From: INTERSENSORY PERCEPTION AND SENSORY INTEGRATION
Edited by Richard D. Walk and Herbert L. Pick, Jr.
(Plenum Publishing Corporation 1981)

8

Gait Perception as an Example of How We May Perceive Events

JAMES E. CUTTING and DENNIS R. PROFFITT

1. Time, Movement, and Their Place in Event Structure

In *An Essay Concerning Human Understanding,* John Locke (Locke, 1690/1959, II:8:9) listed five primary qualities of objects in the world around us: solidity, extension, figure, number, and motion. These attributes have served us well in the study of perception. Solidity, for example, has been studied in terms of transparency and depth in vision, texture in touch, and stereophony in audition—stereophonic literally means "solid sound." Extension, or size, has been very important, particularly in vision, where psychologists have studied effects of relative and absolute size, and proximal and distal size. Figure, interpreted as shape or form, has undoubtedly been the most important quality of objects for all perception, providing impetus for Gestalt and other psychologists. Number, especially when logically extended to include composition and the relation of parts to wholes, has also played an important role. Motion has also played some part; however, unlike the others, motion has most often been excised from the aggregate by the perceptual psychologist. It appears that we have spent most of our efforts in the study of *static* figures of a certain size, composition, and apparent solidity. It is clear that motion is a quality unlike the others; it

JAMES E. CUTTING • Department of Psychology, Uris Hall, Cornell University, Ithaca, New York 14853. DENNIS R. PROFFITT • Department of Psychology, University of Virginia, Charlottesville, Virginia 22901. Supported by NIH Grants MH33087 and MH35530.

invokes time. Our general purpose is to help integrate motion and time back into the study of perception.

Gibson (1950, 1966, 1979) and Johansson (1950, 1973, 1975) have argued that we should have spent more time with movement and less with static perception. Neisser (1976) and Turvey (1977) have lent the force of their recent apostasy from the usual tenets of information processing to reinforce the general views of Gibson and Johansson. It is amusing to note, however, that while these researchers have inveighed against the use of the tachistoscope to reveal psychological processes, the *inventor* of the modern tachistoscope, Raymond Dodge, presented this same view more than 60 years earlier:

> The tendency to reduce the physical exposure time to a minimum . . . is a methodological mistake, based upon a psycho-physical misconception. It introduces unusual conditions altogether foreign to the natural fixation pause, and leads, or may well lead, to a distorted analysis of the processes of apprehension; making the conclusion insofar as they are referred to normal perception, not merely valueless, but false. . . . (p. 32) Anything approximating a threshold exposure instead of simplifying the consequent psychological process, really complicates it and renders it more uncertain. (1907, p. 35)

Alas, fascination with the possibility of taking snapshots of mental life has left psychologists unmoved by Dodge's admonition. The reason for this may be twofold: First, psychologists have proceeded for the most part as if motion perception is achieved by simple compilation of static forms, much as a moving picture is assembled from separate frames; second, when somewhat more reflective, psychologists have simply stated that motion is too complex to be studied until we have a better grasp of the laws of form, composition, solidity, and size.

The first reason views dynamic forms as special cases of static forms. However, Gibson and Johansson suggest reversing this relation, and we agree. The value of considering static events as special cases of dynamic events is evidenced in the history of analytic geometry and calculus, mechanics, and physics. Johansson (1974, p. 142) has noted that, "The initial theoretical structure for physics was of a static type, and Zeno's paradoxes are well-known examples of the logical impossibility of dealing adequately with motion within this framework." He goes on to suggest that conventional approaches to perception are "rather similar to the pre-Galilean type of physics." Indeed, the conceptual breakthroughs of Galileo and Descartes were grounded in considerations of dynamic events, whereas the speculations of their predecessors related almost exclusively to statics (see Boyer, 1968; Mach, 1893). Thus, perhaps we should follow the lead of physics and consider dynamics to be primary with statics as a special case where time is held null.

The second reason for omitting motion from psychological study is equally damning. It asserts that psychologists are simply incapable at this time of a dynamic perceptual psychology. We think this is untrue. In our opinion, the basic elements necessary for a psychology of event perception are at hand.

Both reasons, however, converge on the same fact—the motion of relatively complex objects has remained largely unstudied. Our purpose in this chapter is to demonstrate that these two rationales are incorrect: (a) Movement is not the simple compilation of static forms, and (b) it is not so complex as to be unamenable to analytic study. In fact, we hope to show that there is a beautiful simplicity in dynamic forms that may contribute to their perception.

Our domain of study is that of *event* perception, so named by Johansson (1950). Before beginning our exposition, however, it behooves us to explain what we mean by an event. It seems best to follow Webster, where an event is defined as: "That which occupies a restricted portion of four-dimensional space-time." By extension, *perceptual* event would entail perception of movement or recognition of change in an object in one, two, or three spatial coordinates through the added dimension of time. What this definition accomplishes for us is that it states that events have a structure. Moreover, it implies that this structure separates an event from the rest of the world.

Wolfgang Köhler, among others, has parsed event structure. In speaking of objects and their movements, he stated

> In a physical system events are determined by two sorts of factors. In the first class belong the forces and other factors inherent in the processes of the system. These we will call the *dynamic* determinants of its fate. In the second class we have characteristics of the system which subject its processes to restricting conditions. Such determinants we will call *topographical* determinants of its fate. (1947, p. 107)

A *dynamic invariant* of an event can be thought of as some algorithm that describes change throughout the course of the event, and governs its temporal structure. A *topographic invariant* can be thought of as the abstract physical relations that remain constant throughout the event, governing its spatial structure. For a rolling ball, rolling can be said to be the dynamic invariant (subject to forces of gravity and impetus), and a spherical shape is the topographic invariant (subject to deformations by dynamics). Pittenger and Shaw (1975) have parsed events in a similar manner, postulating transformational and structural invariants. We prefer Köhler's terminology, because it is historically older and because the term *structural* as it includes dynamics, invites confusion of purely topographic concerns with all of event structure.

There are problems, however, with the type of parsing suggested by Köher and by Pittenger and Shaw. One problem is that this distinction systematically confuses two types of movements, those occurring within the object and those occurring with respect to the observer. Wallach (1965) calls these object-relative displacement and angular displacement. For our rolling ball, for example, the action of rolling has two components. One is object-relative. That is, the ball turns circles around itself. The other is observer-relative. That is, the ball moves through a visual angle of measurable degree as it passes in front of the observer. The differentiation of these two motions, rotation and translation, is important because, for event perception to occur, it appears that object-relative displacement is perceived *before* angular displacement, as suggested by Wallach (1965), and contrary to Johansson (1973). Proffitt, Cutting, and Stier (1979) and Proffitt and Cutting (1979) give a detailed account of why this is so. In the final two sections of this chapter we will discuss these two types of dynamic motions in more detail.

