

# Developmental Changes in Infants' Perceptual Processing of Biomechanical Motions

**Jeannine Pinto**  
Psychology Department  
University of Virginia  
Charlottesville, VA 22903  
jmp8p@virginia.edu

**Jeff Shrager**  
Xerox  
Palo Alto Research Center  
Palo Alto, CA 94304  
shrager@xerox.com

**Bennett I. Bertenthal**  
Psychology Department  
University of Virginia  
Charlottesville, VA 22903  
bib@virginia.edu

## Abstract

In order to process the reduced information in point-light images of human movement, observers rely upon general processing heuristics as well as representations more specific to human gait. This paper explores changes in the perception of structure from motion in young infants. We re-examined data from 17 experiments, involving infants of 3- and 5-months old, to determine which stimulus features of point-light motion infants use to organize percepts, and how perception changes. By combining discrimination and encoding information we provide a picture of developing perceptual processes. Five-month-olds encode the stimuli more quickly than 3-month-olds, while the younger infants discriminate pairs of stimuli more frequently. Infants of both ages use phase information to discriminate displays. Three-month olds discriminate canonical forms from modified forms when the stimuli are organized about a vertical axis, whereas 5-month olds discriminate these forms only when one of the figures take on a human-like configuration. These results support a view in which differential skill in what is encoded characterizes development. Furthermore, this work may help guide the integration of theory-formation models with heuristic and constraint-based models, into a more complete account of perception.

Motion can provide observers with important information about the 3-dimensional world (Gibson, 1950). From motion the visual system can extract information about depth, grouping, and other components of structure. Johansson (1973) introduced mov-

ing dot displays of human movement into the exploration of structure from motion. These displays are created by attaching "point lights" to a person's major joints, and filming the person moving in a dimly-lit room. The resulting film provides only the motion information characteristic of human movement; features such as texture, color, and explicit contour have been eliminated. Though a human form is hard to recognize from a single, static frame of a point-light film, Johansson (1973) demonstrated that adult observers can recognize the moving displays as moving persons in as little as 1 second.

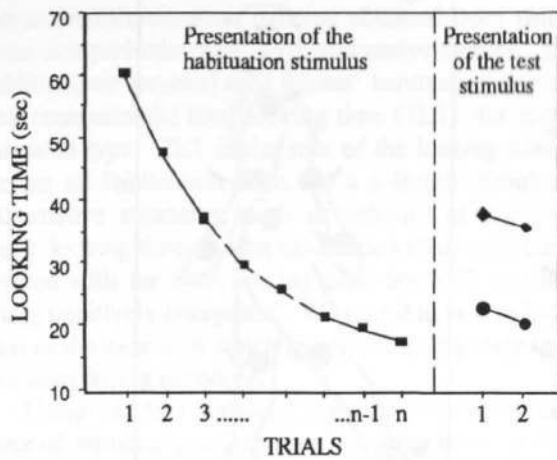
Adults almost always identify the point-light displays as depicting human movement, yet the number of possible connections between the 11 point-lights<sup>1</sup> that compose a human form is very large. This suggests that sensory information alone cannot organize a visual array into an object; the visual system organizes the scene with the aid of constraints, or organizing heuristics. Previous research suggests that the visual system exploits the relative motions among point-lights. Human gait cannot be identified in the absolute motion of any individual element. Rather, form emerges as a product of the relative motions of the elements, or motion of the elements in relation to one another. Some relative motions are more salient than others. Wallach & O'Connell (1953), for example, have suggested that, whenever possible, the visual system assumes that relative motions of the elements reflect rigid relations and thus tries to find the simplest interpretation under which the points are rigidly related in 3D. In their account, though non-rigid interpretations are possible, the rigidity assumption privileges one set of interpretations over another.

---

This work conducted under NICHD grant HD16195 and NICHD Career Development Award HD00678 to the third author.

---

1. The point-lights appear to be attached to the head, shoulders, elbows, wrists, hips, knees and ankles. In a side view of the walker, one shoulder and hip are always occluded by the rest of the body.



