Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

# The Robot in the Crib

## A developmental analysis of imitation skills in infants and robots

Yiannis Demiris[1] and Andrew Meltzoff[2]

[1]Department of Electrical and Electronic Engineering, Imperial College London

South Kensington Campus, Exhibition Road, London, SW7 2BT, UK

Telephone: +44 (0) 2075946300, Fax: +44 (0) 2075946274

y.demiris@imperial.ac.uk

[2]Institute for Learning and Brain Sciences,

CHDD Building, Room 373

University of Washington, Box 357920

Seattle, WA 98195

Telephone: 206-685-2045 Fax: 206-221-6475

meltzoff@u.washington.edu

**Keywords: Imitation, robot learning, developmental robotics, "like-me" hypothesis, active intermodal matching**

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

## Abstract

Interesting systems, whether biological or artificial, develop. Starting from some initial conditions, they respond to environmental changes, and continuously improve their capabilities. Developmental psychologists have dedicated significant effort to studying the developmental progression of infant imitation skills, because imitation underlies the infant's ability to understand and learn from his or her social environment. In a converging intellectual endeavor, roboticists have been equipping robots with the ability to observe and imitate human actions because such abilities can lead to rapid teaching of robots to perform tasks. We here provide a comparative analysis between studies of infants imitating and learning from human demonstrators, and computational experiments aimed at equipping a robot with such abilities. We will compare the research across the following two dimensions: (a) initial conditions – what is innate in infants, and what functionality is initially given to robots, and (b) developmental mechanisms – how does the performance of infants improve over time, and what mechanisms are given to robots to achieve equivalent behaviour. Both developmental science and robotics are critically concerned with: (a) how their systems can and do go "beyond the stimulus" given during the demonstration, and (b) how the internal models used in this process are acquired during the lifetime of the system.

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

## Introduction

Robot designers are frequently faced with a problem that could be seen as an instance of the Nature vs. Nurture debate in the natural sciences. There is a general consensus that tabula rasa algorithms for robot learning are unlikely to be successful; unguided learning without any support will have difficulty overcoming the complexity involved even in simple tasks. However, there is a growing consensus among robot designers that pre-imposing their own conceptions on the robot's control structure might prove counterproductive as they strive toward more adaptive robots that can learn to operate in changing human societies. Imitation, which is seen as a fundamental avenue of learning in humans (Meltzoff, 2007b), has been proposed as a promising method for a compromise between the two approaches. The robot architecture is designed to adapt in order to benefit from other agents' knowledge; imitation can be used as a mechanism of learning for the robot in human and robot societies (Demiris and Hayes 2002; Schaal et al 2003). But where do we start? What bootstrapping mechanisms should we strive to implement in robotic systems and how flexible are such systems as a result?

Developmental psychologists have been studying human infants for more than a century, at first through observation and in the recent decades more systematically through experimental techniques, to determine what capabilities infants are born with and how these capabilities develop over time and with experience. Imitation invokes and coordinates the perceptual, representational, memory and motor neural systems of the infant, and as such is seen as a window into the infant’s sensorimotor and cognitive abilities. Imitation is proving to be useful in the diagnosis of, and theorizing about, disorders such as autism (Dawson, Meltzoff, Osterling, and Rinaldi, 1998; Nadel, 2002;

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

Rogers, 1999) as well as visuoimitative apraxia (Goldenberg and Hagmann, 1997), due to its rich involvement of the underlying neural systems (Carey, Perrett, Oram, 1997). Infants gradually improve their abilities over the first years of their lives to reach a stage where they not only imitate the actions of demonstrators, but also the underlying intentions and goals of demonstrators (Meltzoff, 2007b).

