion relative to the eye, while the action of the attended episode remained more consistently on the fovea.

For empirical and logical reasons, however, this difference cannot have been responsible for the successful selection we observed. Empirically, it appears that these episodes can be followed with peripheral vision alone. We have conducted pilot trials requiring subjects to follow one episode and ignore another while fixating a point at the edge of the screen. A videotape record of the subject's eye positions showed that only three or four inadvertent losses of fixation per trial occurred under these conditions. Nevertheless, while the observed miss rates in these pilot studies were somewhat higher than in the main experiment, selection remained relatively efficient. Thus, it is not peripheral registration per se which prevents subjects from noticing unattended events.

Moreover, logical considerations suggest that eye movements cannot be the principal mechanism of selective attention; they are more nearly one of its consequences. In order to follow one episode rather than the other with his eyes, a subject must already have selected the former rather than the latter. Without some prior basis for selection, his eyes would be drawn to features or movements in either episode with equal probability. An appropriate eye movement is as much a selective response as the press of a button. Therefore the assumed foveal status of an attended episode cannot explain how the subject managed to follow it, but rather requires explanation in itself. Such an explanation must eventually be given in terms of the general mechanisms of event perception, whatever they may be.

It is not surprising that dichoptic presentation afforded no advantage to our subjects. In general, attentive selection must depend on the structure of events, not on the separation of sensory channels. In the special case of vision, it is particularly obvious that the nervous system functions to provide an efficient combination of the inputs from the two eyes rather than a separation between them. Indeed, dichoptic viewing may present peculiar difficulties of its own, as the higher miss rate in our dichoptic condition suggests. If the scenes presented to the two eyes had been at different distances, as they are in many optical devices, conflicting cues for depth of accommodation might well have led to eye-strain and other problems. The design of some existing optical systems may not be optimal from this point of view.

Our results show that the subjects had great difficulty in keeping track of both episodes at once, and rarely noticed odd events in the unattended episode. There are several possible reasons for this. First, the unattended episodes must have been imaged peripherally much of the time. Given the poorer acuity of the peripheral retina, it may be that fewer details of these episodes were available to the observers. Second, our subjects had to choose between two responses in the double condition; disjunctive reactions are known to be more difficult than simple ones. A third possibility is more basic. Event perception may be so organized that when a particular structured flow of information is being followed, or a particular representation constructed, the perceiver cannot follow or construct an unrelated one. The results of the selective listening studies suggest that this is true, at least for certain levels of complexity and certain stages of learning. Further research will be needed to clarify these issues.

Finally, some introspective comments should be presented. The reader may wonder, as we did, whether the unattended episode "really disappears" or not. Can people fail to see what is directly before their eyes? The subjects' answers to such questions were somewhat inconsistent, and as unsatisfactory to them as to us. Moreover, no subject served in both the binocular and the dichoptic conditions, and so none was able to clarify his introspections by comparing the two. We were able to do this ourselves, however, and offer the following observations.

In the dichoptic case, the unwanted episode really does disappear (or parts of it do), and we can attend to its disappearance. This is the familiar phenomenon of binocular rivalry; the disappearances are so visible that one can count and report them. These fluctuations are most obvious when we look passively at the display. They are less apparent when we are deliberately attending to one episode rather than the other, and least of all if we are actually responding to it with the microswitch. Under these latter conditions we do not see the unwanted episode disappear; it simply "isn't there." In all dichoptic viewing we experience a certain feeling of effort when we switch our attention from one episode to the other, as if some inertia has to be overcome.

The binocular case is quite different. We would not say that the unwanted episode ever "disappears," although it is not really seen. There is no moment at which we are aware that it is going or has just gone out of sight. The term "disappearance" can refer either to the visible vanishing of images in binocular rivalry (as in our dichoptic case) or to the way a real object can leave the field of view, perhaps by going behind an obstacle or by being destroyed. No such disappearances are perceived in the binocular condition. It is also noteworthy that when we attend to one of the episodes and ignore the other, we almost always remain aware that "something else is going on." The total experience is not like looking at a single episode with the other screen turned off. Sometimes, how-