**Figure 1.** A hypothetical habituation curve. In the infant-control paradigm, the number of trials is variable, determined when the infant reaches approximately half of the initial 3 trials. On the right, two lines illustrate discrimination (diamonds) and lack of discrimination (circles) of the test stimulus.

In addition to general processing heuristics, like the rigidity assumption, evidence from studies of adults suggests that observers also rely upon processes or representations more specific to human gait to help organize point-light walker displays. Adult observers can identify various movements, such as jogging, jumping, or dancing (Johansson, 1975), and the gender of the walker (Cutting, 1978). In so far as this skill is acquired through extensive experience of movement, adults are likely to develop robust strategies and redundant constraints for interpreting human movement. Thus, examinations of the development of the perception of biomechanical motions can provide insights into early perceptual constraints and their interactions with cognitive development.

This paper explores the perception of structure from motion in young infants. Previous work has used measures of discrimination between point-light displays to assess infants' subjective experience. By summarizing a number of discrimination findings, we have developed a picture of the perception of biomechanical motions at 3- and 5- months of age. We have also sought converging evidence for that picture from data about infants' encoding processes as they are presented with point-light displays. We conclude with some general implications of our analyses for computational views of the development of perception.

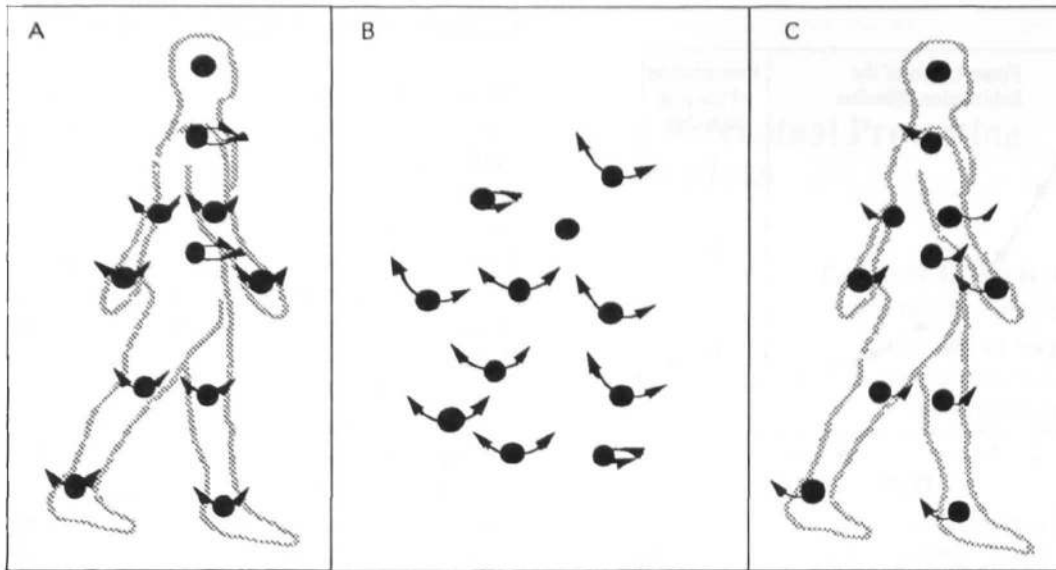
## Tapping Infants' Subjective Experience

The analysis reported here is a re-examination of 17 experiments conducted in the Laboratory for Infant Studies at the University of Virginia, involving 387 infants in two age groups: 3-months old (186 infants) and 5-months old (201 infants)<sup>2</sup>. The goal of these studies was to determine which stimulus features of point-light motion infants use to organize the images, and how this skill develops between 3 and 5 months of age.