This paper presents a developmental analysis of imitation skills in infants and robots, focusing on a comparison of the work of the two authors (space limitations preclude a full survey of this burgeoning field, see (Breazeal and Scassellati 2002, Nehaniv and Dautenhahn, 2007, Schaal, 1999, Schaal et al 2003), for interesting surveys). Developmental studies of infant imitation skills can provide a source of inspiration for roboticists: new algorithms for tackling the problem of transferring skills from humans to robots can be derived by studying the hypotheses put forward with respect to infant imitation skills. Conversely, the drive towards precise descriptions so that mechanisms can be implemented on computational and robotic platforms serves to expose gaps in hypotheses and can suggest human developmental experiments for filling these gaps.

**System 0.0: In the Beginning**

In the analysis of any developing system, the initial conditions play a crucial role. For infants this translates into investigating what imitation behaviour they are capable of displaying during the beginning of their postnatal life - not always an easy task. For robots this task is easier, since it simply involves describing the algorithms and initial knowledge that are pre-programmed into the robot before the imitation experiments begin.

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

*Perceptually similar vs. perceptually opaque acts*

Imitation involves motor action towards a goal. The imitator, whether infant or robot, must be able to compare its current state with this goal in order to drive its motor system towards achieving this goal. This leads us to a distinction between types of imitation where the goal and the current state can be represented in the same modality, allowing a "direct" comparison, such as vocal imitation or imitation of hand gestures versus imitation of opaque acts that the agent cannot see or hear itself perform, such as facial imitation. From the robotics perspective the first case is relatively easy since a simple control mechanism that is trying to reduce the error between goal state and current state will suffice; in infant literature, given the relative immaturity of the infant's motor system, issues such as imitation of hand gestures are difficult to investigate, although the available literature suggests it emerges quite early. The second case, facial imitation, is particularly interesting since it can illuminate the nature of representational and comparative mechanisms that neonatal infants posses and utilise during the imitation process. It is also significantly harder to achieve in robots since a mapping between states in different modalities needs to be established before an error-reducing controller can be implemented. Infants have been shown to imitate facial gestures as early as 42 minutes following birth (Meltzoff and Moore, 1983, 1989), implying the existence of a basic mechanism for comparing states across modalities (see Meltzoff and Moore, 1997, for detailed arguments concerning a possible role for prenatal movement experience).

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

*From Observing to Imitating*

A mechanism that could accomplish both facial and manual imitation was provided by Meltzoff and Moore (1997) and termed the Active Intermodal Matching (AIM) model. A diagrammatic version of the model is shown in Figure 1a. The first computational implementation of this mechanism in a robotic platform was performed by (Demiris et al, 1997) who used a robotic head to observe and imitate human head movements, a diagrammatic version of which is shown in figure 1b. More recently, Rao, Shon, and Meltzoff, (2007) proposed a probabilistic formulation of this mechanism, which has the potential of improving the robustness of such mechanisms to the uncertainty of the physical world, and the noise of the sensors used to perceive its current state.

INSERT FIGURE 1 HERE

The experiments of Demiris et al 1997 demonstrated that the AIM mechanism is computationally implementable, and sufficient for achieving imitation of movements, but also raised a number of interesting questions regarding the underlying representational mechanisms.

(a) How can visual and proprioceptive representations be compared during imitation? The robotic experiments used the posture of the demonstrator as the "supramodal" unit of comparison, fulfilling the role of what (Meltzoff & Moore 1997) termed "organ relations";

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

head joint angles were used to achieve the results in Demiris et al 1997. Visual information regarding the state of the demonstrator arrive at the infants' visual neural system, undergoing a series of representational transformations starting from the infant's retina up to the higher level cortical areas. A similar process occurs in the proprioceptive system. AIM hypothesizes that these transformations eventually provide neural representations of "organ relations", the supramodal representations that can be compared directly between the two modalities. Similarly, in Demiris et al 1997, visual and proprioceptive information are converted into postural representations, which can be used both for representing the state of the demonstrator as well as the state of the imitator. From experiments in monkeys we know that postural units of representation exist (neurons in the STS area fire according to the postural state of a demonstrator's body parts (Carey et al 1997), (Allison et al 2000)), but whether these are innate or develop with exposure to such stimuli is not known yet, and there has been less work in humans than in monkeys about the neural coding of body parts. Thus the progress in both developmental psychology and robotics suggests that neuroscience explorations of infant body representation will be an especially fruitful area, especially if we seek a fuller account of the mechanisms underlying human imitation.