What is new to our approach to perceptual events is a factor that derives from topographic and from object-relative dynamic invariants: the center of moment. The center of moment is a functional point within the topography of the object, around which object-relative dynamics occur. For example, an axle is the center of moment for a wheel, and a fulcrum is the center of moment for a lever arm. Our plan in this chapter is to demonstrate that this point is perceptually useful and that, in fact, it appears to guide the viewer to perceive a dynamic whole. In this manner we capitalize on two of the term's meanings: that is, on *moment* as it refers to movement in the study of mechanics, and also as it refers to importance (as in the term *momentous*).

2. Gait Perception without Familiarity Cues

Using a technique that began with the work of Marey (1895/1933), Johansson (1973) mounted points of light on a person and observed that a perceiver had no difficulty determining the activity of that person, whether walking, running, bicycling, painting, doing push-ups, or dancing. We became fascinated with this kind of dynamic array. By placing lights on the joints and by darkening the surround, an experimenter can effectively remove all familiarity cues from the stimulus. In other words, the type of clothing, hairstyle, and facial expression of an individual are omitted, and only the relations among the moving points of light can be seen. Thus, one perceives the presence of a human being simply through the relational structure of the moving lights. This relational structure is, we believe, a deep structure, analogous to that of a sentence. In the final section of this chapter we will outline what we mean by this.

Our starting point in this research program was the commonly held belief that people can recognize friends by their walk. Unfortunately, this belief and the previously existing research on the topic (Wolff, 1943) are confounded by surface cues and probabilities of seeing someone at a given place and time. To counteract these contradictions, we employed six Wesleyan University undergraduates, three males and three females of approximately the same height and weight who lived together in university housing (Cutting & Kozlowski, 1977). We affixed glass-bead retroreflectant bicycle tape to their joints and videorecorded their gait as they passed in front of a television camera. A bright light was mounted near the camera and focused on the walking area. Contrast was turned to near maximum, brightness to near minimum, and the camera was placed slightly out of focus so as to make more circular the images of the reflectant patches. Figure 1 shows static representations of this technique, with brightness turned up so that one can see where the tape is mounted on the walker. Two months after recording many examples of the gait of each walker, we randomized the dynamic-stimulus items on a test tape and presented the test to the subject. Their task was to identify each person in a 2.7-sec dynamic display of lights. Chance performance was near 17% and all viewers exceeded this value, with a mean of 38% correct identification across all trials. Without feedback, their performance improved over the course of the task from 28% on the first third of all trials to 59% on the last third. Thus, viewers can recognize themselves and others from a display that presents only the time-varying relational structure of gait, without the usual cues of hair style, clothing, and facial expression.

Although we continue to be fascinated with this result, we have not found it particularly amenable to further study. The problem lies in the fact that individual identity can be thought of as composed of many aspects: height, weight, age (all of which were controlled for in our study), gender, and various difficult-to-define personality variables. To further our study of gait perception, then, we chose to focus on one particular aspect of different walkers. We selected gender in part because Maas, Johansson, Jansson, and Runeson (1971) asked the following question on a study guide: "Do you think that subjects could detect male-female differences in the motion pattern of walkers, using the 'pinpoint of light' technique?" Aggravatingly, they did not answer the question, so it fell to us to do so experimentally (Kozlowski & Cutting, 1977).

We took the same videotape sequence used previously and presented it to 30 different undergraduates who had no knowledge of the identity of the six walkers. Where chance for determining gender would be 50%, performance was 63% correct, significantly better than chance. Moreover, when we replicated the results with 14 new walkers (seven males and seven females) and 57 new viewers, performance in judging males

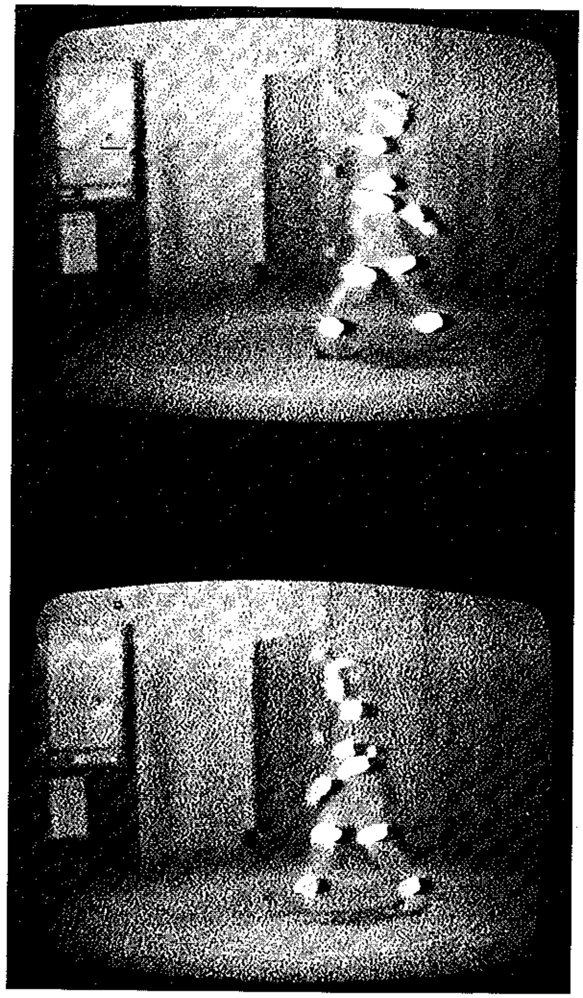


Fig. 1. A walker in two most-outstretched positions in the step cycle. Brightness of the image is turned up to reveal the approximate locations of the reflectant tape. Blurring and fuzziness result from the nature of the video medium.

and females was 67% correct (Barclay, Cutting, & Kozlowski, 1978). In general, across many experimental manipulations, observers, and stimuli, performance appears to hover between 60% and 70% on this task. Thus, while the effect is not large, it is systematic and generalizable.

Among these results were a number of interesting and surprising findings. First, static displays taken from the dynamic sequences were insufficient for revealing gender cues. In fact, most subjects had no idea that the constellations of lights represented human walkers—instead, they saw them as either clusters of stars or as Christmas trees viewed at night. Second, at least 2 sec of display time in the dynamic stimulus was necessary before viewers could give any accurate judgments of gender. This is an order of magnitude longer than the duration needed to see that the moving lights represent a human being (Johansson, 1976). From this fact we argued that two-step cycles, or four complete steps, appeared necessary for gender recognition. Third, no particular lights seemed sufficient for gender identification. It may be that even the ankles alone are sufficient (but see Kozlowski & Cutting, 1978), and certainly the ankles and right hip alone or the wrists, elbows, and right shoulder are sufficient.

These results suggested to us a number of things. First, the dynamic aspect of the display is crucial. Second, it took an extraordinarily long time for viewers to perform the task—at least with respect to the usual displays used in information-processing tasks. This generally excluded reaction time as a dependent measure and the use of brief displays as stimuli if we were to pursue our interest in this domain. Finally, and most importantly, the information available to the viewer appeared to be everywhere in the display, distributed from head to toe in the dynamic stimulus. This precluded the usual divide-and-conquer approach to cue finding in the stimulus array. Somehow the information that was relevant to gender identification was holistic in nature.

