

# Coordinating Multiple Sensory Modalities While Learning to Reach

Matthew Schlesinger<sup>1</sup> & Domenico Parisi<sup>2</sup>

<sup>1</sup>Psychology Department  
Southern Illinois University at Carbondale, Illinois  
matthews@siu.edu

<sup>2</sup>Institute of Psychology  
National Research Council, Rome, Italy  
parisi@www.ip.rm.cnr.it

## Abstract

By the onset of reaching, young infants are already able to coordinate vision of a target with the felt position of their arm [7]. How is this coordination achieved? In order to investigate the hypothesis that infants learn to link vision and proprioception via the sense of touch, we implemented a recent computational model of reaching [22]. The model employs a genetic algorithm as a proxy for sensorimotor development in young infants. The three principal findings of our simulations were that tactile perception: (1) facilitates learning to coordinate vision and proprioception, (2) promotes an efficient reaching strategy, and (3) accelerates the remapping of vision and proprioception after perturbation of the multimodal map. Follow-up analyses of the model provide additional support for our hypothesis, and suggest that touch helps to coordinate vision and proprioception by providing a third, correlated information channel.

## 1. Introduction

Until recently, the prevailing view among motor development researchers was that infants develop hand-eye coordination by visually guiding their hand as they learn to reach [5, 6, 18, 21, 24]. According to this view, infants learn to coordinate the felt position of their arm (via proprioception) with the seen position of a target object by watching their hand as it approaches the target. A number of recent studies, however, suggest that infants do not rely on visual feedback from the hand as a coordination strategy during early reaching. For example, 5-month-old infants rely on ballistic rather than visually-guided reaching strategies when targets are displaced during a reach [1]. In addition, 6-month-olds produce comparable kinematic reaching patterns regardless of whether or not visual feedback of the hand is available [8].

Perhaps the strongest evidence against the “visual-guidance” hypothesis is offered by Clifton, Muir, Ashmead, and Clarkson [7], who found that at the onset of

reaching (around age 12 weeks), infants are just as likely to begin reaching for a target in a fully-lit room as they are to reach for a luminous target in a dark room. These findings are important in several respects. First, they suggest that infants do not learn to reach by visually guiding their hand toward a seen target. Second, they also reveal an early intermodal coordination between sight of the target and felt position of the arm, which is presumably not achieved through visual feedback from the hand during reaching. How then are vision and proprioception coordinated in young infants?

One possibility is that infants possess an innate multimodal map that computes a two-way transformation between the visual and proprioceptive spaces. Support for this hypothesis comes from studies of newborn infants, who show a tendency to generate reflex-like preparatory reaching movements (“prereaches”; see [9, 10, 12]). An alternative hypothesis is that infants learn to coordinate vision and proprioception during the development of reaching. According to this second account, learning to reach creates a task-specific context for infants to coordinate multiple information streams.

In this paper we investigate the second possibility, namely, that visual-proprioceptive coordination emerges as infants learn to reach. In particular, we focus on the role of tactile feedback from the target as a third information source that helps link vision and proprioception. The rest of the paper is divided into four sections. In the following section, we consider visual and tactile feedback as sources of sensory information available to infants as they learn to coordinate vision and proprioception. Next, we briefly describe a recent model of sensorimotor development in young infants [22] that we use to simulate early reaching. We then present a series of simulation studies that contrast the roles of visual and tactile feedback while learning to reach.

## **2. Linking Vision and Proprioception**

If infants are not born with an intermodal map linking vision and proprioception, then how does this coordination develop? We propose that infants learn to “triangulate” the position of their hand by combining information from the target and the arm with a third information source. In principle, at least two sensory channels may provide this third source of information. One possibility is that infants use the sight of their hand to help determine its position relative to a target. The advantage of using visual feedback from the hand to coordinate eye and arm movements is that it is continuously available during a reach. However, vision of the hand is potentially ambiguous (e.g., the hand may occlude the target without touching it), and may require additional computation to determine the hand’s position relative to the target. In addition, monitoring of the hand position also requires attentional resources above and beyond those available for keeping track of the target and controlling arm movements. These reasons may help explain why young infants do not appear to use visual feedback from the hand as an initial reaching strategy.

