



Topics in Cognitive Science (2012) 1–14  
Copyright © 2012 Cognitive Science Society, Inc. All rights reserved.  
ISSN: 1756-8757 print / 1756-8765 online  
DOI: 10.1111/j.1756-8765.2012.01218.x

# Embodied Cognition for Autonomous Interactive Robots

Guy Hoffman

*Media Innovation Lab, School of Communication, Interdisciplinary Center Herzliya*

Received 28 December 2010; received in revised form 25 March 2012; accepted 18 April 2012

---

## Abstract

In the past, notions of embodiment have been applied to robotics mainly in the realm of very simple robots, and supporting low-level mechanisms such as dynamics and navigation. In contrast, most human-like, interactive, and socially adept robotic systems turn away from embodiment and use amodal, symbolic, and modular approaches to cognition and interaction. At the same time, recent research in Embodied Cognition (EC) is spanning an increasing number of complex cognitive processes, including language, nonverbal communication, learning, and social behavior.

This article suggests adopting a modern EC approach for autonomous robots interacting with humans. In particular, we present three core principles from EC that may be applicable to such robots: (a) modal perceptual representation, (b) action/perception and action/cognition integration, and (c) a simulation-based model of top-down perceptual biasing. We describe a computational framework based on these principles, and its implementation on two physical robots. This could provide a new paradigm for embodied human–robot interaction based on recent psychological and neurological findings.

*Keywords:* Embodied cognition; Robotics; Artificial intelligence; Perceptual symbols; Human–robot interaction; Cognitive computational models; Anticipation

---

## 1. Introduction

Among the various applications of computer science, robots are tautologically embodied. However, it seems that many of the computational models underlying intelligent robotic platforms are, to a large extent, unaware of that fact. When embodiment does appear in robotics research, it is usually applied to simple systems, tackling such issues as low-level dynamics and navigation. In contrast, most of the research dealing with high-level intelligent, autonomous, and interactive robotics is still representational: abstract symbol

---

Correspondence should be sent to Guy Hoffman, School of Communication, Interdisciplinary Center Herzliya, P.O. Box 167, Herzliya 46150, Israel. E-mail: hoffman@idc.ac.il

processing systems that view the robot's "body," its sensors and actuators, as proxies to numerical manipulation at best, and as noisy nuisances to circumvent at worst.

At the same time, a growing body of research in psychology and neuroscience is moving away from abstract symbolic models of cognition, emphasizing instead embodied aspects of intelligence (Wilson, 2002). According to this view, human perception and action are not mere input and output channels to an abstract symbol processor or rule-generating engine, but instead decision making, memory, perception, and language are intertwined and grounded in our physical presence (Barsalou, 1999; Pecher & Zwaan, 2005; Wilson, 2001). Importantly, embodied principles are not limited to rudimentary perceptual, physical, and memory tasks but apply to increasingly complex and high-level cognitive processes, such as social cognition, communication, and the coordination of joint activities (Barsalou, 2008; Barsalou, Niedenthal, Barbey, & Rupert, 2003; Sebanz & Bekkering, 2006).

In this article, we suggest that core principles from recent EC research can and should be transferred to a wider range of robotic systems. In particular, embodied robotic cognition research could transcend simple robotic systems, navigation, and dynamics, and be applied to autonomous interactive robots that act in meshed joint activities with humans.

The amodal and modular focus of autonomous robotics can be explained by its roots in so-called good old-fashioned AI (GOFAI) and cybernetics. GOFAI is exclusively concerned with abstract symbol processing and has had its most notable successes in logic, mathematics, game playing, data mining, and classification (for a review of the history of classical AI, see Pfeifer and Bongard [2007]). When AI was beginning to tackle robotics, it stayed true to this symbolic tradition, combining it with ideas from cybernetics, which drew a clear separation between input (sensors), decision making (processors and control), and output (motors and other actuators).

As a result, much of the last 50-year history of robotics adopted this modular view, according to which sensory input is filtered into features, which are analyzed and classified into abstract amodal symbols representing an external world state. The control, learning, or decision-making processes use these states, with additional symbolic knowledge, to choose one or more actions. Actions are then executed by actuators altering the external world, in turn causing new sensory perception. Information thus flows in a unidirectional stream from the world, to sensors, to decision making, to action, and back to the world. Each stage of processing is clearly separated, and approached as a distinct problem, often providing for a whole subfield of AI and robotics.

