

# Evolving Imitating Agents and the Emergence of a Neural Mirror System

Elhanan Borenstein<sup>1</sup> and Eytan Ruppin<sup>1,2</sup>

<sup>1</sup>School of Computer Science, Tel Aviv University, Tel-Aviv 69978, Israel

<sup>2</sup>School of Medicine, Tel Aviv University, Tel-Aviv 69978, Israel

borens@post.tau.ac.il

## Abstract

Imitation is a highly complex cognitive process, employing vision, perception, representation, memory and motor control. The underlying mechanisms that give rise to imitative behavior have attracted a lot of attention in recent years and have been the subject of research in various disciplines, from neuroscience to animal behavior and human psychology. In particular, studies in monkeys and humans have discovered a neural mirror system that demonstrates an internal correlation between the representations of perceptual and motor functionalities. In contradistinction to previous engineering-based approaches, we focus on the evolutionary origins of imitation and present a novel framework for studying the *emergence* of imitative behavior. We successfully develop evolutionary adaptive autonomous agents that spontaneously demonstrate imitative learning, facilitating a comprehensive study of the emerging underlying neural mechanisms. Interestingly, some of these agents are found to embody a neural “mirror” device analogous to those identified in biological systems. Further analysis of these agents’ networks reveals complex dynamics, combining innate perceptual-motor coupling with acquired context-action associations, to accomplish the required task.

## Introduction

Imitation is an effective and robust way to learn new traits by utilizing the knowledge already possessed by others. The past twenty years have seen a renewed interest in imitation in various fields of research such as developmental psychology, experimental studies of adult social cognition, and most important, neurophysiology and neuropsychology (Prinz and Meltzoff, 2002). Research in this last field had led to the exciting discovery of *mirror neurons*. These neurons, found in the ventral premotor cortex (area F5) in monkeys, discharge both when the monkey performs an action and when it observes another individual making a similar action (Gallese et al., 1996; Rizzolatti et al., 2001). An analogous mechanism, whereby cortical motor regions are activated during movement observations was also demonstrated in humans using TMS, MEG, EEG and fMRI (e.g. Iacoboni et al., 1999). Imitation of motor skills requires the capacity to match between the visual perception of a demonstrator’s action and the execution of a motor command. The neural mirror system, demonstrating an internal correlation between

the representations of perceptual and motor functionalities, may form one of the underlying mechanisms of imitative ability.

Learning by imitation has already been applied by researchers in the fields of artificial intelligence and robotics in various experiments. Hayes and Demiris (1994) presented a model of imitative learning to develop a robot controller. Billard and Dautenhahn (1999) studied the benefits of social interactions and imitative behavior for grounding and use of communication in autonomous robotic agents. Borenstein and Ruppin (2003) employed learning by imitation to enhance the evolutionary process of autonomous agents. For an up-to-date introduction to work on imitation in both animals and artifacts see the cross-disciplinary collection (Dautenhahn and Nehaniv, 2002b). Furthermore, some researchers, motivated by the recent discovery of a neural mirror system, have implemented various models for imitative learning, embodying neurophysiologically inspired mechanisms. Billard (2000) presented a model of a biologically inspired connectionist architecture for learning motor skills by imitation. The architecture was validated through a mechanical simulation of two humanoid avatars, learning several types of movements sequences. Oztop and Arbib (2002), focusing on the grasp-related mirror system, argued that mirror neurons first evolved to provide visual feedback on one’s own “handstate” and were later generalized to understanding the actions of others. They have conducted a range of simulation experiments, based on a schema design implementation of that system, providing both a high-level view of the mirror system and interesting predictions for future neurophysiological testing. Other researchers (Marom et al., 2002) claimed that the mirror system structure can be acquired during life through interaction with the physical or social environment and demonstrated models whereby perceptual and motor associations are built up from experience during a learning phase.

The studies cited above, however, assume that the agents’ basic ability and incentive to imitate are innate, explicitly introducing the underlying functionality, structure or dynamics of the imitation mechanism into the experimental system.