An alternative information source comes from tactile feedback during contact with the target. Unlike vision of the hand, tactile feedback is unambiguous. That is,

once contact is made with the target, the hand's location can be equated with the seen position of the target. Thus, tactile feedback provides a unique sensory signal for associating seen locations in space (i.e., the target) with felt positions of the arm. However, compared to vision of the hand, which provides continuous feedback during a reach, tactile feedback is only available *after* a successful reach. In addition, because this coordination begins to develop *prior* to the onset of consistent reaching, only infrequent or sporadic sensory updates (after successful reaches) are available for coordinating sight of the target with felt position of the arm.

### 3. A Model of Reaching in Young Infants

The Schlesinger et al. model [22] is designed to represent two core features of infant sensorimotor development: (1) learning by trial-and-error exploratory movements [3, 11, 23], and (2) use of coarse feedback for producing subsequent movements (i.e., the success or failure of a reach). Accordingly, the model implements an unsupervised, variation-and-selection learning algorithm analogous to evolutionary learning [13] as a stochastic optimization procedure. Specifically, a population of artificial neural networks is used to represent the diverse set of reaching movements available to infants. Because the networks inhabit a simple ecology (including their bodies) we refer to them as “econets” [20]. Like the movements generated by infants, econets adapt under selective pressure to produce the most direct, efficient reaching patterns.

It is important to stress that the model simplifies many of the details of sensorimotor control systems, and should not be interpreted as a biologically plausible model of sensorimotor development in infants. Rather, our principal goal is to use the model as a framework for abstracting out the most relevant aspects of multimodal sensory control, and simulating the developmental relations among these factors.

The model replicates several key aspects of reaching development in infants (see [22] for a complete description), including (1) limited use of redundant joints [4]; (2) a predominance of muscle co-contractions during early reaching [4, 23]; (3) the early emergence of stereotyped reaching movements [4, 15]; and (4) an early suppression of pre-reaches [11]. We present here a brief overview of the Schlesinger et al. model. Interested readers may refer to [22] for a full description, including the details of the learning algorithm and a complete list of all major parameter values.

#### 3.1 The econet

Figure 1a presents a schematic diagram of the econet in its two-dimensional reaching workspace. The econet remains at the center of the 100 x 100 unit workspace. The target is a small object (triangle), placed randomly in the workspace at the start of each trial. The econet has a monocular visual system with a 64° visual field (see Figure 1a). The econet also has a two-segment arm that spans the workspace. The trunk, head, and eye are positioned along the same axis, and so move as a single unit with the rotation of the eye. As Figure 1a illustrates, the eye

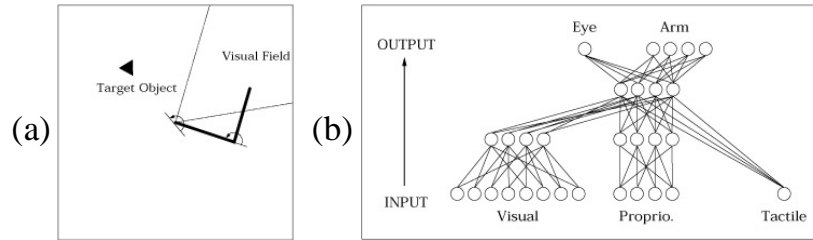


Figure 1: (1a) the econet in its 2D workspace; the left and right edges of the visual field are shown, as well as the rotational limits of the shoulder and elbow joints. (1b) the 4-layer, feedforward neural network used to control the movements of the econet (not all connections shown).

can rotate  $360^\circ$  in either direction while the arm's movement is limited to  $180^\circ$  rotation at the shoulder and elbow joints. A multi-layer feedforward neural network (see Figure 1b) uses visual input from the eye, proprioceptive input from the arm, and tactile input from the hand to control the movements of the eye and arm. Note that the visual and proprioceptive sensory inputs first pass through an intramodal hidden layer before combining at the second intermodal hidden layer. This partially modularized architecture helps minimize crosstalk or interference effects between sensory modalities, and results in improved performance [14, 22].

