Perception & Psychophysics
1973, Vol. 13, No. 3, 451-454

Visual stream segregation*

ALBERT S. BREGMAN† and ANDRÉ ACHIM
McGill University, Montreal, Quebec, Canada

A visual analog of auditory stream segregation occurs when a dot, moving in discrete jumps, alternates between positions on two regular trajectories. At slow speeds, one dot in irregular motion is seen. At higher speeds, two dots are seen, each moving in a regular trajectory.

The present study concerns visual apparent motion. It is an attempt to show an analogy in vision to what Bregman and Campbell (1971) have referred to as "primary auditory stream segregation" (PASS). This refers to the capacity of the auditory system to take a complex waveform of sound that has been produced by several different co-occurring sources of sound (for example, instruments in an orchestra) and to decompose it back into its component simple sounds. Sometimes this process segregates a signal into two streams when it really all came from a single source. False segregation effects can be produced when a single instrument rapidly alternates between two lines of melody, a higher one and a lower one. If this is fast enough, each line of melody is heard as a separate entity. Composers in the Baroque period used this phenomenon to create interesting counterpoint effects in their music.

Bregman and Campbell (1971) spliced six tape-recorded sine tones into a single repeating loop, with each tone lasting for 100 msec. There were three high and three low tones in a mixed order. Ss heard two separate streams of sound, one including the three high tones and the other composed of the three low tones. They could not judge the position of the high tones relative to the low tones in the series. The two streams of sound were perceptually segregated—they could listen either to one or the other, but not to both. In unpublished studies, the first author has found that at *low* speeds (say 3 tones/sec), Ss can perceive all the tones as part of a single stream. At higher speeds, however (5-10 tones/sec), the high stream splits away from the slow stream.

There is, in vision, an analogous phenomenon. Suppose one illuminates two lamps alternately in a dark room. If these lamps are separated by an appropriate time interval, the viewer experiences "apparent movement" between them. This is the familiar "beta" movement extensively investigated by psychologists and reviewed by Aarons (1964) and Graham (1965).

*This research was sponsored by grants to the first author from the National Research Council of Canada (APA-127) and from the Defense Research Board of Canada (9401-40).

†Requests for reprints should be sent to A. S. Bregman, Psychology Department, McGill University, P.O. Box 6070, Montreal 101, Quebec, Canada.

Now let us describe a phenomenon we have frequently observed in our laboratory. Suppose that four lamps, spatially distributed as in Fig. 1 (left-hand side), are used as a display. They are in a vertical row, but grouped by twos—the two upper ones are close together and the two lower ones also are. They are illuminated in a continuously repeating cycle in the order 1, 2, 3, 4, as shown in the diagram: a high one, a low one, a high one, and a low one. An oscillating pattern of apparent motion will move through Positions 1, 2, 3, and 4, in that order. This is shown by the dotted arrows in the time graph in Fig. 1. If we now speed up the sequence, the apparent motion splits into two substreams, as indicated by the solid arrows. Here, positional grouping (high vs low) acts like pitch similarities in audition (high vs low), producing an organization into co-occurring streams. One light seems to move back and forth between Positions 1 and 3, and another will oscillate between 2 and 4. We call this phenomenon "visual stream segregation" (VISS) and postulate that it will respond to a wide class of variables in the same manner that PASS does. If true, this would suggest either that information processes at some higher level of the central nervous system were responsible for both effects or that the two sensory systems had evolved in similar ways for similar reasons.

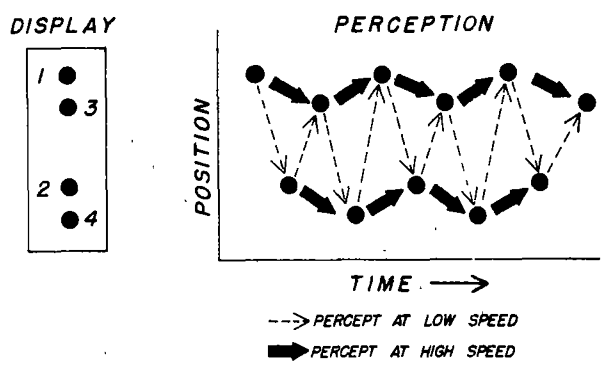

Fig. 1. Demonstration of visual stream segregation with four lamps, lit in the order 1, 2, 3, 4 at low and high speeds. The motion percept at low speeds is given by the dotted arrows, and at high speeds by the solid arrows.

