Bulletin of the Psychonomic Society
1977, Vol. 9 (5), 353-356

Recognizing friends by their walk: Gait perception without familiarity cues

JAMES E. CUTTING and LYNN T. KOZLOWSKI
Wesleyan University, Middletown, Connecticut 06457

Viewers can recognize themselves and others in an abstract display of their movements. Light sources mounted on joints prominent during the act of walking are sufficient cues for identification. No other information, no feedback, and little practice with such a display are needed. This procedure, developed by Johansson, holds promise for inquiry into the dimensions and features of event perception: It is both naturalistic and experimentally manageable.

People often believe that they can recognize friends by their walk. Unfortunately, this belief and the previous research on the topic (e.g., Wolff, 1943) are confounded by familiarity cues, size and shape cues, or other nongait sources of information such as probabilities of seeing a person at a given place or time. We demonstrate that viewers can recognize themselves and others in a dynamic display of their movements when these factors are controlled.

We were stimulated by the work of Johansson (1973, 1975), particularly his films (Maas & Johansson, 1971a, b). When viewing them, one sees people stripped of familiarity cues such as clothing and hairstyle; people are presented as arrays of point-light sources moving across a screen in an orderly fashion. Johansson's technique seemed to be ideal for the study of how ecological events are perceived.

A partial taxonomy of events has been proposed by Shaw, McIntyre, and Mace (1974). Some relevant distinctions are those of (1) fast vs slow events, where the critical feature is whether dynamic change can be perceived directly or only inferred, (2) reversible vs irreversible events, (3) rigid vs plastic events, and (4) events associated with animate vs inanimate sources. Most psychologists have concentrated on the perception of fast, reversible, rigid, inanimate events (e.g., Börjesson & von Hofsten, 1973; Johansson & Jansson, 1968). A few, however, have begun to study slow, irreversible, elastic, animate events, such as the aging of faces (Pittenger & Shaw, 1975a, b). Walking is an intermediate type of event: It is fast, animate, irreversible, and also rigid—that is, composed of a hierarchy of rigid pendular motions.

Gibson (1950) has argued that the perception of any moving shape can be thought of as the perception of formless invariant relations displayed over time. The study of gait or any other system of events should consider the interrelation of two component invariants: the underlying dynamic aspect of the event, or the transformational invariant, and the underlying unity of the structures involved, or the structural invariant (Pittenger & Shaw, 1975a; Shaw & McIntyre, 1974). In the present paper we observe whether a particular aspect of the structural invariant (the identity of the walker) is sufficiently presented through the transformational invariant (walking) for recognition.

Supported by research grants from Wesleyan University to both authors. We thank Robert J. White and Deborah A. Cassidy for technical assistance. Requests for reprints should be sent to the authors at the Department of Psychology, Wesleyan University, Middletown, Connecticut 06457. J. Cutting is also a staff member of the Haskins Laboratories.

METHOD

Our study of gait used glass-bead retroreflective tape wrapped around walker's joints, video-tape recording equipment, and bright lights focused on the walking area and mounted close to the lens of the television camera. The contrast of the image on the television monitor was turned to maximum, and the brightness to minimum, so that only the reflectant patches could be seen (see Johansson, 1973). Static approximations to our stimuli can be seen in Figure 1. Figure 2 shows one of our walkers with the image brightness turned up.

Six Wesleyan University undergraduates, three males and three females, served as walkers. Each had a normal gait. They were approximately the same height and weight, and they lived together in university housing. All wore tight-fitting dark clothing during the recording session. We wrapped 5-cm-wide commercially available reflectant tape around their wrists, around their arms just above the elbow, around their ankles, and around their legs just above the knee. We affixed 5 x 18.5 cm patches to their belts at the hip and to their shoulders as epaulets, half on the shoulder and half on the upper arm. No patch was placed on the head. Each individual walked at a normal pace for several minutes until we were satisfied that he or she was not "performing" before the camera. We then recorded side views of each as he or she walked in front of the camera 8 m from the lens. Each individual walked back and forth 10 times, while his or her friends waited in another room. Individuals were on camera for five strides (±¼ stride) and a mean of 2.7 sec during each pass across the viewing field. The camera was fixed and did not pan to follow the walker.