Each of the 17 studies used a habituation paradigm to test infants' discrimination of one or more stimulus dimensions. Behavioral habituation is a common tool for studying infant perceptual and cognitive processing. In this paradigm, infants are familiarized, or habituated, to one stimulus and then presented with a second stimulus. The infant's differential response to the two stimuli is used as the measure of pairwise discrimination. The studies reported in this paper were conducted using an infant-control habituation paradigm (Cohen, 1973). In this paradigm, infants are repeatedly presented a visual stimulus (the "habituation" stimulus) and the time they spend looking at this stimulus is measured by trained observers. Gradually, over successive presentations of the same habituation stimulus, infants' looking time per trial declines (Figure 1). When looking time reaches a criterion value, a variant of the habituation stimulus is presented for two trials (the "test" stimulus). The infant's time looking at the test stimulus is then compared to its time looking at the last two presentations of the habituation stimulus. If the infant dishabituates (i.e., looks longer -- with renewed attention) to the test stimulus, the infant is said to discriminate between the two stimuli. Since the stimuli are generally constructed to differ on a single dimension, evidence of discrimination provides evidence of the infant's sensitivity to that dimension, from which we infer that that dimension is involved in processing of the image.

Researchers use pairwise discrimination patterns across many stimulus pairs to piece together theories of what infants organize, perceive, and know, but less attention is generally paid to the habituation data itself. To the extent that performance in the habituation phase of the study is tapping the process of encoding the habituation stimulus, habituation itself can tell us something about perception. In comparisons across stimuli, the habituation trials offer a measure of the relative ease with which stimuli are

2. Each infant was tested only at one age and in one condition.



**Figure 2.** Point-light walker displays. A. Canonical walker. The arrows indicate the pendular motions of each point-light. B. Scrambled walker. The elements preserve only the absolute motions of the canonical walker. C. Phase-shifted walker. While the point-lights on a limb of the canonical walker move together, the point-lights of the phase-shifted walker move asynchronously, sometimes in opposition to one another, as the arrows suggest.

organized during encoding. Ease of encoding, operationalized as the sum of an infant's looking times across habituation trials, is a reflection of the interaction between the infant's organizational processes and the structure of the stimulus. Thus, by the combination of the discrimination information available from the habituation/ dishabituation analyses, and the encoding information available in the habituation trial data, we can develop a fairly detailed picture of the overall processes of perception and its development.

### Point-light Displays Of Human Gait

The studies we selected for analysis all involved computer-synthesized point-light walker displays (Cutting, 1978; Bertenthal & Kramer, 1984) and foils in which selected stimulus parameters had been modified. The standard, or canonical, displays (Figure 2a) were comprised of 11 dots that mimic the kinematics of human gait, approximated by the harmonic motions of hierarchically-nested pendula. Adults identify this form as depicting a person walking on a treadmill (Bertenthal & Davis, 1988). Infants show the same pattern of looking during habituation to these synthesized displays as to the displays created using a human model. Foils were constructed varying the relations among the point-light elements, the orientation, and global form.

**Relations among the elements:** The *scrambled* display (Figure 2b) was comprised of 11 point-lights with individual motion vectors corresponding to those in the canonical display, however, the spatial configuration among elements was random. The *phase-shifted* display (Figure 2c) maintained the individual motion vectors as well as the hierarchical spatial configuration of the canonical form, but altered the relative motions among the elements. In the canonical displays, point-lights cycled synchronously together or in opposition. In phase-shifted displays, point-lights no longer maintained synchronicity, appearing much like a marionette. Researchers examining motor behavior have suggested that phase relations are an important invariant in human locomotion (Schmidt, 1985). In *non-rigid* displays (not shown), the distance between the point-lights was randomly varied from frame to frame.

**Orientation:** Displays were oriented upright, inverted, or horizontal in the picture plane.

**Global form:** In addition to human forms, displays looking much like four-legged "spiders" were created. These forms were comprised of four limbs, identical to human limbs, radiating from a single point.

Each of the 17 studies was designed to test infants' sensitivity to the spatiotemporal properties of point-light walkers across the dimensions just de-

scribed. Sensitivity was assessed on the basis of habituation/dishabituation patterns obtained from pairwise comparisons made over successive studies. In addition, we obtained each infants' habituation curve, and computed the total looking time (TLT) for each stimulus type. TLT is the sum of the looking times across all habituation trials for a selected stimulus. Alternative measures, such as velocity of decline, peak looking time, or first habituation trial, were correlated with the total looking time but TLT can be more intuitively interpreted. We take it to be a reflection of the ease with which an infant can organize the stimulus during encoding.