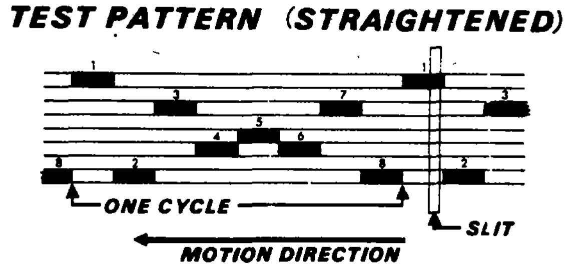

Fig. 2. Pattern of arcs drawn on the test disks straightened so that arcs are shown as straight lines.

A special case of VISS can be seen with only two lamps repeatedly lit in alternation. At slower speeds, an oscillating beta motion is seen. At higher speeds, no motion is seen. The two lamps are simply seen as two separate blinking sources of light. This separation of the perception of the changes in illumination of the two lamps can be seen as a special case of perceptual segregation, where Positions 1 and 3 of Fig. 1 are collapsed into one location and Positions 2 and 4 into a second one. Thus, the flickering upper limit of beta motion, when the lights are alternated too quickly, is not a *breakdown* in beta, but a redefining of the perceived objects in the situation as being two in number rather than one (with new potentialities for apparent motion which are not seen because the two "objects" are stationary).

The present study was an attempt to explore the effects of event rate and total duration of the display in VISS. We can view VISS as an organization of successive moments of visual experience into the perceptual framework of "objects in motion."

The visual system probably uses spatial proximity and temporal separation to decide on the grouping of successive events. One formulation of how spatial displacement and time are used was given by Körte (1915), reviewed by Graham (1965). Körte's "third law" stated that as the time between successive events (e.g., illuminations of lamps in our example) increased, to obtain smooth apparent motion the separation in positions would have to be increased, and vice versa. This can be seen as a law which would tend to produce veridical organization of discrete visual appearances. If an object disappears behind a fence, the longer the fence, the greater the time it should take before it comes out the other side.

In the case of VISS, displays with a higher event rate should segregate better, because, if Körte was right, the optimal spatial displacement to obtain good apparent motion will be less, favoring the short motions seen when the motion splits into substreams. To test this hypothesis, the present experiment varied the rate of changes in the location of a colored spot in the visual field.

A second variable investigated was how long it takes for stream segregation to occur. Obviously, it cannot occur until at least two dots are seen. But is that sufficient, or does the visual system require a succession of events before it consistently sees the sequence as two dots in motion rather than one? To answer this question, we varied the number of cycles that S saw of a repeating display and asked him to report "the way he saw the display at the end of each presentation."

METHOD

Stimuli

The S was presented with a sequence of visual events in a tachistoscope. Each one was a colored "dot" at a different discrete position on a vertical axis. Position on the horizontal axis was constant. The sequence of eight dots was such that the successive dots alternated between positions on two approximately sinusoidal trajectories which were 180 deg out of phase. If S followed the true temporal sequence, he would see an irregular, fast, up-and-down motion of a single dot. However, if S grouped spatially adjacent dots, he would see a smooth slow, up-and-down motion of two separate dots, which would cross one another at the center of the field or would approach one another and then bounce apart (Fig. 2).

In this way, in addition to using grouping by proximity, we used the continuity of the dots on two different smooth trajectories to promote the segregation of two perceptual objects. Continuity on a trajectory is a factor known to promote *auditory* stream organization (Heise & Miller, 1951).

Conditions

The experimental variables were (1) the duration of the dot at each position, and (2) number of cycles of the eight-dot pattern seen. The durations were 227, 114, and 57 msec per position, and the number of cycles were 1, 2, or 4. This gives the nine conditions shown in Table 1. In this table, the diagonals contain conditions with equal total duration of the display. Hence, we can look also at the effects of total duration.

Each S, tested individually, saw all nine conditions 10 times each, for a total of 90 trials. Conditions were randomized in blocks of nine, in which each condition appeared once. They were rerandomized for each S.

Procedure

The S was first shown two movement patterns similar to the ones he would see on test trials, except that these two were unambiguous continuous real motion. One was one dot in motion and the other was two dots in motion. He was told that on test trials, the perception might not be as clear as in the illustration, but that he was to decide whether he saw one dot in motion or two, recording his answer by checking one of two boxes ("one" or "two") on a rating sheet and checking his confidence on a 6-point scale ranging from "very certain" to "very uncertain." He was informed that the exposures would vary in duration from less than ½ sec to 7 or 8 sec, and that if his perception changed during the interval he was to report the last one. He knew there would be a trial every 30 sec.

On each trial, S saw a preexposure field, containing a blank slit, for 2 sec. This served both as a warning signal and as a fixation point. Then he saw a movement pattern appear for some time behind (i.e., through) the slit. Then he saw a totally blank field for 1 sec. He then made his written response.

The onsets of trials were at 30-sec intervals, with a rest interval between blocks of nine trials of from 1 to 2.5 min while E arranged the stimuli for the next block.