Initially, however, we ignored these implications and took a particularistic approach. Our first hypothesis was that arm swing might dictate to the viewer who is male and who is female. Indeed, we found that in our sample and in the normal population in general, women swung their arms more than men. Thus, we brought back our original set of walkers, and controlled arm swing. We found that viewers could recognize gender regardless of arm swing, though to a diminished extent. Thus, arm swing may be sufficient, but it is not necessary for gender recognition. Next, we thought that walking speed might suggest to the viewer who is male and who is female. Indeed, in our sample and in the general college population women take more steps per minute than men. We brought back our original walkers again, controlled walking speed, and found that viewers could still recognize gender, though again to a somewhat lesser degree. We also investigated step size—males generally take bigger steps than females—but this, too, was not a necessary cue. Thus, with such a particularistic approach we seemed to be thwarted at every turn. This, more than the logic of our previous results, forced us to drop back and reconsider the stimulus as a whole (Cutting, Proffitt, & Kozlowski, 1978).

2.1. A Search for a Biomechanical Invariant

We were convinced that there was something in the dynamic display that viewers were using to make their judgments. Thus, with no particular prejudice as to what we might find, we took a whole battery of physical measurements from our walkers—shoulder width, hip width, torso length, height, leg length, arm length, etc. Juggling these figures in virtually all possible ways, we rediscovered a fact that has been known since the beginning of art: Men have proportionately wider shoulders than women, and women have proportionately wider hips. This type of data can be found in Gray's Anatomy (1901/1977), Albrecht Dürer's sketches (Dürer, 1528/1972), and in contemporary sources such as Faust (1977, p. 67). The standard measure that reflects these differences is a ratio of shoulder width to hip width. Depending on where the measurements are taken, the ratio for men in our sample and for men in general is about 1.1, for women in our sample and women in general, about 1.0. This ratio for each of our walkers was highly correlated with a second measure, the percentage of all trials in which each was identified as male. That is, in general female walkers in our sample had shoulder–hip ratios near 1.0 or below and were identified as male on a mean of 35% of all trials. Male walkers, on the other hand, had ratios near 1.1 and were identified as male on a mean of 65% of all trials. This correlation across all walkers was remarkably strong, $r = .84$.

We thought that we might be well on our way toward describing a biomechanical invariant for gait perception, but then realized that there was a basic problem with our description: Viewers made judgments of walkers' gender from *sagittal* displays, where only one shoulder and one hip were visible at any one time. Thus, while our description accounted for most of the variance in our response measure, we were embarrassed by the fact that this description could not possibly be used, because it was not directly calculable from the visual display. We did feel, however, that we might be on the right track and that something related to the shoulders and hips may be invariant in sagittal, and perhaps other, displays.

Our second tack, then, was to determine what the shoulders and hips were doing in the step cycle. Studying the work of Carlsöö (1972), Murray (1967), and Saunders, Inman, and Eberhart (1953), we realized that the hips and the shoulders work in opposition to one another, as shown in Figure 2. *Torso* and *torsion,* after all, are derived from the same word. Thus, at the point within the step cycle where the right leg is forward and both feet are on the ground, the right hip is forward and the right shoulder is back. Both joints move counterclockwise along the path of an ellipse so that the next time both feet are on the ground, with

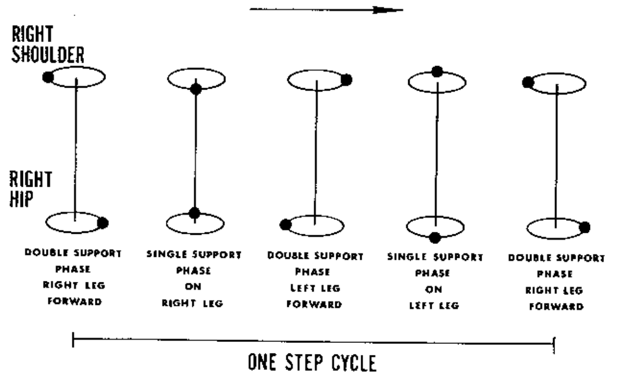


Fig. 2. A schematic representation of ellipsoidal movements of the right shoulder and right hip during one step cycle.

weight equally distributed, the right shoulder is forward and the right hip is back. Notice also that when the person is standing solely on the right foot, the length of the torso as measured from the hip to the shoulder on the right side is at its shortest extent, and that when standing only on the left, this length is at its greatest. This latter motion derives from a description by Murray, Kory, and Sepic (1970, p. 647). However, we concentrated on the former—the oscillation back and forth of the hips and shoulders. We argued that if male shoulders are broader than female shoulders, this difference ought to be visible as a greater shoulder swing for males than females. Moreover, if the shoulder difference is greater than the inverse difference in hip widths—and it is—then one could derive a measure of the various hip and shoulder swings from the visual display. This measure is shown in Figure 3.

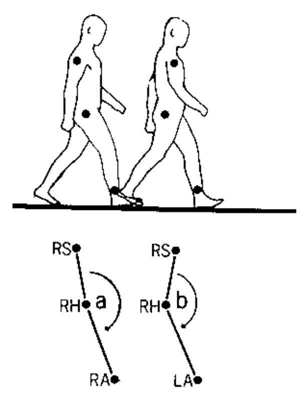


Fig. 3. Viewed from the side (as a sagittal projection), the torque in the torso can be measured by subtracting Angle B from Angle A. RS = right shoulder; RH = right hip; RA = right ankle; LA = left ankle.

Consider just three points of light: those for the right shoulder and the right hip, and that for the forward leg as it is planted on the ground. Connecting these points of light one can derive two angles, one when the right leg is forward and the angle is largest and one when the left leg is forward and the angle is smallest. The difference between these angles, we argued, would reflect the differences in the amount of shoulder swing, and hence might reflect gender. More particularly, male walkers should have larger angle differences than females. Indeed, careful measurement of all stimuli for all walkers yielded this difference. The correlation between this index and the percent of all trials in which each walker was identified as male was high, $r = .76$. This correlation is slightly lower than that for torso structure and identifiability, and it is as high as it is only because one walker was omitted from the sample. She was omitted because she did not swing her arms at all, had no difference between her two angular measurements, and was identified as male nearly 50% of the time. In other words, we argued, she was off the scale of angular differences—females had a mean of about 9° and males about 18°, and she was near 0°.

We were not happy with the need for her exclusion, but we were even more unhappy with our torsion index and its relation to a previous result. Remember, not all joints need to be represented in the dynamic display for accurate gender identification to occur. For example, the arms (shoulders, elbows, and wrists) are sufficient by themselves—leaving no hip and no ankles to be measured against. In addition, the legs (the hips, knees, and ankles) are also sufficient—leaving no shoulder to be measured against. Thus, while our torsion index may be appropriate for some standard displays in which all joints are represented by lights, it was clearly insufficient for others in which gender was determined almost as well as in the standard condition.

