The mental representation of movement when static stimuli are viewed

JENNIFER J. FREYD
Stanford University, Stanford, California

If the representation of movement is a fundamental organizing principle of cognition, as hypothesized here, it should be possible to find cases in which static stimuli induce a dynamic mental representation. Subjects viewed frozen-action photographs, and their memory for these scenes was tested. They found it harder to reject distractors when the distractors were photographs of the same scene shot later in time than when the distractors were photographs shot earlier in time. In a second study, an asymmetry in goodness of apparent motion was found between forward and backward action sequences. Both results support the hypothesis that people represent the motion implicit in a photograph.

The human perceptual system is strikingly able to detect and recognize the motion of dynamic stimuli. This has led some researchers to suggest that the visual system may be organized for analyzing changing things, not static things (see Gibson, 1966; Johansson, 1975). The hypothesis presented here is that the representation of motion, or change, plays an important role in human perception and cognition, independent of whether the stimuli are dynamic or static. This hypothesis is in concert with Shepard's (1981) suggestion that the perception of objects depends upon a knowledge of the objects' possible transformations.

Some studies have shown that movement perception may be important in cases in which it was previously thought that mainly static, or configural, information was used. For instance, although we usually assume that we recognize an acquaintance by recognizing his or her facial features, body shape, and other aspects of physical appearance, in fact viewers can recognize themselves and friends without that configural information if abstract movement information is provided: Cutting and Kozlowski (1977) have shown that a person can be identified in the dark by friends if the person has a small number of tiny lights attached to his or her body and he or she moves. Similarly, some linguistic information can be conveyed, in the dark, to perceivers of American Sign Language, if the signer has a few tiny lights attached to his or her fingers and the hands move (Poizner, Bellugi, & Lutes-Driscoll, 1981). The claim that movement perception is fundamental to perceptual organization is also supported by a recent demonstration that one of the important ways that infants determine what makes a given object distinct and unified is by determining what moves as a single object (Spelke, 1982).

Perhaps the strongest test of the importance of movement perception would be the case in which, technically, no movement is actually present, but is simply implied. That is, if it can be shown that the representation of movement occurs under static conditions, it will indeed suggest that the mental representation of movement is a fundamental organizing principle for human perception and cognition. This is different from "fooling" motion detectors, as in apparent motion or the autokinetic effect. Those are perceptual illusions of movement rather than cognitive representations of movement.

In a recent study (Freyd, in press), I tested the hypothesis that, in perceiving handwritten letters, we represent dynamic information although only static information is present. I proposed that processes of recognizing static handwritten letters might use tacit knowledge of the manner in which those letters were drawn. In the experiment testing this hypothesis, subjects saw artificial characters drawn in real time and were later asked to identify distorted versions of the same characters presented statically. Subjects were faster on static characters distorted in a manner consistent with the drawing method they had witnessed than on static characters equally distorted but inconsistent with the drawing method, suggesting one way in which humans can use dynamic information in the perception of static forms.

The research presented here tests the proposal that the representation of movement also occurs...
under static conditions using “frozen-action” photographs. It was hypothesized that photographs that capture or “freeze” some object in the process of motion might induce a mental representation of movement. That is, when a person views a photograph of an object undergoing a unidirectional movement, the person might represent the object as continuing in motion. If this were true, we would expect that the person's memory for the photograph would be distorted in the direction of remembering the object as being further along the path of movement. (In general, a mental representation of movement need not necessarily have a preferred direction; however, in the case of irreversible motion, as used in the following experiments, we might expect this asymmetry to be implicitly represented in the form of a preference for forward motion.) This prediction was tested in Experiment 1 by comparing responses to distractors that were photographs of the same scene but shot slightly later in time, with distractors that were photographs of the scene shot slightly earlier in time. The hypothesis was also tested using an apparent motion paradigm in Experiment 2. If the photograph does induce a mental representation of movement in one direction, then there should be an asymmetry in goodness of apparent motion between forward and backward action sequences.