Apparatus

A mask with a narrow vertical slit was placed over a record-player turntable so that arcs of circles drawn on paper disks would appear as stationary dots in the slit. If arcs of circles of different radii followed one another on the disk, a series of different dot positions would be seen in the slit. A test pattern of eight consecutive arcs was drawn once, twice (each arc half as long), or four times (arcs again half as long) on paper disks. On a nominally 33 rpm turntable, these disks generated the same movement pattern at three different speeds: approximately 227, 114, or 57 msec per dot, respectively. Figure 2 shows the pattern of arcs on the turntable. The arcs were drawn as follows. Nine concentric circles were drawn 5 mm apart on paper disks, forming eight circular 5-mm-wide tracks between the circles. The tracks were numbered from 1 to 8 from the inside. They were divided into 8, 16, or 32 equal-length intervals (arcs). For each consecutive interval, one track was filled with red ink, in the following pattern: 8, 1, 6, 3, 4, 3, 6, 1, which was repeated when necessary. This is shown diagrammatically in Fig 2.

The training display (unambiguous movement) was made out of arcs just as in the ambiguous display except that diagonal lines were drawn on the disks connecting some of the arcs, producing continuous (but somewhat stepwise) changes of dot position as the disk rotated (see Fig. 3). In the unambiguous "one-dot" case, eight successive arcs were drawn, just as on the test disks except that consecutive arcs were joined by diagonal lines. In the "two-dot" case, two arcs were present in the slit at once, one for the upper dot and one for the lower dot, except where the two dots' trajectories overlapped in the center position.

Both the unambiguous illustrations ("one dot" and "two dots") were drawn on the same disk; thus S saw both through slits in a mask over a turntable at the same time. This display was viewed directly, with the turntable sitting on a table. The dots took a new position every 167 msec, except for the lower dot at its lowest position in the two-dot case, which stayed for 334 msec at this position.

The apparatus for the test trials was a Polymetric three-field tachistoscope (V-0565-3A-TR). The Field 2 back was removed and replaced with a 45-deg mirror so that a turntable apparatus covered by a mask with a slit in it appeared as if vertically mounted at the end of the visual tunnel when Field 2 was lit. The Field 1 back was removed and replaced with a replica of the mask which covered the turntable in Field 2. The only difference was that this slit had a blank sheet of paper a few millimeters behind it. Each slit was about 7.6 cm long and .2 cm wide and was recessed in a 2.5 x 7.6 cm rectangle.

The slit in Field 1 served as a fixation point. When the illumination shifted to Field 2, it appeared that a series of dots suddenly appeared behind the unchanged slit. Since the slit was 2 mm wide and the arcs were 5 mm thick, each dot was actually a vertically oriented 2 x 5 mm rectangle. There were eight 5-mm-thick arcs: thus, the display "moved" within a region subtending approximately 3 deg 13 min at the eye, the visual path being 71 cm long

Subjects

The Ss were 10 university students, 5 males (2 with eyeglasses) and 5 females (3 with eyeglasses), who were paid for participating

Table 1
Total Time of the Display (in Seconds) in Each of the Nine Conditions

Number of Cycles	Time at Each Position (in Seconds) .057	.114	.227
1	.46	.91	1.82
2	.91	1.82	3.64
4	1.82	3.64	7.28

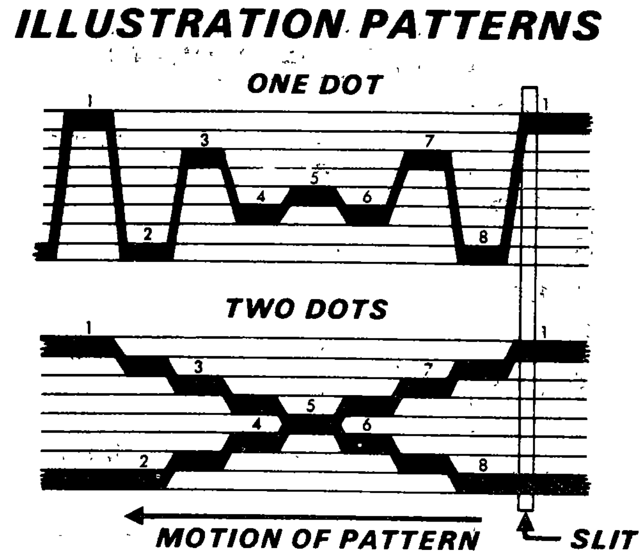

Fig. 3. Unambiguous illustration patterns which generated continuous changes in the position of one or two dots in motion.