(b) What is retained in memory during the demonstration? Since the demonstration is a time-varying (spatiotemporal) representation, the robotic experiments revealed a number of choices: retaining the entire sequence of representation units (at a potentially high sampling rate), retaining only the final goal/endpoint, or as a compromise, retaining only representative units along the trajectory.

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

In infant development research, there is a growing body of work about 'goals', but this has been largely concentrated on older infants and children (e.g., Bekkering, Wohlschläger, and Gattis, 2000; Meltzoff, 2007b; Williamson and Markman, 2006). However as it was shown in (Meltzoff and Moore 1983), newborns are indeed able to imitate static gestures, thus the goal/endpoint is clearly important, and in fact infants seem to be able to identify the means of achieving the endstate when they see the endstate.

On the other hand, we also know that infants observe the particular action and not just an endpoint or goal, at least not broadly construed. They do imitate the particular means an adult uses. This is shown in the (Meltzoff, 1988) study of head touch where 14-month-olds copied the novel way the adult turned on a light (by touching a panel with his head). The adult did it using an unusual means, and the infants adopted that means instead of using their more familiar means, their own hands. So infants seem to be able to imitate based on the goal and also copy the means, although the timing of the onset of these mechanisms is an active topic of research (e.g, Williamson & Markman, 2006).

**Developmental change**

The imitation skills of infants do not remain static but rather improve over time and with experience. For example, older infants are capable of imitating not only the surface behaviour of demonstrators, but also understand their underlying goals and intentions. In an experiment by (Meltzoff, 1995, 2007b), 18-month old children were shown unsuccessful acts involving a demonstrator trying but failing to achieve his goal, i.e. the children did not see the successful end-state. Yet children did not replicate the unsuccessful surface behaviour of the adult but proceeded to imitate the intended goal,

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

even when it was never shown to them. Inferring the goals of the demonstrator enables infants (and robots) to go beyond the surface behaviour, and to engage in a more productive imitation. Children under about 1-year-old and younger seem to fail at this task (Bellagamba and Tomasello, 1999; Meltzoff, 2007b), although with a different paradigm and much easier acts such abilities might be demonstrable. An important development between 1 yr and 1.5 years seems to be taking place in the infant's ability to infer goals (Meltzoff, 1995, 2007b). The conditions under which infants imitate exact actions versus reaching goals are receiving increased empirical scrutiny (e.g., Gleissner, Meltzoff, and Bekkering, 2000; Williamson and Markman, 2006).

The concentration on the goal while ignoring the surface behaviour implies that infants have acquired knowledge that is put into use during this inference process. Meltzoff (2007b) argued for the existence of a self-learning development phase, where infants' bodily acts are associated with underlying mental states. Once the infants have built up this experience, they can use it to infer how the external state of the demonstrator relates to its internal one. They can understand others "by analogy with the self" (Meltzoff 2005, 2007a 2007b). For robotics, this is an exciting proposition since it provides a pathway through which robots can autonomously discover how to behave within human societies. This proposition is not without difficulties, for example, finding a principled way for selecting the next dimension to explore among the many available in the state space. An example of this could be a principled way of selecting the next action (from the ones available to the agent) to execute for exploration purposes. This can become a key example of how developmental science and artificial intelligence can be mutually informative.

**Understanding others by analogy with the self**

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

How can this process, "understanding others by analogy with the self" be implemented in robotic systems? In (Demiris and Hayes 2002, Demiris and Johnson 2003, Demiris and Khadhouri 2006) we proposed to do so through a computational architecture termed HAMMER (Hierarchical Attentive Multiple Models for Execution and Recognition).