Although we believed that we were still on the right track, our description was surely not the most apt. We had come to believe that our biomechanical invariant should be one that was appropriate to all experimental conditions. Moreover, we began to be wary of the fact that, whereas our previous description might be appropriate to walkers, it was completely irrelevant to other types of moving displays. Thus, we wanted our invariant not only to capture the essence of the dynamics of gait, but to be applicable to other perceptual events.

2.2. An Invariant Found: A Center of Moment

Further study of existing research on the dynamics of gait revealed the systematic patterns involved. As one arm swings forward, the other swings back; as one leg swings forward the other swings back; crossed limbs work in phase synchrony; ipsilateral limbs work in opposition. Thus, dynamic symmetry is the key. It occurred to us, then, that if gait was symmetrical we should be able to find planes of symmetry. Certainly the midsagittal plane, the plane of biological symmetry, was important. However, we found two others: a horizontal plane of symmetry between the hips and the shoulders near the level of the navel, and a frontal plane dividing the front of the body from the back. These three planes intersect at a point, somewhere between the shoulders and the hips, so we began to focus on trying to find it.

Dividing the dynamic walker into planes is no easy feat, so it occurred to us that perhaps the environmental coordinates of the space around the walker may not be the most appropriate. We decided to consider the coordinates of the walker instead. This view assumes simply that the walker stays stationary and the environment moves around him or her. Functionally, it is identical to a description that assumes that the walker moves through the environment, but it turns out to have some notational simplicities about it, not captured by the other. Besides, Johansson (1973, 1976) had begun to use environmental coordinates in his vector-analytic approach to gait perception and, to us, it looked hopelessly cumbersome.

By using walker coordinates we assumed that there is a point within the walker around which everything moves. Indeed, this point will move through the environment, but this is an irrelevant concern at this time; we are simply interested in finding a "ground zero" for the movement of gait. It turns out that since the arms and legs are working in dynamic symmetry they can be temporarily omitted from consideration. That is, whatever movements happen to be going on in an arm are being counteracted by the other arm and the ipsilateral leg, but reinforced by the opposite leg. Forces and movements in the limbs nullify one another, leaving only the torso to be considered.

The torso has the general shape of an isosceles trapezoid, oscillating as a flat torsion spring. This general shape is shown in Figure 4 for a male walker and a female walker. When the torso oscillates, one can

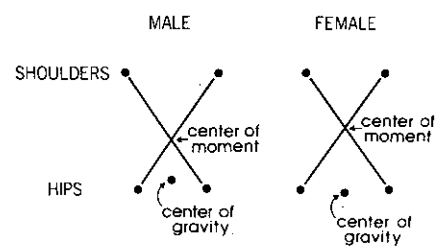


Fig. 4. Schematic representation of the torsos of a male and female. In general, males have slightly wider shoulders and narrower hips than females. The intersection of stress lines across the diagonals of the torso is the center of moment. Note that this point is not the center of gravity.

derive stress lines across the diagonals of the torso. These intersect at a point that we call the *center of moment,* a reference around which all movement in all parts of the body has regular geometric relations. Its relative location can be determined by knowing only the relative widths of the hips and shoulders. In general, males' center of moment is lower than females'. It is important to note, however, that for walkers the center of moment is *not* the center of gravity. The center of gravity concerns distributions of mass, whereas the center of moment concerns distributions of movement.

To make this point clear, consider a rolling wheel. The center of gravity for the wheel happens also to be its center of moment. That is, all motion occurs around the point around which all mass is distributed. However, consider these centers for a pendulum. Here, the center of gravity lies in the bob, oscillating back and forth, whereas the center of moment is the pivot. In similar fashion, the centers are separated in the human body. Certainly the center of gravity plays a crucial role in maintaining the balance in a human body, but we suggest that it is the center of moment that is perceptually the more important, because its location within the torso is strongly correlated with viewer performance, $r = .86$. Happily, this correlation includes the one difficult-to-identify female walker that was previously excluded.

However, a note of caution is in order. We acknowledge that there are an indefinite number of possible mathematical descriptions for our stimulus events, each of which would be similar to this one and account for the data just as well. Thus, there are at least two avenues of logic that we must pursue. First, we must garner some empirical evidence showing that our theory is plausible, and second, we must discover if our description applies to other types of perceptual events besides walking.

2.3. Gait Synthesis

We had found that fully three-fourths of all the variance in viewer judgments could be accounted for by a single variable. One should always be skeptical of such results, and we were. Thus, it behooved us to determine if this claim could be substantiated. The method that occurred to us was gait synthesis—an attempt to generate, on a computer display scope, constellations of dynamic light patterns that mimicked the movements of male and female walkers. Through this synthesis it would be possible to generate stimuli that differed *only* in their centers of moment. That is, the motions involved would differ only because of the difference in the place around which they were generated, constrained by the general forces of gravity and impetus.

Gait synthesis turned out to be relatively easy to accomplish (Cutting, 1978a,c), and it became possible to generate synthetic male and female walkers that were at least as identifiable as such on 80% of all trials. In addition, gait synthesis allowed us to replicate some of our previous results. For example, removing some of the lights from the display, such as the shoulder and the hip, decreased viewer performance to about 60%, yet it remained significantly above chance. Moreover, viewers judged that the synthetic versions of female walkers were more light-footed than the synthetic males. This result correlates nicely with naturalistic observations, and it makes good biomechanical sense. Females, with larger excursions of the hip, absorb more of the vertical bouncing motion of gait. This causes the carriage of the torso to be more even, and creates the appearance of gliding across the floor.

We were most pleased with our gait-synthesis venture by the fact that the synthesis routine is *exactly* that structure derived from our analysis of gait. This analysis-and-synthesis type of research program is analogous to that used so successfully in speech perception (Liberman, Ingemann, Lisker, Delattre, & Cooper, 1959). Its success can be taken as corroborating evidence that our description of the center of moment is theoretically plausible.

We are left, however, with the problem of generalizability. That is, would our center-of-moment analysis work for the perception of other types of events—for moving objects without shoulders and hips?

3. Perceiving Centers of Moment in Other Events

We do not have space to discuss fully our research on nongait movement, but we can indicate the perceptual utility of centers of moment in four domains, each quite different from walking.

3.1. Perceiving Centers of Moment in Wheel-Generated Motions

Consider first the Gestalt demonstration of the perception of rotary motion, or wheel-generated movement. From reading Koffka (1935) most of us remember that placing a light on the rim of a wheel, darkening the surround, and rolling that wheel across a flat surface yields the perception of a hopping light—but *not* the perception of a wheel rolling. Add some lights to the perimeter, or one to the center, and suddenly the dynamic event looks like a rolling wheel. Although the Gestalt psychologists were convincing in such demonstrations, we found little hard data on how good these rolling configurations of lights appeared to viewers. Thus, using a scaling technique, we investigated the apparent goodness for configurations of lights mounted on a rolling circular structure (Proffitt & Cutting, 1979; Proffitt *et al.*, 1979).