**In contrast to this engineering-based approach, we wish to study the neuronal mechanisms and processes underlying imitation from an evolutionary standpoint, and to demonstrate how imitative learning *per se* can spontaneously emerge and prevail.** Clearly, acknowledging the evolutionary origins of imitation and examining the emerging (rather than engineered) imitative learning device can shed new light on the common fundamental principles that give rise to imitative behavior.

In this study, we thus set out to pursue two objectives: Acknowledging the significance of embodied imitation, **we first present a novel experimental framework for evolving context-based imitative learning in evolutionary adaptive autonomous agents** (Ruppin, 2002; Floreano and Urzelai, 2000). We demonstrate the emergence of imitating agents that embody a simple, yet biologically plausible mechanism of imitative behavior. **We then turn to systematically analyze the structure and dynamics of the resulting neurocontrollers.** This analysis surprisingly reveals neural devices analogous to those found in biological systems, including clear examples of internal coupling between observed and executed actions. Further analysis of the network adaptation dynamics demonstrates the innate nature of these internal links with direct bearing on one of the key questions in imitation theory, concerning the ontogeny of mirror neurons (Prinz and Meltzoff, 2002; Hurford, 2003). We conclude with a discussion of the implications of our findings for imitation theory and a description of future work.

## Context-Based Imitation

Learning by imitation, like any cognitive process, must be considered an intrinsically embodied process, wherein the interaction between the neural system, the body and the environment cannot be ignored (Keijzer, 2002; Dautenhahn and Nehaniv, 2002a). In particular, every action, either observed or performed, occurs within a certain *context*. A context can represent the time or place in which the action is made, various properties of the environment, or the state of the individual performing the action. Clearly, there is no sense in learning a novel behavior by imitating another’s actions if you do not know the context in which these actions are made – a certain action can be extremely beneficial in one context, but have no effect (or even be deleterious) in a different context. We hence use the term *context-based imitation* in the sense of being able to reproduce another’s observed action whenever the context in which the action was originally observed, recurs.<sup>1</sup> For example, an infant observ-

<sup>1</sup>Animal behavior and human psychology literature introduces a wide range of definitions of imitation, focusing on what can constitute true imitation vs. other forms of social learning (Billard and Dautenhahn, 1999; Zentall, 2001). Our definition addresses the importance of the observed action’s *context* for a successful *future* behavior.

ing his parents may learn by imitation to pick up the phone (*action*) whenever the phone is ringing (*context*).

Context-based imitation can thus be conceived as constructing a set of associations (or a mapping) from contexts to actions, based on observations of a demonstrator performing different actions within various contexts. These associations should comply with those that govern the demonstrator’s behavior, and should be learned (memorized) so that each context stimulates the production of the proper motor action even when the demonstrator is no longer visible. Such a learning scheme can be seen as an imitation-based analogue of a partially observable hidden Markov model of classical operant learning. It should be noted however, that “*action*” is an abstract notion, and in reality, an imitating individual (agent) should also be capable of matching a *visual perception* of the demonstrator’s action with the corresponding *motor command* that activates this action. The key objective of this study is to gain a comprehensive understanding of the mechanisms that govern such context-based imitative learning and in particular to examine the nature of the associations between visual perception, motor control and contexts that are being formed in the process. To address these questions, we employ an experimental setup that embodies context-based imitation within an evolutionary framework.