*Inverse and Forward Models*

HAMMER makes extensive use of the concepts of inverse and forward models. An inverse model (akin to the concepts of a controller, behavior, or action) is a function that takes as inputs the current state of the system and the target goal(s) and outputs the control commands that are needed to achieve or maintain those goal(s). Related to this concept is that of a forward model of a controlled system: a forward model (akin to the concept of internal predictor) is a function that takes as inputs the current state of the system and a control command to be applied to it and outputs the predicted next state of the controlled system (Miall and Wolpert, 1996).

The building block of HAMMER is an inverse model paired with a forward model (figure 2). When HAMMER is asked to rehearse or execute a certain action, the inverse model module receives information about the current state (and, optionally, about the target goal(s)), and it outputs the motor commands that it believes are necessary to achieve or maintain these implicit or explicit target goal(s). The forward model provides an estimate of the upcoming states should these motor commands get executed. This estimate is returned back to the inverse model, allowing it to adjust any parameters of the action (an example of this would be achieving different movement speeds (Demiris and Hayes

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

2002)). Combinations of inverse and forward models are frequently used in motor control (Narendra and Balakrishnan 1997, Wolpert and Kawato 1998) due to their flexible modular structure, and have been advocated for use in imitation and learning (Demiris and Hayes 2002, Demiris and Khadhouri 2006, Schaal 1999, Schaal et al 2003, Wolpert et al 2003).

*The Development of Inverse and Forward Models*

HAMMER's basic building blocks are inverse and forward models. What is the origin of these in robots, what correspondence do these models have in infants, and how are they developed? Meltzoff & Moore's (1997) concept of 'body babbling' or 'motor babbling' is of fundamental importance here. It is a very powerful concept, derived from the infant data, with relevance and application in robotics. Is the newborn infant capable of executing commands that will drive its motor system towards a desired goal (inverse models), and is the newborn infant capable of predicting consequences of its motor commands (forward models)?

In robotics, forward models can be developed through an exploration process, akin to the "motor babbling" stage of infants. Sending a series of random motor commands to the robot's apparatus, we can associate the motor commands with their visual, proprioceptive or environmental consequences (including the expected time that they will appear (Dearden and Demiris, 2005)). Building associations between actions and consequences through motor babbling means that these associations can be inverted (Demiris and Dearden 2005) to create approximations to basic inverse models. Observation and imitation can play a key role at this stage: once basic inverse models have been created, by

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

observing and imitating others (observation of how certain goals can be achieved) the robot can combine (chain together, and/or place in parallel) basic blocks to form more useful complex inverse models.

In infants the situation is less clear: as we described earlier, infants get some 'goals' from observing others. But they also seem to generate random acts and then want to 'perfect them' themselves (Gopnik, Meltzoff, and Kuhl, 2001; Piaget, 1952). They may watch their hand float across their visual field and then 'want to gain control' of this sight. This causes them to repeat it again and again until they've mastered it. Infants seem to be born with this innate 'mastery motivation'(Gopnik et al., 2001; Meltzoff and Moore, 1997).

*In the demonstrator's shoes*

HAMMER uses an inverse-forward model coupling in a dual role: either for executing an action, or for perceiving the same action when performed by a demonstrator. If HAMMER is to determine whether a visually perceived demonstrated action matches a particular inverse-forward model coupling (figure 2), the demonstrator's current state as perceived by the imitator is fed to the inverse model. The inverse model generates the motor commands that it would output *if it was in that state and wanted to execute this particular action*. In a sense, the imitator processes the actions by analogy with the self – "what would I do if I was in the demonstrator's shoes?"

In the perception or planning modes, the motor commands are inhibited from being sent to the motor system. The forward model outputs an estimated next state, which is a prediction of what the demonstrator's next state will be. This predicted state is compared with the

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

demonstrator's actual state at the next time step. As seen in figure 2 below and the text that follows, this comparison results in an error signal that can be used to increase or decrease the behaviour's confidence value, which is an indicator of how closely the demonstrated action matches a particular imitator's action.