**EXPERIMENT 1**

If the perception and representation of movement serves an important mental organizing function, then, when people are presented with certain static displays, they may form a related dynamic mental representation. They could do this by using their knowledge of dynamic processes in the world and their encoding of implicit dynamic information in the static stimulus. One situation in which people might perceive implicit change is when they are viewing photographs of “frozen” motion, in which implicit change could mean the movement an object would be undergoing were it to be unfrozen. In other words, when people look at, say, a photograph of a tennis player serving, they might understand the motion in the picture by forming some sort of dynamic representation of it. For instance, in the mental representation of the tennis player, the racket might move forward and continue the serve. It was this possibility that motivated Experiment 1.

Irreversible action sequences (including unidirectional human and animal movement and natural processes that move toward greater entropy) were filmed with a movie camera, and pairs of individual frames were selected to use for stills (see Figure 1). The frames in these pairs were separated by 55 to 500 msec in the real world. In the experiment, individual stills were presented to subjects. They were asked to look at one frame for 250 msec and to hold it in memory for another 250 msec, and then to determine whether the second frame was “same as” or “different from” the first.

My prediction was that subjects would “unfreeze” the frozen motion implicit in the frame that they were to memorize by creating a dynamic mental representation of the static photograph; in other words, they would anticipate the motion in the scene. Thus, assuming that discrimination becomes more difficult as the similarity between items increases, subjects should find it more difficult to identify as “different” the distractor frames taken slightly later in time than they would the frames taken slightly earlier in time. If the backward discrimination is easier, subjects should be faster and/or more accurate on those trials than they would be when presented with a forward discrimination.

**Method**

**Subjects.** Sixteen Stanford undergraduates received course credit for their participation in this study.

**Stimuli.** Stimuli were 30 pairs of “before” and “after” photographs. The pairs were created by making glossy stills from
selected frames of movie film. Eight basic action sequences were used: waves crashing, a person walking, a person jumping off a wall, a person throwing a light bulb to the ground, a person tossing a ball, a person pouring water out of a cup onto the ground, a cat playing with string, and a seagull flying. The number of pairs from each action sequence varied from one (a person walking) to eight (a person pouring water out of a cup). In the selection of pairs, an attempt was made to pick shots in which the "motion" was unambiguous. The real-world time separation between frames also varied—from 500 msec (e.g., waves crashing) to 55 msec (e.g., person tossing a ball). The pairs were chosen so as to make the experimental task very difficult, but not impossible. Thus, the real-world time separation chosen depended mostly on how quick the action was: Waves crash slowly, but people toss quickly. An equal number of filler "same" pairs were also made. These 30 pairs were selected from the same action sequences, and in the same ratios. There were also two same and two different practice pairs. All individual shots were then mounted on cards that fit into the tachistoscope used for the experiment.

Procedure. Each subject was seated, with his or her head against a visor, in front of a tachistoscope. The subject rested a finger from each hand on the two response keys in front of him or her. All subjects were told that the task was to look at a photograph for 250 msec, hold it in memory for 250 msec, and then to decide as rapidly as possible whether a second picture was "same as" or "different from" the first. It was explained to the subjects that the "different" pairs were very similar, that the difference was very subtle. They were told to expect the scene to be the same within any pair and the difference, if there was one, to be a small one within the scene. There was no mention of "movement" in the verbal instructions. The subjects were given a few practice trials with feedback. If the subjects made any errors on the practice trials, they had to go through the practice again in a new order, and they went on to the experimental trials only when they had correctly responded to all of the practice trials. The subjects were told to guess "same" if they were not sure. Speed was stressed over accuracy; subjects were told that long reaction times could not be used as data.