RESULTS

Each S gave two responses on each trial, a judgment of "one" vs "two" dots and a rating of confidence on a 6-point scale. These two responses were converted to a single number ranging in steps of .1 from 1.0 to 2.0, the number 1.0 referring to a "very certain" one-dot judgment and 2.0 referring to a "very certain" two-dot judgment. This was done as follows: the judgment of "very uncertain" was assigned the number 1.5 regardless of whether S had judged "one" or "two" dots to be present; then the points on the certainty scale were taken as steps of .1 towards 1.0 or towards 2.0, depending on whether S's judgment was "one" or "two." The resulting score is referred to as a D score.

The mean results for the nine conditions are shown in Fig. 4. Vertical bars extend one standard error above and below the means (N = 10). It is clear that as the rate of presentation increased, the D scores increased. This orderly increase held for all 10 Ss in the one-cycle condition, for 9 out of 10 Ss in the two-cycle condition, and for 9 out of 10 in the four-cycle condition. In the three conditions (1, 2, or 4 cycles) combined, all 10 Ss showed an increase in number of dots as speed increased.

The result for number of cycles is less clear. Three Friedman two-way analyses of variance by ranks were performed on D scores, one at each level of speed. Only the one at 57 msec/dot was significant [$\chi^2(2) = 7.35$, $p < .05$].

The cells with equal total time of presentation show grossly different results, depending entirely on the speed at which the presentation occurred. Thus, total time per se seems to be unimportant.

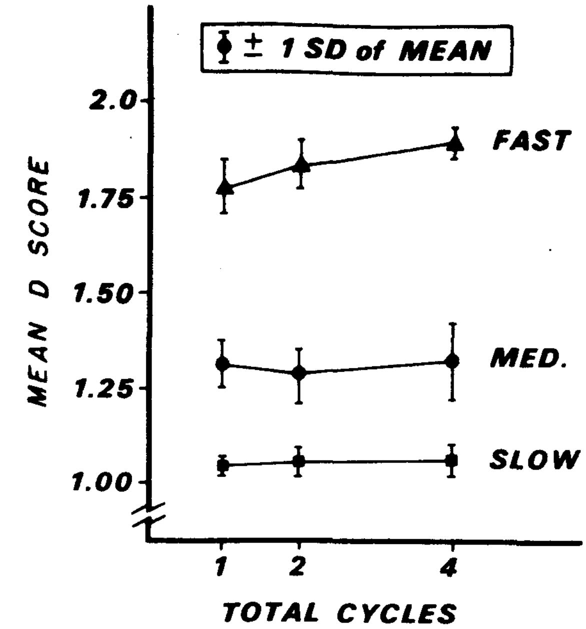

Fig. 4. Mean D scores for the nine conditions. Vertical bars extend one standard error above and below the means (N = 10).

DISCUSSION

The results show clear-cut effects of speed upon VISS when visual angle of the displacement is held constant. This is consistent with Körte's third law, as described in the introduction.

The results of number of cycles were ambiguous. There seemed to be an effect at the fastest speed, but not at the slower ones. Furthermore, the total time of the display seemed unimportant in VISS. However, we should notice that the one place where the number of cycles presented made a difference was at the fastest speed and that these conditions were those that took the least time to present (as a group). The shortest condition in this experiment presented eight dots at 57 msec/dot and lasted .46 sec. It is possible that most of the evidence required by the visual system to stabilize on a "two-dot" description of the input occurs with less than eight displacements of the stimulus or in less than .46 sec. Thus, the question of how long it takes for a two-object percept to emerge has not been settled. It would require a shorter cycle (2 to 4 dots) in a similar experiment to decide this issue. At least we can suggest that the present evidence justifies some further study.

Thus, we have displayed a phenomenon in vision analogous to auditory stream segregation (Bregman & Campbell, 1971) where visual angle plays the same role as pitch differences in producing parallel perceptual streams. In audition, the parallel organizations are perceived as "sources" and the sequential organization as melody or, more generally, as a patterned sound stream. In vision, the parallel organizations are perceived as "objects" and the sequential one as motion. However, perceptual phenomena are similar in the two cases.

REFERENCES

Aarons, L. Visual apparent motion research: Review, 1935-1939, and bibliography, 1955-1963. Perceptual & Motor Skills, 1964, 18, 239-274.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. Journal of Experimental Psychology, 1971, 89, 244-249.

Graham, C. H. Perception of movement. In C. H. Graham (Ed.), *Vision and visual perception.* New York: Wiley, 1965.

Heise, G. A., & Miller, G. A. An experimental study of auditory patterns. American Journal of Psychology, 1951, 64, 68-77.

Körte, A. Kinematoscopische Untersuchungen. Zeitschrift für Psychologie und Physiologie der Sinnesorgane, 1951, 72, 193-296.

(Received for publication January 10, 1973; accepted January 26, 1973.)