### *1.1. An embodied alternative*

In a radical departure from this view, Brooks (1991) proposed an alternative "Intelligence without Representation," in which independent low-level behaviors (such as obstacle avoidance, walking dynamics, and so forth) result in an overall intelligent creature that has no symbolic representation of an external world state. He proposed looking to insects and other simple organisms for inspiration, "growing" intelligence as an emergent property of increasingly complex systems.

Following Brooks, a number of embodiment-inspired subfields of robotics emerged. One segment of robotics steered away from modeling human intelligence altogether and instead focused on biomimetics, especially of insects (e.g., Miki & Shimoyama, 1999; Silva, Teneiro Machado, & Jesus, 2008), snakes (e.g., Hirose & Mori, 2004), and other nonhuman organisms. In parallel, walking dynamics (e.g., Collins, Ruina, Tedrake, & Wisse, 2005; for a review, see Chiel, Ting, Ekeberg, & Hartmann, 2009)—and more recently—grasping mechanisms (e.g., Edsinger & Kemp, 2006) have been explored as embodied, nonsymbolic subfields of robotics.

The “Artificial Life” subfield of AI also investigated a variety of complex behaviors emerging from simple rules (for a review, see Langton, 1995), inspiring a different kind of nonrepresentational approach to robot intelligence. One example is that of “swarm robotics” (Şahin, 2005), in which a large number of independent robots behave collectively through simple local interactions. Another is “modular robotics” (e.g., Zykov, Mytilinaios, Desnoyer, & Lipson, 2007), where the overall behavior of a robot is independently controlled within each of its physical subparts.

These lines of research unseated the exclusive status of symbol processing AI in robotics. Still, such embodied interpretations were mostly demonstrated on either very simple systems or dealt with the solution of specific mechanical and sensory dynamics challenges.

Meanwhile, the fields of Personal Robotics and Sociable Robotics (Breazeal, 2002; Fong, Nourbakhsh, & Dautenhahn, 2003) began to form, returning the focus to human-like and human-interactive systems. Some researchers envisioned robots as teammates in human-robot teams (Hinds, Roberts, & Jones, 2004; Hoffman & Breazeal, 2004), initiating new models for artificial social behavior and interpersonal communication. However, many of these new efforts toward interactive robotics returned to a more classical “good old-fashioned” view of Intelligence, in part because they drew heavily on the linguistics- and logics-based field of “Discourse Theory” to model human-robot interaction (e.g., Rickel, Lesh, Rich, Sidner, & Gertner, 2002; Hoffman & Breazeal, 2004).

In neuroscience and psychology, however, an opposite trend developed: EC was found to be applicable to an increasing number of more sophisticated cognitive tasks. Embodied mechanisms were shown capable of modeling abstract thought, language, mathematical reasoning, and learning, as well as social and moral decision making (Barsalou, 1999, 2008; Chandler & Schwarz, 2009; Lee & Schwarz, 2010). And embodied representations could also account for aspects of social interaction, communication, and coordination (Barsalou et al., 2003; Sebanz & Bekkering, 2006).

### *1.2. Embodied cognition for autonomous interactive robots*

These findings suggest a reevaluation of higher level artificial cognition, and of autonomous interactive robotics, in EC terms. Currently, many of the works in autonomous interactive robotics suffer from the drawbacks of abstract symbol systems, such as discreteness, rigid structure, and slowness. At the same time, EC suggests new models and theories applicable to social interaction. We therefore propose a renewed view of EC in the context of

autonomous robots aimed for human interaction, especially where fluid activity with the robot's surroundings and human counterparts is desired.

Specifically, we believe that the idea of modal perceptual representation, the integration of action, perception, and cognition, and a notion of simulation-based top-down perceptual biasing could inform the design of such robots. To that effect, we present an implementation of these three ideas in a novel computational framework, used in two physical robotic systems. The article concludes by suggesting additional ways in which ideas from EC could apply to interactive robotics.

## **2. Three embodiment principles for autonomous interactive robots**

In this section, we propose three EC principles that could be applicable to autonomous interactive robots.

### *2.1. Modal perceptual representation*

Traditional robotics research uses amodal theories of knowledge, which assert that information is processed from perceptual stimuli into nonperceptual symbols, later used for information retrieval, decision making, and action production. Recent findings in EC challenge this view, suggesting instead a perceptual model of cognition, in which concepts are grounded in modal representations, utilizing some of the same mechanisms used during the perceptual process (Barsalou, 1999; Kosslyn, 1995). This is supported by a range of evidence, for example, perceptual neural activation when a subject is using a concept in a nonperceptual manner (e.g., Martin, 2001; Kreiman, Koch, & Fried, 2000); visual priming by reading a sentence that has an implied visual orientation (Stanfield & Zwaan, 2001); memory recall impairment matching speech impediments (Locke & Kutz, 1975); and the increased speed of comparing visually similar variations of a concept, as opposed to visually distinct variations (Solomon & Barsalou, 2001).