Each visual input unit spans  $8^\circ$  of the visual field, and encodes the presence of either the target or hand (scaled from 0 to 1 as a function of the distance of the object from the econet). The proprioceptive input units encode the amount of stretch in the arm muscles, represented by a linear mapping from the shoulder and elbow joint angles. Finally, the tactile input unit encodes contact of the hand with the target as a binary signal. Movements of the econet's eye and arm are produced by activation on its output units. A single unit moves the eye either left or right, while the remaining four units activate the flexor-extensor muscle pairs of the upper arm and forearm.

### 3.1 The learning algorithm

We employ a genetic algorithm as a proxy for trial-and-error sensorimotor development. An initial generation of econets is produced, with random connection weights. Each econet is given 20 trials (of 100 timesteps) to reach a nearby object; fitness is increased by 1 point for each timestep spent in contact with the target (for a theoretical maximum of 2000 fitness points). The top 20 econets of each generation are selected and used to produce 100 new econets (5 offspring per parent). During the reproduction process, the connection weights of the parent are copied with a 2% chance of random mutations. An 8-bit binary string is used to represent each connection weight; mutations are accomplished by switching the randomly selected bit from 1 to 0 or vice versa. Note that while the connection weights mutate from parent to offspring, all other characteristics (e.g., the neural network architecture,

body parameters, etc.) remain constant across generations. Also, no other variation mechanisms (sexual reproduction, crossover) are employed.

## 4. Contrasting Visual and Tactile Feedback

We now present a series of simulation studies that allow us to contrast the relative roles of visual and tactile feedback while learning to reach. In the first set of simulations, we illustrate how tactile feedback facilitates learning to reach. The second set demonstrates the reaching strategies that emerge when either visual or tactile feedback is available. In the final set, we investigate how the model responds to a perturbation of the visual-proprioceptive coordination.

### 4.1 Touch facilitates learning to reach

Figure 2a presents the average fitness (i.e., time spent in contact with the target) in two populations of econets, over 10 simulation runs. (In all analyses and figures, we adopt the convention of presenting the results of the 20 most fit individuals from each generation.) Both populations receive visual and proprioceptive sensory inputs. In the “tactile-on” population, sensory input is also received from the tactile input unit (when the target is contacted); while no tactile inputs are received in the “tactile-off” population. Figure 2b presents the average number of trials (out of 20) in which the target was reached. As the figure illustrates, tactile feedback not only results in reaching the target more often, but also in maintaining contact with it for more time.

Can we conclude that tactile feedback leads to better coordination of vision and proprioception? One counterargument is that because the tactile-on econets receive both tactile *and* visual feedback about the location of their hand, it is the *combination* of the two forms of feedback that lead to better intermodal coordination. To test this possibility, we simulated a variant of the tactile-on population that could not see their hands (tactile-on-luminous). This condition corresponds to an infant that learns to reach for luminous objects in a dark room, and

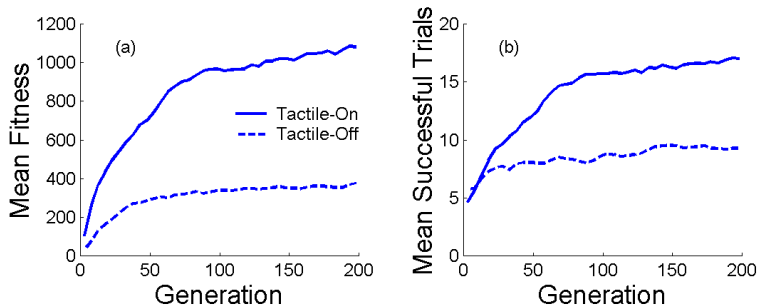


Figure 2: Average fitness (2a) and number of successful trials (2b) in two populations of econets. “Tactile-on” econets receive tactile feedback when they touch the target, while “tactile-off” econets receive no tactile feedback.