Using the TLT measure, we examined the influence of stimulus parameters on encoding times to determine the relative ease of encoding, and by inference, the processing constraints available to the infant. Assessments of discrimination could be made only for stimulus pairs actually tested, but assessments of ease of encoding could be made by way of pseudo-experiments. In these analytic studies, we compiled all habituation curves produced in response to a stimulus exhibiting the property we wished to examine. For example, in order to examine the relative ease with which infants encode canonical versus scrambled forms, we gathered the habituation curves from all infants who had seen either stimulus as the

familiarization display, and compared the total looking times for these data.

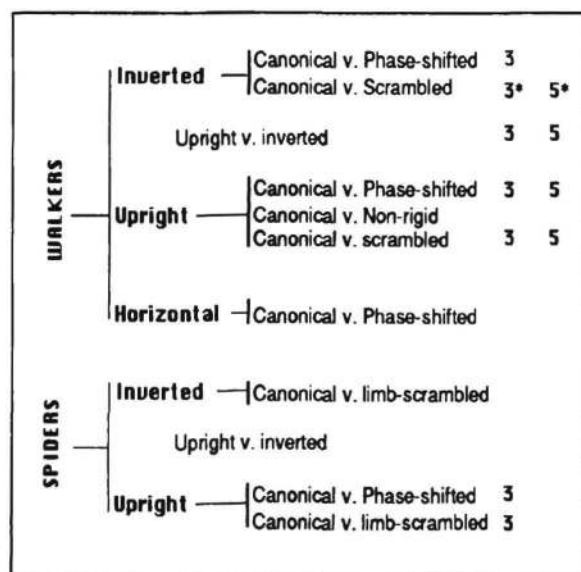
## Results

Several researchers (e.g., Bronson, 1982; Sokolov, 1960) suggest that infants' visual attention is guided by knowledge and expectancy. If this is the case, 3- and 5- month-olds, who are likely to differ in knowledge, should exhibit different patterns of encoding and discrimination. Moreover, we expected to find a systematic relationship between differential encoding times and discrimination.

Data were organized according to age and stimulus parameters, specifically global form (walker, spider), orientation (upright, inverted, horizontal), and element relations (canonical, scrambled, phase-shifted, nonrigid). The results of a re-analysis of pairwise discrimination by each age group are summarized in Figure 3.

In general, when assessed by either discrimination or encoding behavior, 3- & 5- month-olds respond differently to the point-light displays. As shown in Figure 3, three-month-olds discriminated 7 of the 11 pairs of stimuli, whereas 5-month-olds discriminated only 4. Though differences in encoding times did not mirror discrimination patterns, total looking time also differed systematically by age. Three-month-olds ( $M=203.1$ ), across stimulus parameters, looked longer during the habituation process than did 5-month-olds ( $M=120.68$ ),  $F(1,385)=59.23$ ,  $p<.01$ . This general change in performance may result from maturing neural and motor processes (Johnson, Posner, & Rothbart, 1991), from increased skill or knowledge, or from changes in motivational factors. Though this data cannot directly discern between these accounts, it contributes additional support to positions which posit global change; 5-month-olds habituated faster than did 3-month-olds across all stimuli.

Three-month-olds discriminate forms with differing relations among the elements when those stimuli are organized about a vertical axis. These infants discriminate canonical from phase-shifted forms and canonical from scrambled forms when they are upright or inverted, but not when they are horizontal. Bornstein & Krinsky (1985) report evidence which suggests that 3-month-old infants may show sensitivity to the properties of a form only when it is organized about a vertical axis. Interestingly, at 3 months of age TLT to vertical stimuli ( $M=208.3$ ) was significantly greater than TLT to horizontal stimuli ( $M=120.8$ ),  $F(1,184)=5.58$ ,  $p<.05$ , suggesting a corre-



**Figure 3.** Summary of discrimination analyses. The number to the right of the tree indicates the age group exhibiting discrimination, as indicated when the mean of the test trials is significantly greater than the mean of the last 2 habituation trials. (The \* indicates an order effect.)

spondence between higher looking time and subsequent discrimination of stimulus properties.