A test tape consisting of all tokens of all walkers was created by recording onto a second video tape. We used two helical-scan recorders, a monitor connected to one recorder (on which the source tape was played), and a television camera focused on the monitor at close range and connected to the second recorder (on which the test tape was recorded). Each token was selected in random order and recorded onto the test

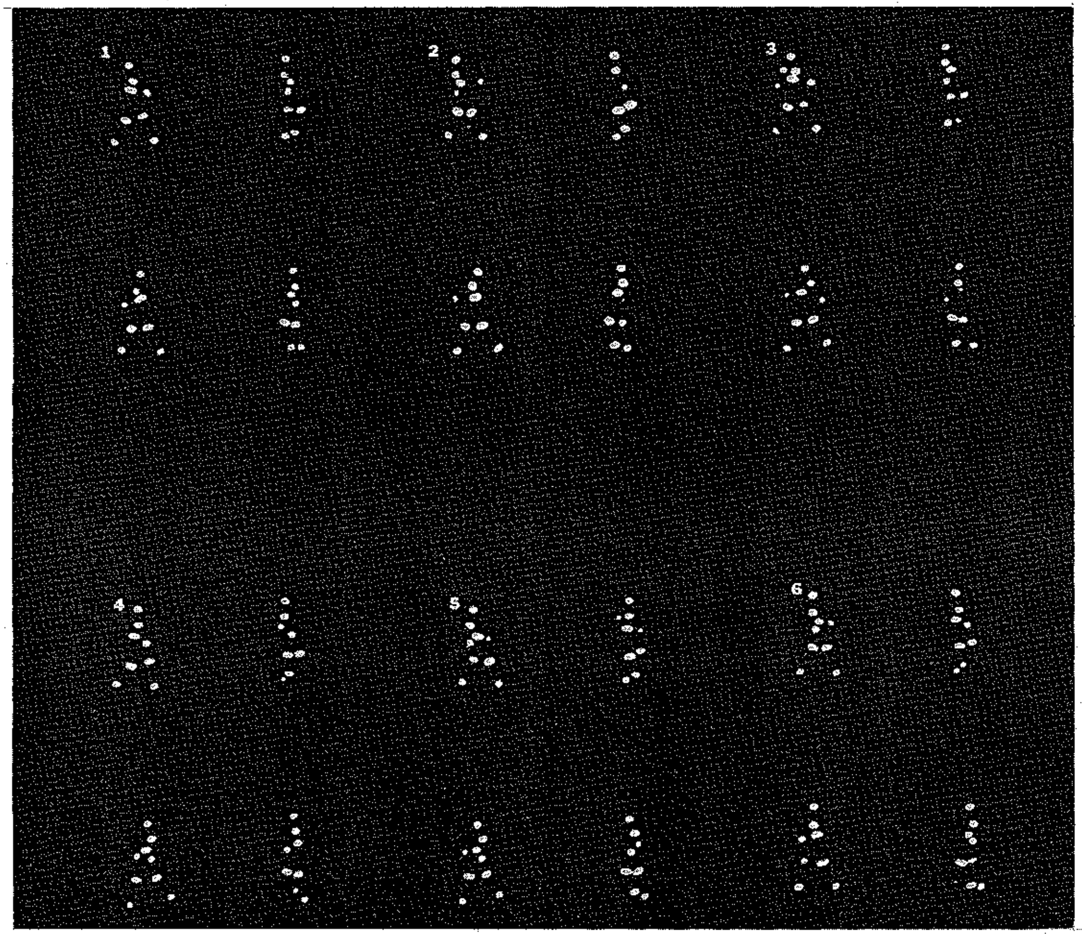

Figure 1. Each of the six walkers is portrayed in a cluster of four point-light configurations. Those at the top of each cluster show the individual walking from left to right, those at the bottom from right to left. The two configurations on the left portray the walker with arms and legs most outstretched, and those on the right with body most aligned, on one stiff leg with the other slightly raised. Walkers 1, 2, and 3 are female; 4, 5, and 6 are male. Dynamic arrays, but not the static ones shown here, are adequate for recognition of a particular person.

tape, yielding 60 trials: (6 walkers) × (10 separate tokens per walker). Each trial consisted of the individual walking from left to right across the monitor screen, followed by a 3- to 5-sec pause and the same person walking from right to left. An interval of 15 to 20 sec occurred between trials to insure that no "flop-over" of image on the monitor would occur within a trial. Barrel distortion and other degradations of the image due to rerecording were minimal.

Two months after the recording session we invited the six walkers back to determine if they could recognize one another. A seventh undergraduate, who knew the six well, also served as a viewer. All sat in a dimly lit room and viewed the test sequence on a monitor. For all viewers, the dynamic shapes subtended a visual angle of about 5 deg measured vertically. For each trial, viewers wrote the name of the walker and indicated confidence in their responses using a five-point unipolar scale. Before viewing, they ranked the individuals for how easy they would be to recognize by their walk; after viewing, they wrote a few phrases about how they had made their decisions about the identity of each walker. The entire session lasted about 40 min.