## The Experimental Setup

### The Environment

The agents in our simulation inhabit a world that can be in one of several *world states*  $\{s_1, s_2, \dots, s_n\}$ . In each time step, the world state is randomly selected from  $\{s_1, s_2, \dots, s_n\}$  with a uniform distribution. These states can represent for example the presence of certain food items or the size of an observed object and hence form the *context* in which actions are observed and performed. An additional set,  $\{a_1, a_2, \dots, a_m\}$ , represents the repertoire of motor *actions* that can be performed by the agent or by the demonstrator. Within the simulations described below, both  $n$  and  $m$  are set to 4. A *state-action mapping* is also defined, assigning a certain action as the proper action for each world state  $s_i$ . Regularly performing the proper action assigned to the current state of the world is deemed a successful behavior and confers a positive fitness. It is assumed that the environment is also inhabited by a demonstrator (teacher), successfully performing the proper action in each time step. However, the world state and demonstrator are not visible in every time step and can be seen with probabilities 0.6 and 0.2 respectively. **Furthermore, the above mapping, from world states to actions, is randomly selected anew in the beginning of each agent’s run in the world.** The motivation for this state-action mapping shuffle is twofold. First, it prevents such a mapping from becoming genetically determined. To demonstrate a successful behavior, agents must *learn* the proper mapping by observing the demonstrator, promot-

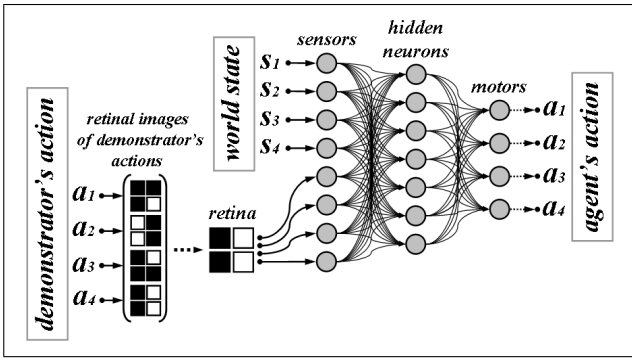


Figure 1: The agent’s sensorimotor system and neurocontroller. The sensory input is binary and includes the current world state and a retinal “image” of the demonstrator’s action (when visible). The retinal image for each possible demonstrator’s action and a retinal input example for action  $a_4$  are illustrated. The motor output determines which actions are executed by the agent. The network synapses are adaptive and their connection strength may change during life according to the specified learning rules.

ing an imitation based mechanism to emerge. Second, it represents a scenario of a changing environment, wherein novel world states appear over time (new food sources, other species, etc.), making prior state-action mappings obsolete.

## The Agent

Figure 1 illustrates the structure of the agent’s sensorimotor system and neurocontroller. The agent’s sensory input in each time step includes the current world state (if visible) and a 4-cell retinal “image” of the demonstrator’s action (if visible). The retinal image is determined according to a predefined mapping from actions to retinal binary patterns which remains fixed throughout the simulation. Each of the agent’s output neurons represents a motor command, determining which actions (if any) will be executed by the agent.

Each agent embodies a simple feed-forward neural network as a neurocontroller. These networks however are *adaptive*, whereby the genotype of each individual encodes not only the initial synaptic weights but also a *Hebbian learning rule* and *learning rate* for each synapse. In particular, each synapse in the network,  $(i, j)$ , connecting neuron  $j$  to neuron  $i$ , is encoded by 4 genes, defining the following properties: (i)  $w_{ij}^0$  - the initial connection strength of the synapse (real value in the range  $[0,1]$ ); (ii)  $s_{ij}$  - the connection sign (1 or -1); (iii)  $\eta_{ij}$  - the learning rate (real value in the range  $[0,1]$ ); and (iv)  $\Delta w_{ij}$  - the learning rule applied to this synapse.  $\Delta w_{ij}$  encodes 1 of 5 learning (modification) rules: no learning, plain Hebb rule, postsynaptic rule, presynaptic rule and covariance rule. Each synaptic weight  $w_{ij}$  is initialized with  $w_{ij}^0$  at the beginning of the agent’s life and is updated after every time step (a sensory-motor cycle) ac-

ording to  $w_{ij}^t = w_{ij}^{t-1} + \eta_{ij} \Delta w_{ij}$ . For a detailed description of the adaptation dynamics see Floreano and Urzelai (2000). The network topology is static throughout the process and for the purpose of our simulation was set to 8-7-4 (i.e., 8 input neurons, a hidden layer with 7 neurons, and 4 output neurons), with an additional threshold unit in each layer. Such *evolutionary adaptive autonomous agents*, inspired by those presented in Todd and Miller (1991) and Floreano and Urzelai (2000), demonstrate a learning process that is supervised only indirectly, through natural selection.