The subjects saw all 60 pairs once, with the order randomized between subjects. Of the 30 "different" pairs, the subjects saw 15 in real-world temporal sequence (before-after) and 15 in reverse order (after-before). The temporal sequence for any given pair was counterbalanced across subjects. Trials began when the experimenter placed the cards in the tachistoscope and the subjects verbally indicated that they were ready. Then the first photograph came on for 250 msec; this was followed by 250 msec of darkness and then by the second photograph. The second photograph stayed on until the subject pressed a response key. Responses and reaction times were then recorded. After each subject had responded to all of the 60 trials, the experimenter asked the subject whether any continuous motion between photographs had been observed. All subjects reported that they had not seen any apparent motion, but most subjects mentioned that the individual photographs were objects in motion. (Two experienced observers of apparent motion were asked to view the display, and both agreed with the subjects that no apparent motion was visible. In fact, no apparent motion can be seen with these stimuli unless the presentation period of the first member of the pair is about 1,000 msec.) The complete session took approximately 45 min.

Results and Discussion
As predicted, subjects took longer to indicate correctly that the second frame was different when the pair was in "forward," or real-world, temporal order (the mean across subjects was 847 msec) than when the pair was in "backward," or reverse, order (788 msec). The reaction time difference was found in 13 of 14 subjects and was statistically significant [t(13) = 2.49, p < .05]. Also, the difference between real-world and reverse order was obtained for pairs from seven of the eight action sequences, averaged across subjects. The overall error rate was 31%. Although there was a difference in error rates in the predicted direction, it was not statistically significant.

This result, that it is harder to reject distractor photographs from later in time than those from earlier in time, supports the hypothesis that people represent the motion implicit in the frozen-action scenes. Note that this finding cannot easily be accounted for by demand characteristics, since optimal performance on the memory task required that subjects hold a static image in memory. Moreover, the result also suggests an interesting possibility: if these photographs induce a mental representation that partly continues the motion, then this phenomenon should influence apparent motion under appropriate conditions. That is, there should be an asymmetry in the goodness of apparent motion between forward and backward motion, corresponding to the asymmetry found in Experiment 1. This possibility was tested in Experiment 2.

EXPERIMENT 2

Visual apparent motion is usually defined to be the perception of continuous motion when, in fact, the display is merely a flickering of two or more stimuli at certain time intervals. The classical apparent-motion display uses two lights that go on and off asynchronously, and the critical parameters for seeing continuous motion are degree of asynchrony and distance between the lights. However, apparent motion can also be found between shapes, which is a necessary condition for seeing continuous motion when watching a movie. Now, if, as Experiment 1 suggests, people form a dynamic mental representation of frozen-action photographs, then this should have an effect on the perception of apparent motion between photographs. Consider the following experimental setup, used in Experiment 2: Subjects are presented one shot from an action sequence for 1,000 msec, after which there is a short interstimulus interval (ISI) followed by a second shot from the action sequence. If the subjects begin mentally to continue the movement implied in the first shot, they should find it easier to perceive apparent motion between shots when the pair is in real-world temporal order.

Experiment 2 is somewhat similar to experiments by Jones and Bruner (1954) and Toch and Ittelson (1956). Both investigations were interested in showing that perceived meaning, as well as spatial and temporal factors, had an effect on apparent motion. Jones and Bruner, for instance, compared
perceived apparent motion using nonsense and meaningful stimuli. They found that meaningful stimuli led to the perception of smoother, or better, motion than nonsense stimuli did. However, they did not control for the simple perceptibility of the static stimuli; their illustrations suggest that the static nonsense stimuli are more confusing and harder to understand or perceive than the meaningful stimuli. Jones and Bruner also tested the hypothesis that the path of motion perceived depends on the meaning of the motion by comparing reported motion seen when two men threw balls at each other and when two men shot billiard balls at each other. In the throwing case, the subjects were more likely to see the balls as "crossing over" in midair, while in the billiard case, the subjects were more likely to see the balls as bouncing off at an angle. Unfortunately, the exact positioning of the balls before and after they passed, or collided, was not the same in the two cases. Between the two conditions, it was not only the context, or meaning that changed but also the angular distances. Similarly, the physical environment depicted in the two conditions differed: in the throwing case, the force of gravity should play a significant role in determining the path of the balls. Judging from the illustrations, there is no way the two thrown balls could have collided and then gone off in the directions indicated without violating conservation of momentum (especially given any effect of gravity). Even if the explanation of the result is that subjects see motion that conserves momentum, and not what they expect to see, as Bruner and Jones suggest, this is at least as interesting a result. Indeed, it suggests that dynamic mental representations may be constrained by laws of physical motion, a point considered in the general discussion.