The notion of perceptual representation can be translated to computational models of cognition in a number of ways. Gray, Breazeal, and colleagues have presented simulation-theoretic models for a robotic system inferring beliefs, intentions, and goals of a human peer (Breazeal, Gray, & Berlin, 2009; Gray, Breazeal, Berlin, Brooks, & Lieberman, 2005). The robot reuses the same perceptual systems that it uses to generate behavior, to simulate the perceptual perspective, action intentions, and task goals of a human collaborator or adversary. Human-subject studies using the system show the robot displaying comparable behaviors to humans in the same situation.

In this article, we propose another computational model of concepts, memory, and decision making that makes use of modal perceptual representations, in the spirit of Convergence Zones (Damasio, 1989; Simmons & Barsalou, 2003). In the framework described in Section 3, decision making happens in the same modal systems that process perception, by biasing perceptual and sensory processing layers that trigger behavior. Learning, too, is modeled by altering attributes and connections of modal perception-processing systems, and

not by the creation of amodal symbolic rules and relations. Through a process of perceptual simulation, the artificial agents' world model is represented by a mixture of real-world and simulated perceptions. This permits mechanisms akin to human priming and practice, and serves as a basis for learning and conditioning.

## *2.2. Action/perception and action/cognition integration*

In addition to a perception-based theory of cognition, there is an understanding that cognitive processes are similarly interwoven with motor activity. Evidence in human developmental psychology shows that motor and cognitive development are not disparate but highly interdependent. For example, research shows that artificially enhancing 3-month-old infants' grasping abilities equates some of their cognitive capabilities to the level of older, already grasping<sup>1</sup> infants (Sommerville, Woodward, & Needham, 2005). Adult behavior expresses similar interdependence: Hand signals have been shown to be instrumental to lexical lookup during language generation (Krauss, Chen, & Chawla, 1996), and an action/cognition relationship is supported by findings of redundancy in head movements and facial expression during speech generation (Chovil, 1992; McClave, 2000). Wilson (2001) points to an isomorphic representation between perception and action, leading to mutual and often involuntary influence between the two.

In contrast, the role of action and motor execution in robotics has traditionally been viewed as a passive "client" of a central decision-making process, and as such at the receiving end of the data and control flow in robotic systems. Even in so-called Active Perception frameworks (Aloimonos, 1993), the influence of action on perception is mediated through the agent changing its surroundings or perspective on the world, and not by internal processing pathways.

Instead, we suggest that action can affect perception and cognition in interactive robots in the form of symmetrical action-perception activation networks. In such networks, perceptions exert an influence on higher level associations, leading to potential action selection, but are also conversely biased through active motor activity. This close integration between motor activity and perceptual processing could lead to more highly meshed activities between robots and human collaborators, as we argue in Section 3.

## *2.3. Simulation-based top-down perceptual biasing*

The two principles outlined above, as well as a large body of related experimental data, give rise to the following insight: Perceptual processing is not a strictly bottom-up analysis of raw available data, as it is often modeled in robotic and AI systems. Instead, simulations of perceptual processes affect the acquisition of new perceptual data, motor knowledge is used in sensory processing, and intentions, goals, and expectations all play a role in the ability to parse the world into meaningful objects. In other words, sensory-perceptual systems are highly penetrable by cognitive and action processes.

Much experimental data support this hypothesis, suggesting not only that perception is often a predictive activity but also that top-down simulation is a viable model for this

predictive behavior. To give just a few examples: in visual perception, information is found to travel both upstream and downstream, causing object priming to trigger a top-down biasing of lower level mechanisms (Kosslyn, 1995). Similarly, visual lip reading affects the perception of auditory syllables, indicating that the sound signal is not processed as raw unknown data (Massaro & Cohen, 1983). High-level processing seems also to be involved in the perception of human figures from point light displays, enabling subjects to identify “complex actions, social dispositions, gender, and sign language” from sparse visual information (Thornton, Pinto, & Shiffrar, 1998). For a thorough review of related findings, see Wilson and Knoblich (2005) and Barsalou (1999).