Two abstract centers are of importance here: one is the center of the wheel, and the other is the center of the configuration of lights, or centroid, which, in our view, is directly analogous to the center of moment in a walker. For a one-light configuration the centroid is simply the center of the single light; for a two-light system the centroid is the center of the triangle formed by the three lights, derived by the method of medians. This point is the point on which the triangle would balance if it were placed on top of a pencil tip. We discovered that it is the relation between these two points, the centroid and the center of the generating wheel, that dictates how wheellike the motion of lights will appear. That is, if the two centers are very far apart, the system of lights will appear to be a hobbling or bouncing structure. If, on the other hand, the two systems are very close together, the structure will look more like a rolling wheel. When the two centers are in exactly the same locale, viewers typically judge the stimuli to be perfect, or near-perfect, wheels. Thus, it is the center of moment (centroid) of a configuration of lights, and its relation to the center of the generating wheel, that perceivers appear to use in making their judgments of goodness concerning rotary motion.

3.2. Perceiving Centers of Moment in Aging Faces

A second nongait event in which we have found a center of moment useful is the perception of aging in human profiles. Pittenger and Shaw (1975) describe the process of facial aging as one of cardioidal transformation. That is, from the work of Thompson (1917) and Enlow (1975), they derived a mathematical description of the elastic shape changes undergone by the human skull. These changes are rather closely approximated by a heartlike (cardioidal) change in which the proportionately larger size of the cranium in the child is supplanted by a proportionately larger size of the lower face in the adult. Fortuitously, this transformation must occur about a point, and the most likely place for this point is somewhere inside the skull. Transforming the profile of an early teenager around a matrix of different points in sagittal space, and having viewers make judgments about the goodness of different transformed heads, we derived a point of best transformation (Cutting, 1978b). That is, this point would yield the transformed profiles judged as best examples of younger and older versions of a standard profile. Interestingly, this point corresponds almost exactly with a region found by Enlow from which aging changes are generated. This region is near the foramen magnum, the place where the skull meets the spinal column. Thus, for the perception of transposed heads—with apologies to Thomas Mann—there is a center of moment around which aging best occurs.

3.3. Perceiving Centers of Moment in Flowfields

A third nongait event of interest concerns the flowfield demonstrations of Gibson (1950, 1979). When flying over a terrain, or driving a car through it, one becomes aware that all textures in the visual world around the observer expand centrifugally from a point toward which one is hurtling. The perceptual utility of this point is that its location marks precisely the direction in which one is headed. (A complementary point can also be found directly behind the observer, indicating the direction from which he or she is coming.) Thus, the center of moment for a radiating flowfield can be useful to a perceiver, just as it is in the cases of a walker, a rolling wheel, and an aging face.

3.4. Perceiving Centers of Moment in the Night Sky

A fourth nongait event concerns celestial navigation by migratory songbirds. This entry in the domain of event perception pleases us because, for the first time in our discussion, it concerns points of light that are real objects (stars) rather than schematic representations of such objects. It appears that certain songbirds can use the clear night sky as an aid in migration. Prior to migratory flight, they typically sit in the branches of a tree for a period of hours, then suddenly take off—southward if it is fall and northward if it is spring. Ethologists had postulated that these birds might have some form of innate celestial map. Emlen (1975), however, thought this implausible and set out to discover what source of information in the night sky might aid migration. He suggested that birds watched the night sky rotate over a period of hours and through that rotation found the celestial north pole, Polaris (the North Star). To corroborate his point, he placed birds overnight in round cages within a planetarium, and found that over the course of hours they oriented with respect to the north pole. Moreover, when Emlen artificially rotated the night sky around Betelgeuse, the birds oriented anew as if it were the pole star. Our interest in this phenomenon is that the celestial north pole, whether there is a star there or not, is the center of moment for the rotating manifold of the night sky. Songbirds appear to extract it from the dynamic display, and use it to guide their migration.

Our delight in these disparate findings is that centers of moment appear to be useful in all types of events—those that are either slow or fast, rigid or nonrigid, reversible or irreversible, and biological or nonbiological. Moreover, underlying the richness of varied types of events there appears to be a common structure. This structure, we believe can be thought of as a deep structure, much like the deep structure of a sentence. In the next section we will outline what we mean by this.

4. Toward a Grammar for Perceptual Events

We are continually impressed in our research by the fact that the perceptual system seems to select one mathematical description for an event from among an indefinitely large set of possibilities. For example, two lights mounted opposite each other on the rim of an unseen rolling wheel are seen to revolve about the wheel's unseen center and also to move linearly across the visual field. This perception of *two* motion components for the two lights is one mathematical description of the event; however, the perception of each light moving independently on cycloidal paths is an alternative that is rarely seen. Depending upon the orders of extraction of information, the number of motion components sought, and the points taken as reference, this simple event could be specified in many different ways. In this section we offer a general description, a grammar if you will, for event perception. This "grammar" is an attempt to describe the information available in visual perception in terms of its analysis into the components that are perceived.

We call the information available to observers over time the visual scene.[1] Figure 5 is a deep-structure description of the parsing of the visual scene by the perceptual system. This diagram retains traditional descriptions of static perception by making them a special case of the more general perception of dynamic events. An emphasis is placed on the splitting of movements into two perceived components: internal dynamics within the figure and action of the figure as a whole. The remainder of this section is an elaboration of Figure 5, proceeding from the top of the diagram and working down.

The first division of the visual scene preserves the distinction between figure and ground while making figure a special type of event. Since the early work of Rubin (1915), perceptual psychologists have noted that objects, figures, and events are always embedded within contexts or grounds. Given that the term *figure* has static connotations, we prefer that the primary distinction be made between event and

[1] The visual scene presented to observers in our research to date is an admittedly restricted domain of moving dots of light presented on a two-dimensional TV screen; however, as there is ample evidence that depth can be perceived from similar experimental displays (see especially Green, 1961, and Braunstein, 1976), we take the limitations on the events studied to be less of a problem than might first be supposed.

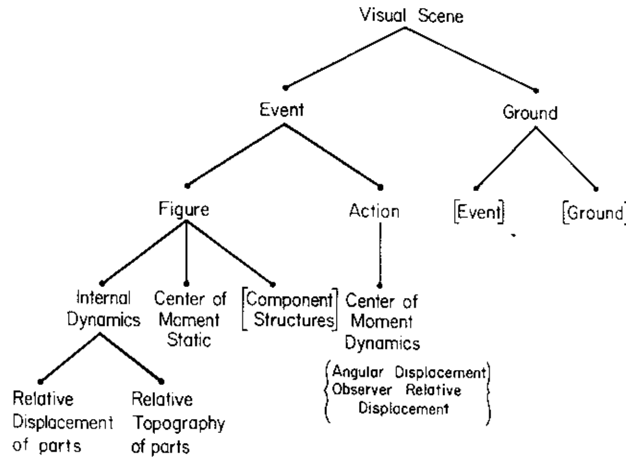


Fig. 5. A grammar for the perception of events, showing the manner in which we believe that the visual perception system parses events and extracts information.

ground. The information defining an event, let us say a walker, is extracted from the visual scene with residual information becoming ground. Within the ground further event–ground distinctions may be embedded. For example, a friend may be recognized walking on a crowded thoroughfare. The walking friend is an event embedded within the ground of the flowing pedestrian traffic. This ground would be rich in possibilities for extracting other events. If the event has no component of motion displacing it relative to the observer, then the event may be rewritten simply as figure. Thus, we define figure as an event without action.