RESULTS AND DISCUSSION

Our viewers did well, although they were far from perfect. Chance performance would be 16.7% correct identification. Overall, correct responses occurred on 38% of all trials [$t(6) = 4.6$, $p < .005$, one-tailed]. The range of performance was 20% to 58%, as can be seen in Table 1. Interestingly, the best viewer was a dancer and she was also second easiest to recognize. Although no feedback was given, performance improved over the course of the task. Viewers increased from 27% correct identification for the first three presentations of

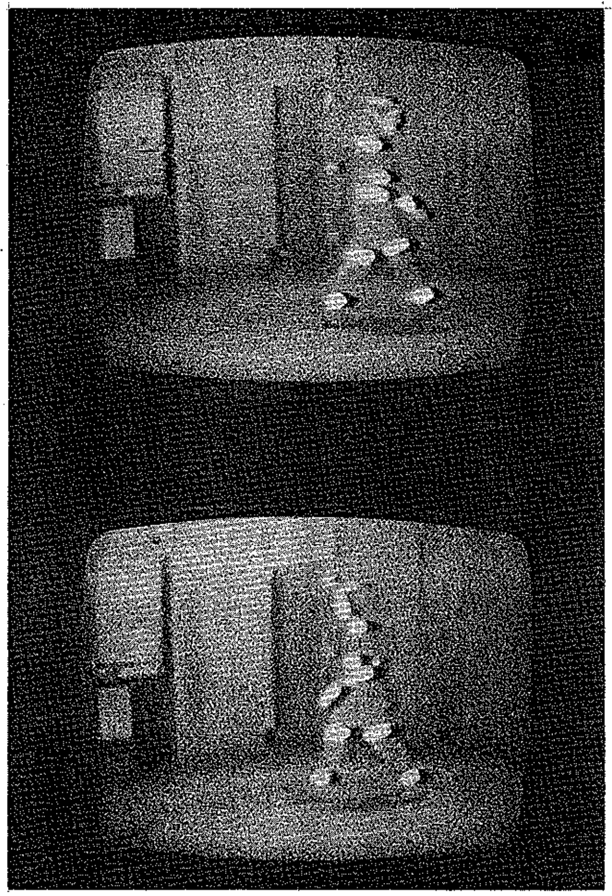

Figure 2. Walker 1 in the most outstretched positions. Brightness of the image is turned up to reveal the approximate location of the reflectant tape. Blurring and fuzziness are due to both the nature of the medium and to the rerecording process.

each of the six walkers (range: 17% to 39%), to 59% for the last three (range: 22% to 95%) [$t(6) = 3.1$, $p < .025$ for this increase]. The confidence judgments reflect the fact that viewers knew what they were doing. When maximally confident, they were 75% correct (range: 31% to 100%) [$t(6) = 6.2$, $p < .001$]; when minimally confident, they were only 24% correct (range: 0% to 43%) [$t(6) = 1.3$, not significantly better than chance]. Finally, self-recognition was not significantly better than the recognition of others (43% vs 36%), a result that fails to replicate Wolff (1943, pp. 88-90).

How Do Viewers Do It?: Direct Percept vs Conscious Problem Solving

From the very first trial, viewers had no difficulty envisioning a person walking in the point-light display: The percept seems to be direct and immediate, as Johansson suggests. Do viewers come to see particular persons walking the point-light presentations or do they learn how to match limited information from the display with stored knowledge of the walkers' gaits? The viewers' answers to the question "How did you recognize each of the walkers?" gives us some indication of how they dealt with the task. It is noteworthy that none said, for example, "It simply looked like Marty walking" or "I saw Elaine walking," although demand characteristics may have contravened any such tendency. Instead, viewers tended to mention as clues certain critical features of the display, such as the speed, bounciness, rhythm of the walker, amount of arm swing, or the length of steps. The most accurate viewers claimed to associate these dynamic aspects of the display with particular individuals. Interestingly, the poorest viewer was the only one who reported using height of the walkers (a nondynamic aspect that was controlled in the selection of walkers) as a way of deciding.

Our subjects ranked how well they would be able to recognize the walkers, but they were strikingly inaccurate in their opinions. The correlation between the ranked ease of recognition and the percent correct responses was small and insignificant (rho = −.06).