While the notion of top-down influences has been explored for some visual tasks in computational systems (e.g., Bregler, 1997; Hamdan, Heitz, & Thoraval, 1999), there is additional potential for using top-down processing in the context of autonomous robots physically interacting with humans, where actions, concepts, and predictions could penetrate lower level perceptual and sensory modules. We moreover believe that simulation-based top-down biasing could specifically be key to more fluent coordination between humans and robots working together in a socially structured interaction.

### 3. Implementation

We exemplify the principles laid out above in a computational framework implemented on two different autonomous interactive robots. The core mechanism, akin to human “priming,” is based on the modeling of the human partner’s activity and the subsequent biasing of perceptual pathways. This relies on two processes: (a) anticipation of the human actions based on repetitive past events, and (b) modeling the resulting anticipatory expectation as modal perceptual simulation, causing a top-down bias of perceptual processes.

To allow for this, concepts are modeled as specific patterns of modal activation and reside within the perceptual streams that process sensory data. Actions are triggered bottom-up through their activation originating in perceptual stimulation, and conversely, anticipated concepts bias the perceptual pathway detecting the properties and features of that concept. This leads to diminished reaction times for confirmatory sensory events, resulting in higher fluency and efficiency in joint actions.

#### 3.1. Modality streams and process nodes

We structure our system in the form of *modality streams* (Fig. 1) built of interconnected *process nodes*. These nodes can correspond to raw sensory input (such as an image frame), to a feature (such as the dominant color or orientation of a data point), to a property (such as the speed of an object), or to a higher level concept describing a statistical congruency of features and properties. This corresponds to the principle of modal concept representation (Section 2.1).

Modality streams are connected to an *action network* consisting of *action nodes*, which are activated in a similar manner as perceptual process nodes. An action node, in turn, leads

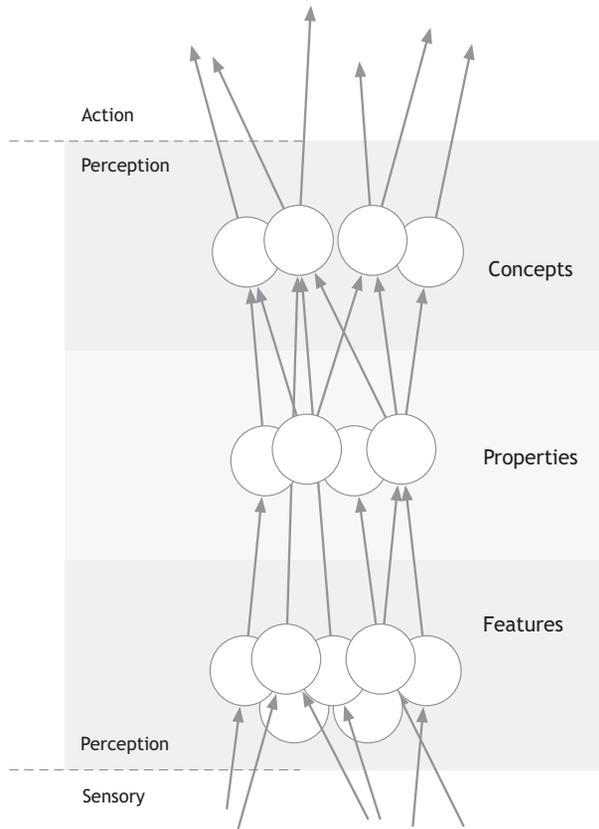


Fig. 1. Schematic of a modality stream.

to the performance of a motor action. Importantly, activation flows in both directions, the *afferent*—from the sensory system to concepts and actions—and the opposite, *efferent*, direction. This is in line with the principle of perception/action integration (Section 2.2).

Each node (Fig. 2) contains an activation value,  $\alpha$ , which represents the processing coming in from the outside world. A separate simulated activation value  $\sigma$  is also taken into account in the node's activation behavior and processing, and it results from top-down processing (as proposed in Section 2.3). The combination of activation and simulation also causes motor action triggers.

### 3.2. Priming

In humans, “priming” is the bias (often measured as a decrease in response time) toward a sensory or memory event. Such priming can occur through cross-modal activation, through previous activation, or from memory recall.