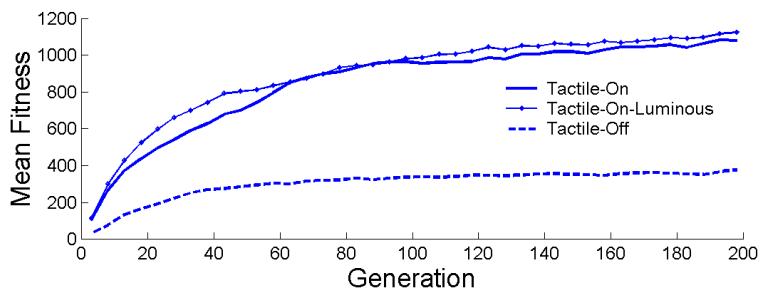


Figure 3: Average fitness in the tactile-on, tactile-off, and a third, tactile-on-luminous population. Tactile-on-luminous econets can see the target, but their hand is invisible.

allows us to compare reaching when only tactile feedback, or both tactile and visual feedback are available.

Figure 3 presents the results of the previous analysis, along with the average fitness in the tactile-on-luminous population. There are two important findings. First, in both of the tactile-on populations, performance and rate of learning are near equal. Thus, visibility of the hand does not appear to play a major role in coordinating vision and proprioception when tactile feedback is available. Second, this result also parallels the findings of Clifton et al. [7], who observed similar onset ages for reaching in normal and luminous-object conditions.

A second question concerns the fact that the tactile sensory signal not only indicates when the hand is touching the target, but also when the econet’s fitness changes. (Recall that touching the target increases fitness.) Thus, touch is confounded with fitness. However, we can separate these two factors by asking what happens when touching the target is no longer rewarded, but remaining close to it is (i.e., increased fitness for “hovering” near the target). Figure 4 presents the average fitness for the tactile-on and tactile-off populations in the “hover” condition (i.e., 0 fitness when touching the target, +1 for remaining near). While the advantage of the

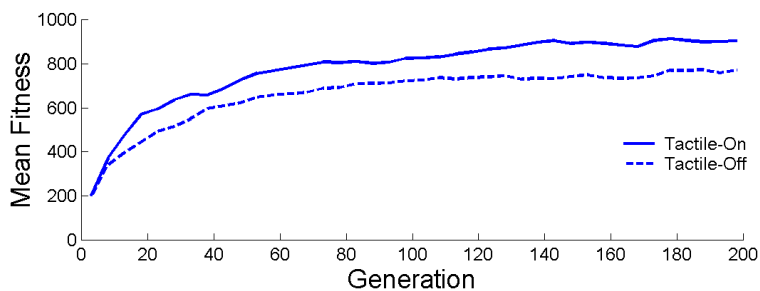


Figure 4: Average fitness in the “hover” condition, where econets are rewarded for keeping their hand near (but not on) the target.

tactile-on population is diminished (compare to Figure 2a), reaching is still significantly better with tactile feedback.

#### 4.2 Touch promotes an efficient reaching strategy

We can explore the coordination between vision and proprioception more closely by analyzing the specific reaching strategies employed by the tactile-on and tactile-off populations. Reaching the target requires first determining the positions of the target and the hand, and then moving the hand toward the target. We might expect different sequencing strategies to emerge depending on how well vision and proprioception are coordinated. Specifically, if they are well-coordinated, econets should employ a “target-first” strategy in which the target is located, and then the hand is brought to it as the target is fixated. However, the target-first strategy depends on knowing the position of the hand, especially when it is outside the visual field. If vision and proprioception are poorly coordinated, we should instead expect a “hand-first” strategy, where econets first bring their hand into the visual field and then search for the target.

Figure 5 presents the relative fixation time of the hand and the target in the tactile-on and tactile-off populations. The tactile-on population employs a target-first reaching strategy: the majority of time is spent fixating the object (59%), while the hand is only fixated near the end of the reach (41%). In contrast, tactile-off econets find the hand, and then visually guide it toward the target: without tactile feedback, the hand is fixated more than the target (62% and 38%, respectively).