Toch and Ittelson investigated the role of meaning in apparent motion with a simple experiment. They compared two displays, each containing three figures on photographic slides. One had three identical bombs pointing downward, and the other display had three airplanes pointing upward. In both displays, the figures were organized vertically. Toch and Ittelson hoped to find that by flashing the figures in either display in a "bab" sequence (i.e., the middle figure preceding the outside figures) they would induce apparent motion between the items. Moreover, they predicted that, with the bomb and the airplane, the perception of motion would be influenced by the meaning of the stimulus. Thus, they predicted, and found, that the bomb was seen as falling while the airplane was seen as rising. One of the major problems with this study is that there was no condition in which bombs were pointed up and airplanes down, a control for the possibility that objects move in the direction to which they structurally point, independent of meaning (both the bomb and airplane have an overall arrow shape). This distinction deserves to be tested empirically, a point I will come to in the general discussion. Still, it is an interesting finding and suggests that perceived apparent motion can be influenced by perceived meaning.

The Toch and Ittelson results, as well as the Jones and Bruner results, can also be interpreted as being consistent with the hypothesis presented here, that subjects mentally represent implicit motion in the picture, which, in turn, influences perceived apparent motion. Experiment 2 tests this hypothesis using the before/after photograph pairs from Experiment 1. If, as Experiment 1 suggests, the individual photographs induce a dynamic mental representation of movement that continues the motion implicit in the picture, then there should be an asymmetry in the goodness of apparent motion between forward and backward sequences. That is, members of the forward pairs should be a shorter "distance" from each other in terms of mental representation than members of the backward pairs. Such a finding might also be predicted from Farrell and Shepard's (1981) suggestion that apparent motion occurs along mentally represented pathways, with the added assumption that people represent the implicit motion of the first photograph.

Method

Subjects. Eight Stanford undergraduates were paid for their participation in this study.

Stimuli. Twelve before-after pairs were selected from the 30 used in Experiment 1. These pairs were selected by a research assistant who did not know which pairs produced the largest effects in Experiment 1. Instead his criteria were two-fold: First, he looked for pairs with the least ambiguous motion, and second, he looked for pairs which produced the best apparent motion for him. Two of the pairs were eliminated because pilot subjects claimed they did not see motion. Of the ten pairs used in Experiment 2, four were from the action sequence of a person jumping off a wall, three were from the sequence of a person tossing a ball, one was from the sequence of a person throwing a light bulb, one was from the sequence of a person pouring water, and one was from the sequence of a bird flying. Two of the pairs had a 55 msec real-world separation, six had a 110 msec separation, one had a 167 msec separation, and one had a 220 msec separation.

Procedure. Subjects were told that they were to judge goodness of apparent motion. They were given a four value rating scale: 1 meant no motion at all was seen; 2, jerky motion; 3, fairly continuous motion; and 4, excellent continuous motion, as in the movies. They also were specifically told not to base their judgments on the plausibility or likelihood of the motion, but instead on the "quality of movement across space." Subjects viewed photographs tachistoscopically, as in Experiment 1. Each trial began with a 2,000-msec inspection period, in which the subject inspected the first photograph, followed by a 1,000-msec blank interval. Then the same photograph was shown to the subject again, this time for 1,000 msec, followed first by a variable ISI, and next by the second photograph for 250 msec. At this point subjects gave their rating verbally. The initial inspection period was used because it was expected that the predicted effect would be stronger if subjects fully comprehended the picture. In pilot studies, it had been found that no apparent
motion could be seen unless the stimulus duration was quite long; hence, the 1,000-msec duration before the critical ISI. (No apparent motion was reported in Experiment 1 in which there was a 250-msec stimulus presentation.)