Fig. 3 exemplifies modal priming for an artificial agent in a simple example: An auditory percept (e.g., the sound “Elmo”) activates a visual memory of the Elmo figure (bottom),

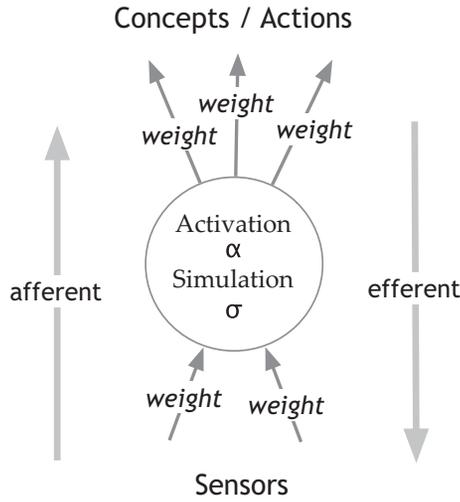


Fig. 2. A process node within a modality stream. Weighted activation travels both up from sensory events to concepts and actions (the afferent pathway), and—through simulation—back downstream (the efferent pathway).



Fig. 3. Top-down processing and cross-modal activation in a perception-based computational framework.

which—using the same pathway utilized in visual perception—resolves into the dominant color of the image (second frame from left). This color is then used as a bias affecting the low-level visual buffer (third frame from left), shifting it toward detection of similarly colored areas, eventually enabling the system to detect the Elmo puppet in the visual field more readily.

In our architecture, the mechanism of artificial priming works as follows: If a certain higher level node  $n$  is activated through priming, the lower level nodes that feed  $n$  are partially activated through the simulation value  $\sigma$  on the efferent pathway. As  $\sigma$  is added to the sensory-based activation  $\alpha$  in the lower level nodes, this simulated top-down penetration

inherently lowers the perceptual activation necessary for the activation of those lower level nodes, decreasing the real-world sensory-based activation threshold for action triggering. The result is reduced response time for anticipated sensory events, and increasingly automatic motor behavior.

### 3.3. Sources of priming

What are the sources of perceptual simulation? We implemented two top-down subsystems to support priming within the proposed perceptual node architecture.

#### 3.3.1. Markov-chain Bayesian anticipatory simulation

The first is a Markov-chain Bayesian predictor, building a probabilistic map of node activation based on recurring activation sequences during practice. This system is in the spirit of the anticipatory system described in Hoffman and Breazeal (2007). It triggers high-level simulation, which—through efferent pathways—biases the activation of lower level perceptual nodes. For example, if a red stop light usually follows a yellow one on a traffic light, then the activation of a yellow traffic light concept activates the red light concept (with a delay), which in turn biases the activation of red feature detectors in the perceptual system, making the robot more responsive to red objects.

If the subsequent real-world sensory data support these perceptual expectations (i.e., the light actually turns red), the robot's reaction times are shortened as described above. In the case where the sensory data do not support the simulated perception, reaction time could be longer and can even lead to a temporary erroneous action, which is then corrected by the real-world sensory data.

#### 3.3.2. Intermodal Hebbian reinforcing

A second mechanism of priming is that of Hebbian reinforcement on existing activation connections. Node connections can be assigned to a connection reinforcement system, which will dynamically change the weights between the nodes. This system works according to the contingency principle introduced in Hebb (1949), reinforcing connections that co-occur frequently and consistently, and decreasing the weight of connections that are infrequent or inconsistent (the “fire together, wire together” principle). The reinforcement of consistent coincidental activations leads to anticipated simulated perception in intermodal perception nodes. For example, the sound of the word “tomato” can, with practice, reinforce the visual concept of “red.” This, again, triggers top-down biasing of lower level perception nodes, shortening reaction times as described above.

### 3.4. Application and evaluation

We have applied this framework on two distinct physical robotic platforms (Fig. 4) designed to operate in synchrony with a human partner. Both robots were evaluated in terms of their action fluency and reaction times with respect to the human's behavior, using repetitive practice. The first robot, Leonardo, is a complex expressive humanoid. We have