#### 4.3 Touch accelerates intermodal recalibration

For the final set of analyses, we ask how econets adapt to a perturbation of the visual-proprioceptive coordination. In particular, we create an intermodal conflict by simulating a 10° prismatic displacement of the visual field. Thus, while touch and

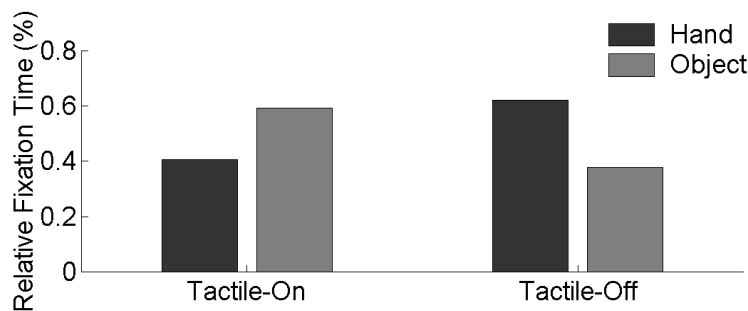


Figure 5: Relative time spent fixating the hand or the object, in the tactile-on and tactile-off populations. Tactile-on econets find the target first, and bring the hand toward it, while tactile-off econets find the target first, and then search for the target.

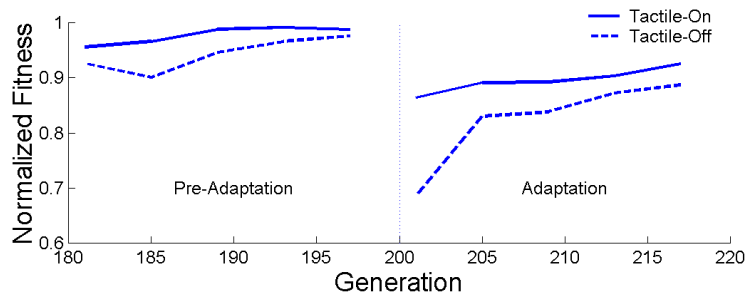


Figure 6: Normalized fitness in the tactile-on and tactile-off populations, during the final 20 generations of training (“Pre-Adaptation”) and for 20 generations of simulated prismatic displacement of visual input (“Adaptation”).

proprioception remain veridical, the visual sensory system now encodes objects in the visual field as  $10^\circ$  to the left of their actual position. This condition is analogous to prismatic adaptation studies with infants, who quickly learn to accommodate their reaching movements while wearing displacement prisms [18, 19].

Prismatic adaptation is simulated by first training the tactile-on and tactile-off econets as before (i.e., for 200 generations). After the initial training phase (“Pre-Adaptation”), prismatic displacement is simulated by shifting the apparent position of all objects in the visual field  $10^\circ$  to the left. Training during the adaptation phase then continues for an additional 20 generations (“Adaptation”). Figure 6 presents the mean normalized fitness in the tactile-on and tactile-off populations, during the pre-adaptation and adaptation phases. There are two major findings. First, prismatic displacement has a larger initial effect on the tactile-off population. Second, the tactile-on population recovers a higher level of performance during the adaptation phase.

## 5. Conclusion

Our simulation results provide support for the hypothesis that tactile feedback facilitates learning to reach. In particular, when both visual and tactile feedback are available, econets tend to rely on tactile feedback. Indeed, visually-guided reaching strategies only emerge when tactile feedback from the hand is unavailable (e.g., tactile-off econets). We also found that like human infants, econets do not need to see their hand in order to learn to coordinate vision of a target with felt position of their arm (e.g., the tactile-on-luminous condition). This result not only replicates the findings of Clifton et al. [7], but also helps us understand how infants may begin to coordinate vision and proprioception *before* the onset of consistent, successful reaching.

In considering these results, however, three important issues should be raised. First, the current findings may depend in part on the specific choice of network architecture and learning algorithm (however, see [22] for other a comparison with

other architectures). It is critical to evaluate the hypothesis that tactile feedback facilitates coordination of vision and proprioception using other network architectures and unsupervised learning algorithms (e.g., reinforcement learning).

Second, the visual-proprioceptive coordination achieved by the econet model is task-specific. Rather than learning a general intermodal mapping that links vision and proprioception, a specific set of behavioral strategies are discovered and employed to solve a particular task. Other modeling approaches may shed light on how a general intermodal coordination might develop (e.g., see [16, 17] for a simulation based on infants' spontaneous arm movements).