Each pair was shown under five different ISI conditions: 90, 110, 130, 150, and 170 msec. These times were selected on the basis of pilot data, because they seemed to fall in the range of optimal apparent motion ISIs for the photographic stimuli. The subjects saw each of the 10 pairs with each of the five ISIs twice, once in "forward," or real-world, temporal order (before-after) and once in "backward," or reverse, order (after-before), generating 100 total trials. The order of the trials was randomized for each subject. The experimental session lasted between 1½ and 2 h.

Results and Discussion

Mean goodness ratings were calculated for backward and forward motion for each of the ISIs. As Figure 2 shows, forward motion was judged to be better than backward motion for each of the five ISIs. This difference was tested statistically with a two-way analysis of variance, in which direction of motion was one factor and length of ISI was the other. There was a significant main effect for direction of motion as predicted: forward motion was rated as better (2.54 with a standard deviation of .33 on the scale on which 4 is ideal motion) than backward motion (2.20 with a standard deviation of .42 on the same scale) [F(1,7) = 11.75, p = .011]. There was no main effect for ISI (F < 1) or for an interaction between direction of motion and ISI (F < 1).

The finding that forward apparent motion is judged to be better than backward apparent motion is quite robust: Seven of eight individual subjects showed the effect, and the eighth simply showed no difference in preference between directions. This predicted effect was also found to hold for 9 of the 10 photograph pairs, and it represents a statistically significant difference when tested in an items analysis (t(9) = 3.43; p < .01). Also, as mentioned, the difference was found for all five ISIs sampled.

One possible problem with the present experiment is that subjects might have been affected by the plausibility of the motion even though they were specifically instructed to make judgments on quality of motion, independently of plausibility. The subjects claimed that they did follow these instructions, and, in fact, most subjects gave some ratings of 4 for some backward trials. Note also that such an explanation cannot account for the results of Experiment 1, in which subjects were asked to hold a static image in memory. To completely rule out the possibility that the subjects were unconsciously affected by the plausibility of the motion in Experiment 2, future experimentation using this sort of paradigm might look for cases in which the optimal ISI for backward motion is different from the optimal ISI for forward motion. That is, if, as hypothesized here, subjects look at the first picture and initiate a mental representation that continues the implicit motion, then the representational "distance" that the backward trials cover should be greater than the "distance" that the forward trials cover. This should mean that the optimal ISI for backward motion will be longer than the optimal ISI for forward motion, because the backward trials have more distance to cover. If this effect were to be found, it would rule out the plausibility explanation for the asymmetry between the goodness of forward and backward apparent motion.

Future research might also investigate whether there is an interaction between the real-world time separation between picture pairs, and goodness of apparent motion given ISI. This would be interesting if it turned out that there was a good correspondence between the real-world temporal intervals and the ISIs for the best apparent motion. This cannot be determined with the present study because the real-world time separations are confounded with both type of motion depicted and the quality of the photographs.

**GENERAL DISCUSSION**

The results of Experiments 1 and 2 support the hypothesis that people form a mental representation of movement when viewing frozen-action photographs. Perhaps the most significant question that emerges from these results concerns the source of movement information in a static stimulus. For instance, with these photographs, the implicit motion information could be at a totally semantic level; that is, subjects understand the meaning of the picture as a whole and deduce the motion involved.
At the other extreme, the motion information could be at a featural level, such that the correct motion can be deduced without any semantic understanding.

Friedman and Stevenson (1980) discuss the possible ways implicit motion can be represented in pictures. Although they do not consider issues of mental representation, they do review evidence that people can, at least, easily state whether a picture implies motion and what that motion is. They also make an introspective comment that is compatible with the hypothesis presented here: "we feel that pictorial movement is compelling, that, over and above its informational value, it can also transmit a sense of movement" (p. 226). Friedman and Stevenson surveyed pictures (paintings, photographs, caricatures, cartoons, and diagrams) to learn about pictorial movement. They came up with a number of ways movement can be conveyed, from paths such as ski tracks in fresh snow to the superposition of multiple viewpoints of the same object at different points of time. They also point out that there is a great deal of redundancy of movement cues within any one picture. It is interesting to note that there are effective "metaphorical" movement indicators, as seen in cartoons, but, according to Friedman and Stevenson's review of developmental literature, such indicators are effective only for older children and adults. Thus, it seems that there are a number of different levels at which movement can be expressed statically. Further research should discover which of the indicators can induce dynamic mental representations.