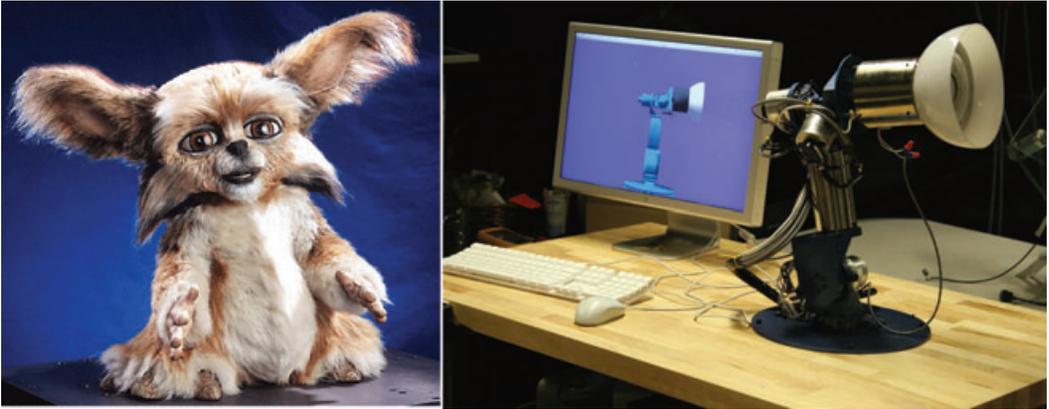


Fig. 4. The robots used as physical platforms for the anticipatory perceptual simulation architecture. The expressive humanoid Leonardo (left), and the nonhumanoid robotic desk lamp AUR (right).

demonstrated that—through repetition—the robot can reduce its reaction times in a play interaction akin to the children’s game “Patty Cake.” The robot’s task was to mirror the hand pattern of the human player facing it, as the human repeated an arbitrary sequence of hand gestures multiple times.

In this application, we only use Markov-chain Bayesian anticipatory simulation to bias the robot’s perceptual system. Before practice, the robot displayed a noticeable lag to the human’s play pattern, but when top-down perception was sufficiently activated, the robot’s movements were near simultaneous to those of the human player.

The second implementation used a nonhumanoid robot, AUR, a robotic desk lamp. In this study, the human and the robot had to solve a joint task of moving around a space together and selecting specific light colors for each area of the joint space. Again, the human led the interaction and repeated the same patterns multiple times. For this setup, we used both Markov-chain Bayesian and Hebbian reinforcement simulation.

In studies with untrained subjects, we showed our framework to be significantly more efficient and fluent than in a comparable system without anticipatory perceptual simulation. We also found significant differences between the two conditions in the human subjects’ sense of team fluency, the team’s improvement over time, the robot’s contribution to the efficiency and fluency, the robot’s intelligence, and in the robot’s adaptation to the task. The results of this study are reported in Hoffman and Breazeal (2010).

#### 4. Discussion

We take EC to refer to a view according to which intelligence is not made up of abstract symbols processed unidirectionally from perceptual processing modules to independent motor devices, but instead that perception, cognition, and action are intertwined and operate in a multidirectional and simultaneous manner.

In recent years, EC research has been expanding to model and interpret an increasing number of cognitive capabilities, including complex perceptual, communicational, and social mechanisms. At the same time, most higher level cognitive functions modeled in robotics revert to a traditional symbol-processing view, in which perception and action are peripheral channels to amodal decision systems. This is particularly true for research in the fields of personal and sociable robots, which engage in verbal and nonverbal discourse with human partners.

In this article, we have proposed three ways in which EC could inspire novel paradigms for autonomous interactive robotics. Specifically, we suggested (a) modal perceptual representation, (b) action/perception and action/cognition integration, and (c) a simulation-based model of top-down perceptual biasing, for EC-based cognitive robot architectures. We also presented an anticipatory perceptual simulation framework exemplifying these principles, and its application to two robotic platforms, used in interaction with untrained human subjects.

Applying embodied cognition principles on any computational system is inherently a challenge, as all computation is, at the lowest level, symbolic, abstract, and modular. That said, one can view artificial intelligence and interactive robotics from a broader point of view, creating ample opportunity to implement the core psychological and neuroscientific findings of EC on robotic systems. And, as robots always act in a physical environment, it is fitting to cease viewing their physical presence as a noisy filter over pseudoperfect information. Instead, and especially when interacting with humans, robots should make use of their embodiment, as the human brain does with its own physical circumstance.

Embodiment naturally found its first venue in robotics in the modeling of simple autonomous systems, and in the solution of specific mechanical and sensory dynamics problems. This signaled a turn from the traditional concerns of AI: language, decision making, planning, and rule learning. However, as EC research matures and begins to account for elements of higher level cognition, autonomous robotics could benefit from a reevaluation of embodiment in these contexts.

Some robotics researchers have begun to integrate embodiment principles in their work. Gray and Breazeal's work on simulation was mentioned above. Gorniak and Roy (2007) have presented a computational theory of language understanding situated in a robot's physical environment. Other robotics research has been concerned with the social-embodied idea of physical behavior expressing internal states (Breazeal, 2002), and the social effects of physical mimicry (Riek, Paul, & Robinson, 2009). In this article, we have suggested a perceptual-simulation interpretation of EC applied to autonomous interactive robots.