By our account, an event is divided into two components, figure and action. A figure is a whole comprised of parts interrelated with static and dynamic invariants. For example, the dynamic configuration of lights that is recognized as a walker is a whole comprised of the lights which are its parts. The movement of the whole relative to the observer (that is, across the monitor screen) is its action.

Consider the pendulum drawn in Figure 6. As an event, a swinging pendulum consists of a bob, suspended below a pivot, that oscillates back and forth through an arc of angle θ. Considered as a whole this event has no observer-relative motion. That is, it may be adequately described as a figure with no action component. Remember, action is the dynamic whole that maintains a static relationship to a stationary observer. Thus, a pendulum is perceived to have motion, but its movements are not of the traveling sort.

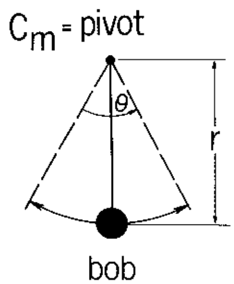


Fig. 6. The movements and structure of a simple pendulum.

We find it convenient, as shown in Figure 5, to consider a figure as comprised of three components: internal dynamics, center of moment, and component structures. For some simple figures, such as a swinging pendulum, there are no component structures; thus, we bracket this constituent to show that it is optional, just as other events are optional in the whole of the visual scene. A walker, on the other hand, is perceived as a nesting of *many* component structures, a description of which entails the elaboration of the two required constituents.

As we see it, the internal dynamics for any figure is comprised of two constituents, relative displacement and relative topography of parts. These constituents are, respectively, dynamic and static relational invariants of figures. Generally speaking, the internal dynamics of a swinging pendulum consists of the motion of the bob about the pivot and the pivot-to-bob distance. More specifically, the relative displacement of parts is the acceleration-motion vector of the bob as it moves through arc θ of radius r about the pivot. The relative topography of parts is r, the length of the arm of the pendulum. If the pendulum were not swinging, the perception of this static figure would be wholly described by noting the positions of the pivot and the bob and the distance between them. Thus, for the special case of static perception, the available figural information in the visual scene is fully defined by the relative topography of the figure's parts.

The second constituent of figure is the center of moment as a static relational invariant. It is that point in space about which the internal dynamics of the figure are specified. The center of moment is the reference point, analogous to the point of origin in a coordinate system, for the description of the motions and locations of the parts of a figure. Its means of determination has been discussed in previous sections for each domain—walkers, wheels, aging faces, flowfields, or stars—that we have examined. As suggested earlier, the center of moment of a pendulum is its pivot.

For dynamic figures, we find that some shapes, like an aging face, change over time. Others, such as a pendulum, generally do not. Our description of figures allows us to state precisely the conditions for the perception of changing shape: If the internal dynamics of a figure has a period of identity, then it can be specified by a group of symmetry operations about the center of moment. A pendulum, for example, has a period of identity every 2θ of swing. A walker, likewise, has an identity period every two-step cycle. The circular rotations of lights on a wheel have a period of identity every 360° of rotation. So, too, do the circular paths of the stars moving about Polaris. In all of these instances the dynamics of the parts do not change the shape of the whole. However, in those cases where the internal dynamics of a figure do not have a period of identity, the figure will be perceived to change shape. An aging face is one instance of such an event.

Just as we do not perceive as separate the individual motions of parts of a rolling wheel, so, too, we do not perceive as separate the individual motions of parts of the body when viewing a human walker— a moving whole is seen in both cases. Figure 7 portrays the vector paths for the head, shoulder, elbow, wrist, hip, knee, and ankle of a computer-

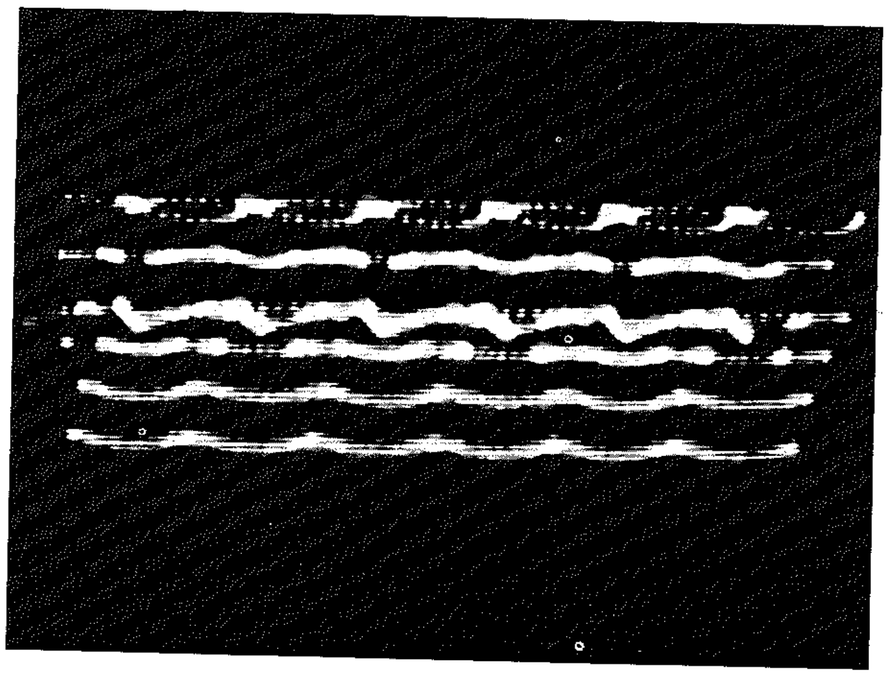


Fig. 7. Vector paths of a synthetic walker generated by the program of Cutting (1978c). The importance here is that we do *not* perceive the vector paths; instead, we perceive a walker. We do this, in part, by segregating relative movements of body parts from the action of the walker as a whole.

UPSIDE DOWN IMAGE

stimulated walker, generated from those used by Cutting (1978c). A similar set of traces may be found in Johansson (1973) for lights attached to a natural walker. We no more see this conglomerate of motion paths of a walker than we see the cycloidal paths of points on a rolling wheel. Event perception is clearly no such compilation of static frames. The grammar for event perception allows us to describe those motion components we do see in a walker by separating the available information into the required constituents and a nesting of component structures.