A third question concerns how our simulation findings should be reconciled with later stages of reaching development during infancy, as well as with reaching behaviors in adults. While infants do not attend to their hand as they learn to reach, visual feedback will come to play a critical role as infants begin to attempt high-precision reaches (e.g., pincer grasp of a small object). In addition, visual feedback plays a central role for adults in recalibrating the visual-proprioceptive map after perturbation [2]. One way these findings may be integrated is by proposing a reliance on tactile feedback during early infancy, when visual perception (e.g., visual acuity, depth perception, visual tracking, etc.) may be no better at locating the hand than proprioception. As visual perception improves and stabilizes, however, infants should transition from the use of tactile feedback to visual feedback [7].

## References

1. Ashmead, D.H., McCarty, M.E., Lucas, L.S., & Belvedere, M.C. (1993). Visual guidance in infants' reaching toward suddenly displaced targets. *Child Development*, *64*, 1111-1127.
2. Beers, R.J. van, Sittig, A.C., & Gon J.J. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology*, *81*, 1355-1364.
3. Berthier, N.E. (1996). Learning to reach: A mathematical model. *Developmental Psychology*, *32*, 811-823.
4. Berthier, N.E., Clifton, R.K., McCall, D.D., & Robin, D.J. (1999). Proximodistal structure of initial reaching in human infants. *Experimental Brain Research*, *127*, 259-269.
5. Bower, T.G.R. (1974). *Development in infancy*. San Francisco: Freeman.
6. Bushnell, E.W. (1985). The decline of visually guided reaching during infancy. *Infant Behavior and Development*, *8*, 139-155.
7. Clifton, R.K., Muir, D.W., Ashmead, D.H., & Clarkson, M.G. (1993). Is visually guided reaching in early infancy a myth? *Child Development*, *64*, 1099-1110.
8. Clifton, R.K., Rochat, P., Robin, D.J., & Berthier, N.E. (1994). Multimodal perception in the control of infant reaching. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 876-886.
9. Ennouri, K., & Bloch, H. (1996). Visual control of hand approach movements in new-borns. *British Journal of Developmental Psychology*, *14*, 327-338.

10. Hofsten, C. von. (1982). Eye-hand coordination in the newborn. *Developmental Psychology, 18*, 450-461.
11. Hofsten, C. von. (1984). Developmental changes in the organization of prereaching movements. *Developmental Psychology, 20*, 378-388.
12. Hofsten, C. von, & Ronnqvist, L. (1993). The structuring of neonatal arm movements. *Child Development, 64*, 1-46-1057.
13. Holland, J.H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, MI: University of Michigan Press.
14. Jacobs, R.A., Jordan, M.I., & Barto, A.G. (1991). Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. *Cognitive Science, 15*, 219-250.
15. Konczak, J., & Dichgans, J. (1997). The development toward stereotypic arm kinematics during reaching in the first 3 years of life. *Experimental Brain Research, 117*, 346-354.
16. Kuperstein, M. (1988). Neural model of adaptive hand-eye coordination for single postures. *Science, 239*, 1308-1311.
17. Kuperstein, M. (1991). INFANT neural controller for adaptive sensorimotor coordination. *Neural Networks, 4*, 131-145.
18. McDonnell, P.M. (1975). The development of visually guided reaching. *Perception and Psychophysics, 19*, 181-185.
19. McDonnell, P.M., & Abraham, W.C. (1979). Adaptation to displacing prisms in human infants. *Perception, 8*, 175-185.
20. Parisi, D., Cecconi, F., & Nolfi, S. (1990). Econets: Neural networks that learn in an environment. *Network, 1*, 149-168.
21. Piaget, J. (1952). *The origins of intelligence in children*. New York: Basic Books.
22. Schlesinger, M., Parisi, D., & Langer, J. (2000). Learning to reach by constraining the movement search space. *Developmental Science, 3*, 67-80.
23. Thelen, E., Corbetta, D., Kamm, K., Spencer, J.P., Schneider, K., & Zernicke, R.F. (1993). The transition to reaching: Mapping intention and intrinsic dynamics. *Child Development, 64*, 1058-1098.
24. White, B.L., Castle, P., & Held, R. (1964). Observations on the development of visually directed reaching. *Child Development, 35*, 349-364.
25. Zernicke, R.F., & Schneider, K. (1993). Biomechanics and developmental motor control. *Child Development, 64*, 982-1004.