By 5 months of age, infants are faster to encode stimuli as well as less apt to show discrimination on pairwise comparisons, suggesting that they attend more selectively to information available in the displays. Five-month-old infants showed few differences in encoding times as well. These infants looked significantly longer at out-of-phase walkers ( $M=153.4$ ) than at in-phase walkers ( $M=102.5$ ),  $F(1,193)=15.6$ ,  $p<.01$ . In addition, 5-month-olds looked longer at out-of-phase walkers than at scrambled forms ( $M=83.83$ ),  $F(1,72)=2.99$ ,  $p<.05$ . Encoding times to upright canonical and scrambled forms did not differ. Five-month-olds, like 3-month-olds, tend to discriminate stimulus parameters only when they are presented in a particular configural context, or frame. While they encode phase information generally, 5-month olds discriminate canonical and phase-shifted forms only when the figures are presented upright and in the context of a walker-like form (hierarchically organized along a vertical axis of symmetry). Phase-relations in spider-like forms or horizontal orientations are not discriminated. Thus, 5-month-old infants may be responding in part on the basis of some familiarity with the characteristic phase-relations of human gait. Indeed, prior research (Pinto & Davis, 1991; Bertenthal, 1992) suggests that 5-month-olds, but not 3-month-olds, respond on the basis of a category of human movement.

## Discussion

We conceive of visual perception as a process whereby information is organized both by general processing heuristics and in accord with knowledge (processes and/or representations) that arises from experience with the domain—in this case, biomechanical motion. We have provided evidence of the development both of general processing heuristics (or the development of their use) and of knowledge that is specific to the perception of biomechanical motion. Both 3- and 5-month old infants use phase information to discriminate point-light displays, but 5-month olds show an encoding difference in phase-shifted vs. canonical displays. Since 3-month-olds discriminate the stimuli along most of the dimensions examined, it is reasonable to interpret their lack of differential encoding times as a reflection of the elevated encoding effort they put into organizing the entire displays. Five-month-olds, on the other hand, may approach the habituation task differently. Rather than encoding every available parameter of the

stimulus, 5-month olds may look only long enough to determine whether or not it has been presented previously (Bronson, 1982) in accord with one or a few key features. Thus, though five-month-olds may be sensitive to phase relations in general, perturbations to the characteristic phase relations of human gait may constitute a meaningful difference only in the context of the familiar form of a person.

Differences in what is encoded by children of different ages or experience has been suggested as an explanation for other developmental phenomena, such as skill at making balance-beam discriminations (Siegler, 1989). Our results support a unified view of some forms of perception with these "higher" cognitive phenomena, in which differential skill in what is encoded is a ubiquitous characteristic of human learning and development.

More specific to perception, our results may help to fill out the way in which a computational account of the extraction of structure from biomechanical motion depends upon both general and domain-specific processing heuristics. A number of computational models are able to use general heuristics to organize broad classes of motion-given information by recovering the spatial or geometric properties of the form (Ullman 1983, Webb & Aggarwal, 1982). However, these general algorithms do not yield complete or unique solutions for biomechanical displays nor do they exhibit the robustness of human visual processes (see Proffitt & Bertenthal, 1988, for a discussion of these models). Some integration of a general heuristic account with a theory of schema-driven processing is needed.

Interestingly, one of the early models of habituation was both schema-based, and implied expectation-based differential encoding. Sokolov (1960) proposed a theory of habituation in which the organism builds a model of the stimulus and then tests this model against the stimulus. In Sokolov's view, habituation results from the repression of an orienting response when the perceptual system determines that there is no new information to be gathered from the stimulus. As the organism comes to expect one hypothesis more strongly than another, the value of additional information is reduced significantly, thus resulting in a decline of looking time. Interpreted computationally, Sokolov's theory is very close to computational methods in model-based diagnosis and in theory-formation methods (see the papers in Davis & Hamscher, 1990, and Shrager & Langley, 1991).