Although this account is tentative and speculative, we are led to believe that the viewers were acting in large part as conscious problem solvers rather than as direct perceivers of particular individuals walking in the dynamic displays. From the start of the session, when their task was made clear to them, they may have tried to think of characteristic features of the walkers. Perhaps many of these features were not actually presented in the abstract arrays, so the viewers had to reassess the features that would allow them to identify the walkers. This process of hypothesis testing might account for the relatively poor performance on the early trials and the improved performance in later trials. Without feedback, viewers learn what they can extrapolate from the displays and use it to better their performance. It is our impression that performance would continue to improve without feedback, and that one or two trials with feedback for each walker would yield essentially perfect performance. What is re-

Table 1
A Matrix for the Identification of Self and Others, as Represented by Percent Correct

	Walkers						
	Nancy	Mary	Elaine	Mark	Marty	Lane	
Viewers	1	2	3	4	5	6	Mean
Nancy	*10*	30	60	30	40	0	28
Mary	10	*20*	40	30	50	50	33
Elaine	78	40	*78*	38	78	40	58
Deborah	44	40	89	13	78	50	53
Mark	10	40	40	*50*	80	10	38
Marty	0	30	10	60	*20*	0	20
Lane	0	20	50	20	80	*70*	40
Mean	21	31	51	35	60	31	38

Note—Self-identification is shown in italics. There are missing data for viewers Elaine and Deborah. Numbers of walkers correspond to those in Figure 1.

markable is how little experience is needed with a dynamic point-light display.

CONCLUSION

Experimental psychology has largely ignored the dimension of time in the stimulus (Jones, 1976). In vision, for example, we have typically tried to stop time, slicing it finer and finer with the tachistoscope. Such snapshots of the visual processing system can reveal many kinds of operations, and their usefulness should not be undervalued. Nevertheless, they remain static views of a fundamentally dynamic system. Over the past 15 years, the application of information-processing techniques to vision has introduced time into the experimental situation, but in an incomplete fashion. Time is taken to be a variable pertinent to the perceiver but not to the stimulus. For example, a brief stimulus may be presented, then masked after an intervening interval so as to disrupt perception in the viewer. But the stimulus is still a snapshot, and very few snapshots appear in our world except on the printed page. A testament to their pecularity is that as pictures the perception of these flat images must be learned (Kennedy, 1974), whereas the perception of movement appears innate. Movement, of course, occurs over time.

According to conventional wisdom, the perception of moving shapes is derived from the perception of static forms. That is, movement is thought to be impleted, or compiled, from snapshots, much as the continuity in film is compiled from its separate frames. The study of apparent movement can be interpreted as a case in point, where geometric forms can appear to move and change shape between stimulus presentations (see Kolers & Pomerantz, 1971). However, the perception of dynamic forms is probably not derived from the perception of static forms. Instead the logic of their relation might be reversed: Static form perception can be viewed as a special case of dynamic form perception where the transformationally invariant information is held null (Pittenger & Shaw, 1975a). This new position easily leads one away from the study of stimuli at a given time and toward the study of events over time. Walking is one such event.

Johansson (1973, 1975) has demonstrated that a dynamic array of point lights is sufficient to recognize the presence of a walker. Our study goes further, demonstrating that the same array is sufficient for the recognition of a *particular* walker. Static or very brief presentations are insufficient. For example, Johansson (1976) has demonstrated that 100-msec presentations of dynamic point-light displays cannot be recognized as people walking or running, whereas slightly longer displays are entirely adequate. Preliminary results of our own corroborate Johansson's findings.

The primary advantage of Johansson's technique is that it is both manageable and naturalistic. With it one can achieve experimental rigor and still deal with common, everyday events. Like others, we think that these are the events that psychologists should study.

REFERENCES

Börjesson, E., & von Hofsten, C. Visual perception of motion in depth: Application of a vector model to three-dot motion patterns. *Perception & Psychophysics*, 1973, **13**, 169-179.

Gibson, J. J. *Perception of the visual world*. New York: Houghton Mifflin, 1950.

Johansson, G. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 1973, **14**, 201-211.

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.

Johansson, G., & Jansson, G. J. Perceived rotary motion from changes in a straight line. *Perception & Psychophysics*, 1968, **4**, 165-170.

Jones, M. R. Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 1976, **83**, 323-355.

Kennedy, J. M. *A psychology of picture perception*. San Francisco: Jossey-Bass, 1974.

Kolers, P. A., & Pomerantz, J. R. Figural change in apparent motion. *Journal of Experimental Psychology*, 1971, **87**, 99-108.

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

Maas, J. B., & Johansson, G. *Motion perception II: 3-dimensional motion perception*. Boston: Houghton Mifflin, 1971 (film). (b)

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. (a)

Pittenger, J. B., & Shaw, R. E. Perception of relative and absolute age in facial photographs. *Perception & Psychophysics*, 1975, **18**, 137-143. (b)

Shaw, R. E., & McIntyre, M. Algoristic foundations to cognitive psychology. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes*. Hillsdale, N.J: Erlbaum, 1974.

Shaw, R. E., McIntyre, M., & Mace, W. M. The role of symmetry in event perception. In R. B. MacLeod & H. L. Pick (Eds.), *Perception: Essays in honor of J. J. Gibson*. Ithaca, N.Y: Cornell University Press, 1974.

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

(Received for publication December 23, 1976.)