## The Evolutionary Process

A haploid population of the agents described above evolve to successfully behave in the environment. Each agent lives in the world for 500 time steps. Fitness is evaluated according to the agent’s success in performing the proper action assigned to the current world state (i.e. activating only the appropriate motor neuron), according to the state-action mapping, in each time step. An agent should perform an action only if the world state is visible and regardless of the demonstrator’s visibility. We use the Mean-Square Error (MSE) measure to calculate the distance between the agent’s motor output and the desired output, averaged over the agent’s life. The agent performance during the first 100 time steps is not evaluated (infancy phase). Fitness value is then calculated as  $(1 - Error)$  and averaged over 20 trial runs in the world.

The initial population is composed of 200 individuals, each assigned a randomly selected genome (encoding the initial connection weights, learning rules and learning rates). Each new generation is created by randomly selecting the best agents from the previous generation according to their fitness, and allowing them to reproduce. During reproduction, 2% of the genes are mutated. The genomes of the top 20% of individuals are copied to the next generation without mutation.

## Results

Figure 2 portrays the fitness of the evolving agents across evolution. As evident, the evolved agents successfully master the behavioral task, regularly executing the proper action in each world state. Obviously, given the way the task is designed, this would not have been possible in the absence of an emerged imitation-based learning strategy.

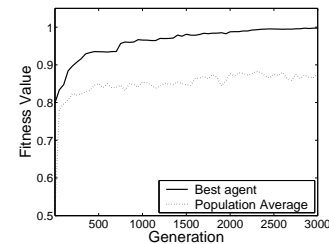


Figure 2: The fitness of the best agent in the population and the population average fitness as a function of generation.

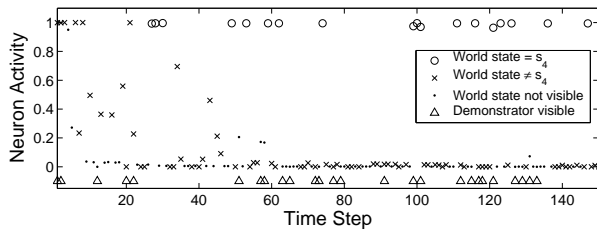


Figure 3: The activation level of one motor neuron ( $m_2$ ) during the first 150 time steps. The different shapes indicate whether the world state was  $s_4$  and whether it was visible. The triangles at the bottom further represent time steps in which the demonstrator was visible.

Having successfully evolved imitating agents, we turned to examine the structure, dynamics and neural mechanisms that these agents embody. In the rest of this section, we analyze one such successful agent – the best agent in the last generation of a specific evolutionary simulation run. Other successful agents, from various simulation runs, were analyzed and demonstrated similar phenomena (not shown here). Direct evidence of the agent’s successful imitative behavior and the resulting learning dynamics are demonstrated in Figure 3, depicting the activity of one of the motor neurons ( $m_2$ ) in different states of the world. In this specific simulation run, the state-action mapping was arbitrarily set so that  $a_2$  is the proper action in world state  $s_4$  and not in any other state. In the beginning of its life, the agent activates motor  $m_2$ , and therefore performs action  $a_2$ , whenever the world state is visible. However, after only a few demonstrations of the appropriate behavior, the proper state-action mapping is learned and this motor is activated only when the world state is  $s_4$ , as expected.

Furthermore, examining the network hidden layer reveals an interesting phenomenon with regard to the internal representation of actions. As stated above, to support imitative learning, wherein associations from contexts to motor commands should be inferred from observations of the demonstrator’s actions, an agent should be capable of matching the visual perception of an observed action to the motor command that generates the corresponding action. Figure 4, depicting the activation level of 3 hidden neurons, attests to the emergence of such inherent perceptual-motor coupling. Apparently, various neurons in the hidden layer are active both when the agent performs a certain action and when it observes the demonstrator making a similar action, **forming internal mirror neurons analogous to those found in biological systems.**<sup>2</sup> Such mirror neurons were found in most