Neither Siegler's theory, nor the models of Sokolov, of Ullman, or of Webb & Aggarwal are complete alone; and their integration requires empirical guidance. Some of this guidance may be usefully de-

rived from developmental data such as that provided here. The present work may thus help us sharpen the view of perception as problem-solving (Rock, 1983), by guiding the integration of theory-formation models with general heuristic or constraint-based models, into a more complete account of general perception.

### Acknowledgments

We would like to thank Margaret Timmins, Sharon Lambert, and Elizabeth Ong for their assistance with this project. We would also like to thank the reviewers for their helpful comments. Please direct correspondence regarding this paper to the second author at System Sciences Laboratory, Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304.

### References

- Bertenthal, B.I. 1992. Infants' perception of biomechanical motions: Intrinsic image and knowledge-based constraints. In C. Granrud (Ed.), *Carnegie Symposium on cognition: Visual perception and cognition in infancy*. Hillsdale, NJ: Erlbaum. Forthcoming.
- Bertenthal, B.I. & Davis, P. 1988. Dynamical pattern analysis predicts recognition and discrimination of biomechanical motions. *Proceedings of the Annual Meeting of the Psychonomic Society*. Chicago, Illinois.
- Bertenthal, B.I., & Kramer, S.J. 1984. The TMS 9918a VDP: A new device for generating moving displays on a microcomputer. *Behavior Research Methods, Instruments, & Computers* 16:388-394.
- Bornstein, M.H. & Krinsky, S.J. 1985. Perception of symmetry in infancy: The salience of vertical symmetry and the perception of pattern wholes, *Journal of Experimental Child Psychology* 39:1-19.
- Bronson, G.W. 1982. *The Scanning Patterns of Human Infants: Implications for visual learning*. New Jersey: Ablex Publishing Co.
- Cohen, L.B. 1973. A two process model of infant visual attention. *Merrill-Palmer Quarterly* 19:157-180.
- Cutting, J.C. 1978. Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception* 7:393-405.
- Davis, R. & Hamscher, W.C. 1990. Model-based reasoning: Troubleshooting. In P.H. Winston & S.A. Shellard (Eds.), *Artificial Intelligence at MIT: Expanding Frontiers*. Cambridge, MA: MIT Press, 380-429.
- Gibson, J.J. 1950. *Perception of the visual world*. New York: Houghton Mifflin.
- Johansson, G. 1973. Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics* 142:201-211.
- Johansson, G. 1975. Visual motion perception. *Scientific American* 232(6):76-88.
- Johnson, M.H., Posner, M.I., & Rothbart, M.K. 1991. Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience* 3 (4):335-344.
- Pinto, J. & Davis, P. 1991. The categorical perception of human gait in 3- & 5- month old infants. *Presented at the biannual meeting of the Society for Research in Child Development*, Seattle, WA.
- Proffitt, D.R. & Bertenthal, B.I. 1988. Recovering connectivity from moving point-light displays. In W.N. Martin, & J.K. Aggarwal (Eds.), *Motion Understanding: Robot and Human Vision*. Boston: Kluwer Academic Publishers.
- Rock, I. 1983. *The logic of perception*. Cambridge, MA: The MIT Press.
- Schmidt, R.A. 1985. The search for invariance in skilled movement behavior. *Research Quarterly for Exercise and Sport* 56(2):188-200.
- Shrager, J. & Langley, P. 1991. *Computational Models of Scientific Discovery and Theory Formation*. San Mateo, CA: Morgan Kaufmann.
- Siegler, R. S. 1976. Three aspects of cognitive development. *Cognitive Psychology* 8:481-520.
- Sokolov, E.N. 1960. Neuronal models and the orienting influence. In M.A.B. Brazier (Ed.), *The central nervous system and behavior: III* New York: Macy Foundation.
- Ullman, S. 1983. Recent computational studies in the interpretation of structure from motion. In J.Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and Machine Vision*. New York: Academic Press.
- Wallach, H. & O'Connell, D.N. 1953. The kinetic depth effect. *Journal of Experimental Psychology* 45(4):205-217.
- Webb, J.A. & Aggarwal, J.K. 1982. Structure from motion of rigid and jointed global forms. *Artificial Intelligence* 19:107-130.