Figure 8 shows the analysis of the figural information for a walker presented as a dynamic display of lights attached to the joints. Because this figure is already rather cluttered, we have elected *not* to add the action component of the event, which would be the translatory motion of the whole figure relative to the observer. One may imagine the

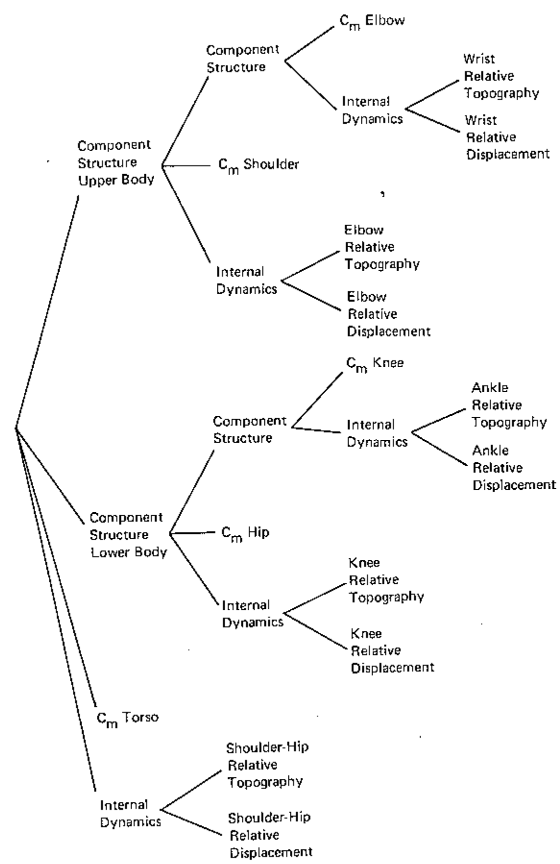
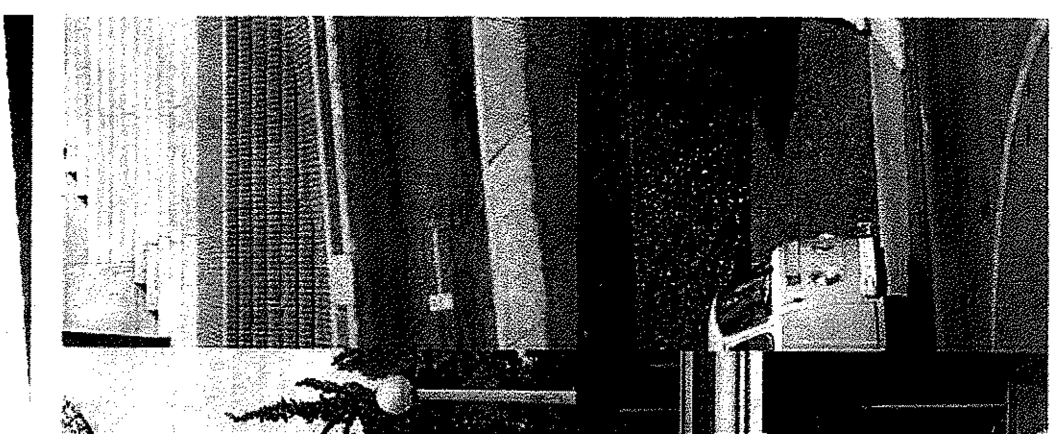


Fig. 8. A partial grammar for the perception of a walker. Only the elements under the heading *Figure* in Figure 5 are shown here.



description to be of a walker on a treadmill, sauntering along but getting nowhere. The motions of the lights attached to the walker are first analyzed into the constituents of the internal dynamics of the torso, the center of moment of the torso, and the component structures of the lower and upper body. The further analysis of these constituents yields a description of the relative motions and topography of the body parts as they are perceived.

The first extraction of information from the dynamic display describes the internal dynamics of the torso occurring about its center of moment. The torso is marked by two lights, one on the shoulder and the other on the hip. These two lights move, 180° out of phase, on elliptical paths of differing sizes. As discussed in Section 2.2, torso dynamics are described as the motions and locations of the hip and shoulder lights as they occur relative to the torso's center of moment. This center serves as a reference point for the figural description of the walker.

Once the perceiver extracts the movement from the hip and shoulders, these points now serve as static centers of moment for the analysis of the component structures of the lower and upper body. Considering first the lower body, one can see the hip as the center of moment for the knee, which moves in pendulum fashion about it. That motion shared by the hip and knee has been subtracted out; thus, the residual motion of the knee is described as an arc vector about the hip with a radius equal to the length of the upper leg. We suggest, as Figure 8 shows, that the internal dynamics of the knee is the relative displacement and location of the knee with respect to the hip. One perceives in the knee only its motion relative to the hip, rather than the more complex motion resulting from compounding this pendulum motion with the elliptical motion imparted to the knee by the hip. With the knee-to-hip relations described, only the component structure of ankle-to-knee remains. That is, the motion of the knee has been subtracted out, thus becoming a static center of moment for the description of the internal dynamics of the ankle. The ankle is related to the knee, roughly as a pendulum bob to its pivot. The internal dynamics of the ankle is described as an arc vector, the knee being its center, of radius equal to the length of the lower leg. The ankle is perceived to move in half-pendulum fashion rather than manifesting the motions imparted to it by the dynamics of the knee and hip. In a similar manner, the motions of the upper body component structures are described as a nesting of pendulum swings. Thus, all of the component structures of the lower and upper body are described as a nesting of pendulums, the knee and elbow being described both as bobs, relative to their hip and shoulder pivots, and as pivots, relative to their ankle and wrist bobs.

We suggest that in perceiving a human walker from a dynamic display of lights attached to joints, we extract information in logical steps that result in the parts of the body being perceived as a system of dynamic nested dependencies. The motions and locations of the hip and shoulder are extracted first as they are related to the body's deepest center of moment, that within the torso. Each step of information extraction that follows takes the previously described part as the center of moment for the determination of dynamic relations of other parts. The tree diagram presented in Figure 8 shows the steps followed in extracting information, and the terminal elements define the perceived components of motion and spatial location for each joint marked by a light.

Recall that we chose to describe a walker in Figure 8 as if he or she were on a treadmill. These motions are also preceived when viewing a walker traversing a terrain. In this latter case, however, one additional component of motion is perceived—the motion of the walker relative to the observer. As early Hollywood productions have shown, a person filmed walking on a treadmill and one filmed walking on solid ground will be indistinguishable if the background in the former situation is moved appropriately. The pendular movements of the arms and legs, together with the torsion of the hips and shoulders about their center of moment, are perceived as wholes in both cases. The motion of a walker relative to an observer is not imparted to each of the body's parts; rather, it is perceived as a distinct component describing the dynamics of the whole.

Remember, we define the movement of a figure relative to an observer as its action. A figure is defined by motions and locations of its parts relative to its center of moment; thus, for the purpose of figural description, the center of moment is a static point of reference. However, with respect to the motion of the figure as a whole, the deepest center of moment serves as that point embodying the dynamics of the whole. The action of a walker is the movement of the torso's center of moment and is described by a fairly uniform translational vector. (The slight up-and-down motion of the torso's center of moment was, since it is an action, ignored in the previous discussion of the treadmill walker, but see Cutting, 1978a, Experiment 5.) Some events have no action, and may thus be adequately described in figural notation. A clock's pendulum, for example, has no action except when the whole clock is moved. Other events, such as walkers, rolling wheels, or falling leaves, are perceived to move relative to the observer and these have both figural and action specifications.