Another issue that emerges from the present research concerns the notion of "dynamic mental representations": What properties of the mental representation are dynamic? One possibility is that dynamic information is represented very abstractly. Such an idea is embodied in Shepard's (1981) suggestion that the representation of possible transformations an object could undergo is a static indication of the space of possible positions or rigid perturbations of an object. Note, however, that the representation of possible transformations could be more concretely dynamic, such that the representation of an object is one in which the object, at some level, is being continually moved. In that way, possible transformations are being dynamically explored. In general, dynamic mental representations might be rather analogous to real-world processes. This possibility is related to Shepard's (1981) proposal that the human mind has internalized certain transformations or processes with some sort of isomorphism with external transformations. Shepard talks about such transformations as processes performed on static representations, whereas I would like to propose that the representations themselves are dynamic (these issues are discussed further in Freyd, Note 1).

By analogy to real-world dynamics, one might look for "representational momentum." That is, if a person has reason to represent dynamic information, the representation might continue to change after the initial force to drive it is removed, much as a car continues to roll forward even if the fuel is cut off. In a recent study, Freyd and Finke (Note 2) induced a dynamic mental representation by showing subjects a static figure at a number of orientations along a possible path of rotation, each orientation being separated temporally. Even with ISIs of 500 msec, subjects reported feeling that the figure was rotating but seeing absolutely no visible motion between orientations. More importantly, subjects' mental representations of the figure seemed to have had "momentum." The subjects were instructed to remember the third orientation they saw, and were then presented with a fourth orientation which was either the same as the third or different. The subjects found it much harder to detect differences when the fourth orientation was a small rotation in the direction of implicit motion than when it was in the opposite direction, suggesting that their memory for the third orientation was distorted in the direction of the motion because of "representational momentum."

The Freyd and Finke results also address one issue that critics of the present research might raise: There might be a possible confounding of temporal ordering and physical changes. According to physical theory, the world is presently moving toward greater entropy as time passes. Thus, time has an implicit directionality. Assuming that this is true, one might argue that before and after shots are consistently physically asymmetric. The Freyd and Finke experiments, however, used a reversible transformation (rotation) and a temporally dependent asymmetry was still found.

Yet another set of issues is whether there is a consistent relationship between time and distance in dynamic mental representations. That is, are dynamic mental representations governed by certain physical laws of motion, either those analogous to real-world constraints or perhaps ones specific for mental representations? For instance, does it matter to our representation of the frozen-action photographs how fast the processes they capture happen in the real-world? This question is related to an issue raised earlier here, and to topics discussed in Shepard (1981): Are physical laws, such as conservation of momentum and gravitational acceleration, internalized in such a way that our mental representations are constrained by them?

In conclusion, the present results support the hypothesis that people represent motion when viewing static stimuli. This, in turn, supports the claim that perception and representation of motion play an important organizing role in the mind. These
results also suggest a number of interesting questions about the way implicit motion information is perceived and about the nature of mental representations.

REFERENCE NOTES


REFERENCES


Notes

1. The short retention interval was chosen because I expected that the predicted effect would be most observable if measured immediately after a person had looked at a photograph, and pilot data confirmed this.
2. In the data analysis, I used a reaction time cutoff point of 2,000 msec. For 10 of the 14 subjects analyzed, there were no reaction times that high, for two subjects there was one, and for two subjects there were two. Two subjects were excluded because they had more than five reaction times over 2,000 msec, which indicated that they were not doing the task as instructed.

(Manuscript received November 12, 1982; revision accepted for publication February 18, 1983.)