However, still more of the recent cognitive and neuropsychological findings and models could be applied to robotics, especially for robots that physically interact with untrained humans. For example, commonsense reasoning—an important field of AI still making broad use of amodal symbolic models (e.g., Havasi, Speer, Pustejovsky, & Lieberman, 2009; Liu & Singh, 2004)—could make use of similar mechanisms humans use when reasoning about the world using their embodied experience. Also, as computer games and digital entertainment become more physically grounded, lessons from embodied practice in sports and performance arts could be utilized to create more responsive and acceptable autonomous game

agents and interfaces; one can even imagine a whole new field of “artificial practice” making use of the embodied techniques of performance artists and athletes. Similarly, EC insights from master craftsmen might shed a new light on industrial robotics solving physical tasks; perhaps a new type of embodied human–robot apprenticeship might emerge. Developmental EC could provide for novel paradigms in robotics aimed at childcare. And similarly, perception–action integration could lead to more natural nonverbal interfaces for nursing robots, and for those systems designed to assist the elderly in their homes.

The multitude of findings in EC, and its increasing reach, in parallel with the growing embodiment of Artificial Intelligence and the greater sophistication of human-interactive robotics, offer a fundamental opportunity for the adoption of EC models in artificial cognitive systems in general, and in particular, in those for autonomous interactive robots.

## Note

1. In the physical sense.

## References

- Aloimonos, Y. (Ed.). (1993). *Active perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Barsalou, L., Niedenthal, P., Barbey, A., & Rupert, J. (2003). Social embodiment. *Psychology of Learning and Motivation*, 43, 43–92.
- Breazeal, C. (2002). *Designing sociable robots*. Cambridge, MA: MIT Press.
- Breazeal, C., Gray, J., & Berlin, M. (2009). An embodied cognition approach to mindreading skills for socially intelligent robots. *The International Journal of Robotics Research*, 28 (5), 656–680.
- Bregler, C. (1997). Learning and recognizing human dynamics in video sequences. In *CVPR '97: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition* (p. 568). Washington, DC: IEEE Computer Society.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139–159.
- Chandler, J., & Schwarz, N. (2009). How extending your middle finger affects your perception of others: Learned movements influence concept accessibility. *Journal of Experimental Social Psychology*, 45 (1), 123–128.
- Chiel, H. J., Ting, L. H., Ekeberg, O., & Hartmann, M. J. Z. (2009). The brain in its body: Motor control and sensing in a biomechanical context. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29 (41), 12807–12814.
- Chovil, N. (1992). Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction*, 25, 163–194.
- Collins, S., Ruina, A., Tedrake, R., & Wisse, M. (2005). Efficient bipedal robots based on passive-dynamic walkers. *Science (NY)*, 307 (5712), 1082–1085.
- Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33, 25–62.
- Edsinger, A., & Kemp, C. (2006). *Manipulation in Human Environments*. 2006 6th IEEE-RAS International Conference on Humanoid Robots (pp. 102–109). Washington, DC: IEEE.

- Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42 (3–4), 143–166.
- Gorniak, P., & Roy, D. (2007). Situated language understanding as filtering perceived affordances. *Cognitive Science*, 31 (2), 197–231.
- Gray, J., Breazeal, C., Berlin, M., Brooks, A., & Lieberman, J. (2005). Action parsing and goal inference using self as simulator. In ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005 (pp. 202–209). Washington, DC: IEEE.
- Hamdan, R., Heitz, F., & Thoraval, L. (1999). Gesture localization and recognition using probabilistic visual learning. In *CVPR '99: Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition* (pp. 2098–2103). Ft. Collins, CO.
- Havasi, C., Speer, R., Pustejovsky, J., & Lieberman, H. (2009). Digital intuition: Applying common sense using dimensionality reduction. *IEEE Intelligent Systems*, 24 (4), 24–35.
- Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.
- Hinds, P. J., Roberts, T. L., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human Computer Interaction*, 19 (1&2), 151–181.
- Hirose, S., & Mori, M. (2004). Biologically inspired snake-like robots. In *2004 IEEE International Conference on Robotics and Biomimetics* (pp. 1–7). Washington, DC: IEEE.
- Hoffman, G., & Breazeal, C. (2004). Collaboration in human-robot teams. In *Proceedings of the AIAA 1st Intelligent Systems Technical Conference, Chicago, IL* (pp. 1–18). Reston, VA: AIAA.
- Hoffman, G., & Breazeal, C. (2007). Cost-based anticipatory action-selection for human-robot fluency. *IEEE Transactions on Robotics and Automation*, 23 (5), 952–961.
- Hoffman, G., & Breazeal, C. (2010). Effects of anticipatory perceptual simulation on practiced human-robot tasks. *Autonomous Robots*, 28 (4), 403–423.
- Kosslyn, S. M. (1995). Mental imagery. In Daniel N. Osherson, Stephen M. Kosslyn, & John M. Hollerbach (Eds.), *An invitation to cognitive science: Visual cognition and action* (2nd ed., Vol. 2), pp. 73–97. Cambridge, MA: MIT Press.
- Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 389–450). San Diego, CA: Academic Press.
- Kreiman, G., Koch, C., & Fried, I. (2000). Imagery neurons in the human brain. *Nature*, 408 (6810), 357–361.
- Langton, C. G. (1995). *Artificial life*. Cambridge, MA: MIT Press.
- Lee, S. W. S., & Schwarz, N. (2010). Dirty hands and dirty mouths: Embodiment of the moral-purity metaphor is specific to the motor modality involved in moral transgression. *Psychological Science*, 21 (10), 1423–1425.
- Liu, H., & Singh, P. (2004). Commonsense reasoning in and over natural language. In M. Negoita, R. Howlett, & L. Jain (Eds.), *Knowledge-based intelligent information and engineering systems* (pp. 293–306). Berlin/Heidelberg: Springer.
- Locke, J. L., & Kutz, K. J. (1975). Memory for speech and speech for memory. *Journal of Speech and Hearing Research*, 18, 176–191.
- Martin, A. (2001). Functional neuroimaging of semantic memory. In R. Cabeza & A. A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition*, 1(3) (pp. 153–186). Cambridge, MA: MIT Press.
- Massaro, D. W., & Cohen, M. M. (1983). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 9 (5), 753–771.
- McClave, E. Z. (2000). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics*, 32 (7), 855–878.
- Miki, N., & Shimoyama, I. (1999). Study on micro-flying robots. *Advanced Robotics*, 13 (3), 245–246.
- Pecher, D., & Zwaan, R. A. (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. Cambridge, England: Cambridge University Press.
- Pfeifer, R., & Bongard, J. C. (2007). *How the body shapes the way we think: A new view of intelligence*. Cambridge, MA: MIT Press.

- Rickel, J., Lesh, N., Rich, C., Sidner, C. L., & Gertner, A. S. (2002). Collaborative discourse theory as a foundation for tutorial dialogue. In Stefano A. Cerri, Guy Gouardères, & Fábio Paraguaçu, (Eds.), *Proceedings of the 6th International Conference on Intelligent Tutoring Systems* (pp. 542–551). Berlin: Springer-Verlag.
- Riek, L. D., Paul, P. C., & Robinson, P. (2009). When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry. *Journal on Multimodal User Interfaces*, 3 (1-2), 99–108.
- Şahin, E. (2005). Swarm robotics: From sources of inspiration to domains of application. In E. Sahin & W. M. Spears (Eds.), *Swarm robotics* (Vol. 3342, pp. 10–20). Heidelberg, Germany: Springer.
- Sebanz, N., & Bekkering, H. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, 10 (2), 70–76.
- Silva, M., Teneiro Machado, J., & Jesus, I. (2008). Research issues in biological inspired sensors for flying robots. In L. Bruzzone (Ed.), *27th IASTED Int'l Conference for Modeling, Identification, and Control* (pp. 262–267). Calgary, AB: ACTA Press.
- Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, 20 (3-6), 451–486.
- Solomon, K. O., & Barsalou, L. W. (2001). Representing properties locally. *Cognitive Psychology*, 43, 129–169.
- Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, 96 (1), B1–B11.
- Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153–156.
- Thornton, I. M., Pinto, J., & Shiffrar, M. (1998). The visual perception of human locomotion. *Cognitive Neuropsychology*, 15, 535–552.
- Wilson, M. (2001). Perceiving imitable stimuli: Consequences of isomorphism between input and output. *Psychological Bulletin*, 127 (4), 543–553.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9 (4), 625–636.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460–473.
- Zykov, V., Mytilinaios, E., Desnoyer, M., & Lipson, H. (2007). Evolved and designed self-reproducing modular robotics. *IEEE Transactions on Robotics*, 23 (2), 308–319.