5. Summary

We suggest that event perception proceeds through a logical series of steps for extracting stimulus information, each step resulting in the definition of distinct components. The first step isolates an event from the ground in which it is embedded. The second step divides the event into the components of figure and action. A figure is described by the internal dynamics of its parts about its center of moment taken as a static entity. Action is the dynamic property of the figure's center of moment. Our account is in accord with the Gestalt dictum that one does not perceive the motions of a uniform whole by perceiving individually the motions of its various parts; however, we go further in asserting that the motions of a whole are perceived as the movements of parts. *More particularly, the internal dynamics of a whole are perceived as the movement of parts about a center that is analytic to the whole, whereas the observer-relative displacement of the whole is perceived as the dynamics of that point.*

We believe that this general description holds for a whole variety of events, including the perception of humans walking, wheels rolling, faces aging, oneself flying, and the night sky revolving. We have only begun to explore the domain that we call event perception; we recognize that there are myriads of questions that our analysis raises; and we anticipate that our grammatical analysis may not be appropriate to all those things that psychologists may want to call events. Nevertheless, we are excited about the prospect of a structural description of events that follows the organizing principles of visual perception. The cornerstone of this approach is the recognition that dynamics is primary and that static arrays must be fit into a dynamic scheme.

6. References

Barclay, C.D., Cutting, J.E., & Kozlowski, L.T. Temporal and spatial factors in gait perception that influence gender recognition. *Perception and Psychophysics*, 1978, *23*, 145–152.

Boyer, C.B. *A history of mathematics*. New York: Wiley, 1968.

Braunstein, M.L. *Depth perception through motion*. New York: Academic Press, 1976.

Carlsöö, S. *How man moves*. London: Heinemann, 1972.

Cutting, J.E. Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception*, 1978, *7*, 393–405. (a)

Cutting, J.E. Perceiving the geometry of age in a human face. *Perception and Psychophysics*, 1978, *24*, 566–568. (b)

Cutting, J.E. A program to generate synthetic walkers as dynamic point-light displays. *Behavior Research Methods and Instrumentation*, 1978, *10*, 91–94. (c)

Cutting, J.E., & Kozlowski, L.T. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 1977, *9*, 353–356.

Cutting, J.E., Proffitt, D.R., & Kozlowski, L.T. A biomechanical invariant for gait perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, *4*, 357–372.

Dodge, R. An experimental study in visual fixation. *Psychological Review Monograph Supplements*, 1907, *8*(4), 1–95.

Dürer, A. *The human figure: The Dresden sketchbooks*. New York: Dover, 1972. (Originally published, 1528.)

Emlen, S. The stellar-orientation system of a migratory bird. *Scientific American*, 1975, *233*(2), 102–111.

Enlow, D. *Handbook of facial growth*. Philadelphia: Saunders, 1975.

Faust, M.S. Somatic development of adolescent girls. *Monographs of the Society for Research in Child Development*, 1977, *42*, Serial No. 169.

Gibson, J.J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.

Gibson, J.J. *The senses considered as perceptual systems*. Boston: Houghton Mifflin, 1966.

Gibson, J.J. *The ecological approach to visual perception*. Boston: Houghton Mifflin, 1979.

Gray, H. *Anatomy, descriptive and surgical*. 15th Ed., 1901. (Facsimile Ed., New York: Bounty, 1977.)

Green, B.F. Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, 1961, *62*, 272–282.

Johansson, G. *Configurations in event perception*. Uppsala, Sweden: Almqvist & Wiksell, 1950.

Johansson, G. Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 1973, *14*, 202–211.

Johansson, G. Projective transformations as determine visual space perception. In R.B. MacLeod & H.L. Pick, Jr. (Eds.), *Perception: Essays in honor of James J. Gibson*. Ithaca: Cornell University Press, 1974.

Johansson, G. Visual motion perception. *Scientific American*, 1975, *232*(6), 76–89.

Johansson, G. Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research*, 1976, *38*, 379–393.

Koffka, K. *Principles of Gestalt psychology*. New York: Harcourt, 1935.

Köhler, W. *Gestalt psychology*. New York: Liveright, 1947.

Kozlowski, L.T., & Cutting, J.E. Recognizing the sex of a walker from a dynamic point-light display. *Perception and Psychophysics*, 1977, *21*, 575–580.

Kozlowski, L.T., & Cutting, J.E. Recognizing the gender of walkers from a dynamic point-light display: Some second thoughts. *Perception and Psychophysics*, 1978, *23*, 459.

Liberman, A.M., Ingemann, F., Lisker, L., Delattre, P.C., & Cooper, F.S. Minimal rules for synthesizing speech. *Journal of the Acoustical Society of America*, 1959, *31*, 1490–1499.

Locke, J. *An essay concerning human understanding*. New York: Dover, 1959. (Originally published, 1690.)

Maas, J.B., Johansson, G., Jansson, G., & Runeson, S. *Motion perception I: 2-dimensional motion perception*. Boston: Houghton Mifflin, 1971 (film).

Mach, E. *The science of mechanics*. 1893. Reprinted, LaSalle, Ill.: Open Press, 1960.

Marey, E.J. *Movement*. New York: Arno Press & New York Times, 1972. (Originally published, 1895.)

Murray, M.P. Gait as a total pattern of movement. *American Journal of Physical Medicine*, 1967, *46*, 290–333.

Murray, M.P., Kory, R.C., & Sepic, S.B. Walking patterns of normal women. *Archives of Physical Medicine and Rehabilitation*, 1970, *51*, 637–650.

Neisser, U. *Cognition and reality*. San Francisco: W.H. Freeman, 1976.

Pittenger, J.B., & Shaw, R.E. Aging faces as viscal-elastic events: Implications for a theory of nonrigid shape perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, *1*, 374–382.

Proffitt, D.R., & Cutting, J.E. Perceiving the centroid of configurations on a rolling wheel. *Perception and Psychophysics*, 1979, *25*, 389–398.

Proffitt, D.R., Cutting, J.E., & Stier, D.M. Perception of wheel-generated motions. *Journal of Experimental Psychology: Human Perception and Performance,* 1979, *5,* 289–302.

Rubin, E. *Synoplevede figure.* Copenhagen: Gyldendalske, 1915.

Saunders, J.B., Inman, V.T., & Eberhart, H.D. The major determinants in normal and pathological gait. *Journal of Bone and Joint Surgery,* 1953, *35-A,* 543–558.

Thompson, D.W. *On growth and form.* 1917. Reprinted, London: Cambridge University Press, 1969.

Turvey, M.T. Contrasting orientations to the theory of visual information processing. *Psychological Review,* 1977, *84,* 67–88.

Wallach, H. Visual perception of motion. In G. Keyes (Ed.,) *The nature and the art of motion.* New York: George Braziller, 1965.

Wolff, W. *The expression of personality.* New York: Harper, 1943.