<sup>2</sup>As seen in Figure 4, the activation level of mirror neurons during action observation is typically lower than the activation level during action execution. An analogous phenomenon can also be detected in neuronal recording data in the literature, and should be further investigated. However, in our simulation, the relatively small number of hidden neurons may account for this phenomenon, forcing mirror neurons to participate also in motor excitation.

of the agents that evolved in our simulation environment. However, typically, not all actions in the repertoire were associated with a corresponding mirror neuron, and there have been cases where successful agents did not seem to incorporate any identifiable mirror neurons. An additional set of intervention experiments, wherein hidden neurons are externally activated (stimulated) or inactivated (lesioned), was performed (not detailed here). These experiments demonstrated that even actions that could not be associated with a fully localized representation (i.e. a mirror neuron) may still be represented in the hidden layer through a distributed neuronal configuration. These findings are further discussed in the following section.

We finally turn to examine the ontogenic, developmental aspects of the resulting neurocontroller. Our main objective is to identify which components in the neural mechanism are innate and which are acquired during the agent’s life. We first determine which synapses play a significant role in the *learning* process. Clearly, variation in the synapse strength during life or the genetically coded learning rate are not appropriate indicators as they cannot differentiate between learning processes that genuinely adapt the agent to the world and unrelated self-organization processes. We thus measure the variance in the connection strength *at the end of the agent’s life across 1000 simulation runs*. A low

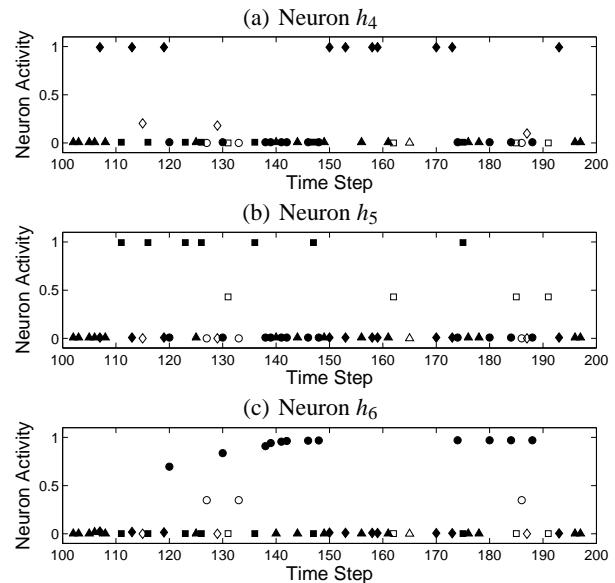


Figure 4: The activation level of 3 hidden neurons ( $h_4$ ,  $h_5$  and  $h_6$ ) during time steps 100-200 with an indication of the executed or observed action. Circles, squares, diamonds and triangles represent actions  $a_1$ ,  $a_2$ ,  $a_3$ ,  $a_4$  respectively. A filled shape indicates that the action was executed by the agent (stimulated by a visible world state), while an empty shape indicates that the action was only observed but not executed. Time steps where actions are both observed and executed are not drawn.

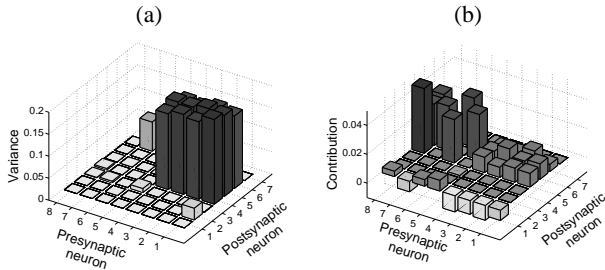


Figure 5: An illustration of the connection strength variance (a) and the overall contribution (b) of the synapses connecting the input layer to the hidden layer.

variance value indicates that the synapse dynamics are independent of the world characteristics (e.g. the state-action mapping), and thus cannot contribute to the learning process that adapt the agent to the world. As demonstrated in Figure 5a, this measure highlights the acquired nature of the synapses connecting the world state neurons (input neurons 1-4), with the mirror neurons we have identified (hidden neurons 4-6). **Clearly, the acquired state-action associations are induced by these synapses.** The markedly lower variance values in other synapses from this layer and in synapses connecting hidden layer neurons to motor neurons (not illustrated here), suggest that these synapses do not play an important part in the learning process. To measure the overall importance of each synapse to the agent’s behavior, we have utilized the Multi-perturbation Shapley value Analysis (MSA), a rigorous way to determine the contributions of system elements (Keinan et al., 2004). The resulting contribution of each synapse connecting the input layer to the hidden layer is illustrated in Figure 5b. Evidently, the synapses that have been identified above as participating in the learning process, possess a non-negligible contribution value. However, the most important synapses are among those connecting the retinal neurons (input neurons 5-8), representing the observed action, with the mirror neurons (hidden neurons 4-6). **These connections manifest the strong innate associations between the visual perception of observed actions and the internal representation of these actions, developed during the evolutionary process.**

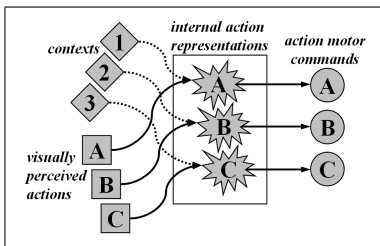


Figure 6: A simple model of context-based imitation. Solid arrows represents innate associations, while dashed arrows represents associations that are acquired during the agent’s life via Hebbian learning.

Based on the findings described above, a simple model of the mechanism that evolved in our settings to support imitative behavior can be inferred (Figure 6). Notably, the required perceptual-motor coupling was not explicitly engineered into the agents, but rather emerged through evolution as an *innate* property. Furthermore, to support an effective mechanism of imitation, visually perceived actions are linked to the corresponding motor commands via fully localized internal elements, representing each action, in the form of mirror neurons. The acquired context-action stimuli can then be constructed through a simple mechanism of Hebbian learning without external supervision or reinforcement signals. This model can account for simple, low-level forms of imitation that exhibit a clear example of innate perceptual-motor link such as infant facial imitation (Meltzoff, 1996).

## Discussion

This study presents an experimental framework for studying the emergence and dynamics of imitation in evolutionary autonomous agents. This framework provides a fully accessible, yet biologically plausible, distilled model for imitation and can serve as a vehicle to study the mechanisms that underlie imitation in biological systems. Our confidence in this framework is motivated by two observations: First, being an evolutionary emerging mechanism, rather than an engineered one, we believe it is likely to share the same fundamental principles driving natural systems. Second, our analysis of the resulting mechanism reveals phenomena analogous to those found biological neural mechanisms.

The model presented in this paper addresses the very essence of questions concerning the mechanism underlying imitative behavior. It successfully demonstrates how the required associations between perceived actions, motor commands and contexts can be constructed within a hybrid adaptation process, combining evolution and lifetime learning. In particular, addressing the ontogeny of mirror neurons, an issue which is currently in the center of imitation theory research, our model offers a simple schema for the origins and dynamics of the neural mirror system.

The mirror neurons that emerged in our simulation also provide interesting insights to the ongoing debate about the role of representation in embedded systems. Our model, promoting the use of observed actions of “others” for learning proper motor actions of “self”, suggests a hypothesis for the origins of internal representation. However, the emergence of fully localized internal representations was not absolute and various distributed representations were also demonstrated, supporting Cliff and Noble’s (1997) call for an operational definition of representation. The evolution of such distributed representations also confirms that the emergence of mirror neurons within our experimental setup is not trivial and we thus wish to use this model to further determine the physical and social environmental conditions that promote the emergence of localized mirror neurons.

The framework presented in this paper can be further enhanced to simulate a more realistic scenario of social learning. In particular, we wish to examine how an extension of the agent's sensory input, and a complex social environment inhabited by demonstrators with varying levels of success, affect the resulting imitation strategy. Questions concerning the dependencies between observed and executed actions and the formation of mirror neurons are especially of great interest: How will the representation of actions that cannot be executed by the observer (e.g. due to different embodiment) differ from those of imitated actions? How will a hierarchical repertoire of actions affect the emerging representation? Can emerging mirror neurons help predict the actions of others (Ramnani and Miall, 2004)? We hope that further extensions of this basic model will allow us to obtain testable predictions regarding imitative behavior in humans and primates, and shed new light on some of the key issues concerning perception, internal representation and cognition.

## References

- Billard, A. (2000). Learning motor skills by imitation: a biologically inspired robotic model. *Cybernetics & Systems*, 32:155–193.
- Billard, A. and Dautenhahn, K. (1999). Experiments in learning by imitation: grounding and use of communication in robotic agents. *Adaptive Behavior*, 7(3/4):411–434.
- Borenstein, E. and Ruppín, E. (2003). Enhancing autonomous agents evolution with learning by imitation. *Journal of Artificial Intelligence and the Simulation of Behavior*, 1(4):335–347.
- Cliff, D. and Noble, J. (1997). Knowledge-based vision and simple visual machines. *Philosophical Transactions of the Royal Society: Biological Sciences*, 352:1165–1175.
- Dautenhahn, K. and Nehaniv, C. (2002a). The agent-based perspective on imitation. In Dautenhahn, K. and Nehaniv, C., editors, *Imitation in Animals and Artifacts*. The MIT Press.
- Dautenhahn, K. and Nehaniv, C. (2002b). *Imitation in Animals and Artifacts*. MIT Press, Cambridge, Mass., USA.
- Floreano, D. and Urzelai, J. (2000). Evolutionary robots with on-line self-organization and behavioral fitness. *Neural Networks*, 13:431–443.
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119:593–609.
- Hayes, G. and Demiris, J. (1994). A robot controller using learning by imitation. In *Proceedings of the 2nd International Symposium on Intelligent Robotic Systems*.
- Hurford, J. (2003). Language beyond our grasp: what mirror neurons can, and cannot, do for language evolution. In Kimbrough Oller, U. G. and Plunkett, K., editors, *Evolution of Communication Systems: A Comparative Approach*. MIT Press, Cambridge, MA.
- Iacoboni, M., Woods, R., Brass, M., Bekkering, H., Mazziotta, J., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286:2526–2528.
- Keijzer, F. (2002). Representation in dynamical and embodied cognition. *Cognitive Systems Research*, 3:275–288.
- Keinan, A., Sandbank, B., Hilgetag, C., Meilijson, I., and Ruppín, E. (2004). Fair attribution of functional contribution in artificial and biological networks. *Neural Computation*, to appear.
- Marom, Y., Maistros, G., and Hayes, G. (2002). Toward a mirror system for the development of socially-mediated skills. In *Proceedings Second International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, volume 94, Edinburgh, Scotland.
- Meltzoff, A. (1996). The human infant as imitative generalist: a 20-year progress report on infant imitation with implications for comparative psychology. In Hayes, C. and Galef, B., editors, *Social Learning in Animals; The Roots of Culture*, New York Academic Press.
- Oztop, E. and Arbib, M. (2002). Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics*, 87:116–140.
- Prinz, W. and Meltzoff, A. (2002). *The imitative mind: Development, evolution and brain bases*. Cambridge University Press.
- Ramnani, N. and Miall, R. (2004). A system in the human brain for predicting the actions of others. *Nature Neuroscience*, 7:85 – 90.
- Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2:661–670.
- Ruppín, E. (2002). Evolutionary autonomous agents: A neuroscience perspective. *Nature Reviews Neuroscience*, 3(2):132–141.
- Todd, P. and Miller, G. (1991). Exploring adaptive agency II: Simulating the evolution of associative learning. In Meyer, J. and Wilson, S., editors, *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, pages 306–315, Cambridge, MA. MIT Press.
- Zentall, T. (2001). Imitation in animals: evidence, function, and mechanisms. *Cybernetics and Systems*, 32(1-2):53–96.