INSERT Figure 2 HERE

HAMMER consists of multiple pairs of inverse and forward models that operate in parallel [Demiris and Hayes 2002]. When the demonstrator agent executes a particular action the perceived states are fed into all of the imitator's available inverse models. As described earlier, this generates multiple motor commands (representing multiple hypotheses as to what action is being demonstrated) that are sent to the forward models. The forward models generate predictions about the demonstrator's next state: these are compared with the actual demonstrator's state at the next time step, and the error signal resulting from this comparison affects the confidence values of the inverse models. At the end of the demonstration (or earlier if required) the inverse model with the highest confidence value, i.e. the one that is the closest match to the demonstrator’s action is selected. Demiris and Hayes, 2002, Demiris 2007 have described the relation of this process to a biological counterpart, the mirror system (Gallese et al, 1996), offering a number of explanations and testable predictions (Demiris and Hayes 2002, Demiris and Simmons, 2006), for example, a predicted dependency of the firing rate of the macaque monkey mirror neurons to the velocity profile of the demonstrated act.

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

**Beyond the stimulus given**

Where do we go from here? As we described, infants do not simply react to the stimulus given to them, and architectures for implementing a model-based (versus purely stimulus reactive) approach in robots have already been proposed. In both disciplines the challenge remains in how human action models are acquired and integrated with the (potentially incomplete or confusing) observed stimulus to derive the demonstrator's goals and intentions. From the robotics perspective, challenging questions that remain unanswered within the multiple inverse-forward models formulation include:

How are the human action models organised? In robots, hierarchical formulations have been proposed and used (Demiris and Johnson, 2003, Tani and Nolfi 1999), but their relation to biological data has not been explored (but see Demiris and Simmons 2006). Recent evidence on how infants encode goals also suggests hierarchical representations (Bekkering et al., 2000; Gleissner et al. 2000), making this a potential interesting area for interdisciplinary collaboration.

- How do the observer's internal models and prior knowledge influence what parts of the stimulus will be attended to? We have suggested elsewhere (Demiris and Khadhouri 2006) a computationally principled method for solving this, based on the motor theory of perception, but again the relation to biological/developmental data requires further exploration.
- The class of theories being described here require the observer to "step into the demonstrator's shoes", essentially take the perspective of the demonstrator (for relevant neuroscience data, see Jackson, Meltzoff, and Decety, 2006). Although algorithmic solutions to this have been proposed (Johnson and Demiris, 2005,

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

Breazeal et al 2006, Trafton et al 2005), higher levels of perspective taking, including beliefs, desires, and intentions remain difficult challenges in robotics. These mechanisms can be informed from developmental work on gaze following, which can be viewed as the lowest end of perspective taking. One year old infants follow the gaze of adults and understand that it is directed at an object, and is not yet another meaningless body movement (Brooks & Meltzoff, 2002, 2005; Carpenter, Nagell, and Tomasello, 1998; Flom, Lee, & Muir, 2006). The evidence points to a use of first-person experience to make third-person attributions; for example, it has been shown that once infants had experience with blindfolds, the interpretation of others who wear blindfolds was changed (Meltzoff, 2007a; Meltzoff & Brooks, 2004). The robotics work on action understanding (Demiris and Johnson 2003), and particularly the perceptual perspective taking architecture is compatible with these findings (Johnson and Demiris 2005); perceptual perspective taking allowed an observer robot to "place itself in the demonstrator robot's perceptual shoes" and engage the inverse models that were compatible with the demonstrator's viewpoint rather than its own viewpoint (Johnson and Demiris 2005). While there is still a lot of work to be done in robotics on this aspect, research on the development of perspective taking and its roots in gaze following (Meltzoff, 2005, 2007a) has the potential of providing robotics researchers with important insights on how these mechanisms can be implemented.

To conclude, according to Meltzoff's (2007a, 2007b) "Like Me" theory, the mechanisms involved in successful imitation are the developmental foundation for the later emergence of other higher-order skills in human social cognition. The child bootstraps from recognizing that they can *act* like another agent to the realization that the other has

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

desires/intentions that generate the actions, just like the self does. The HAMMER architecture (Demiris and Khadhouri, 2006) and the work on learning inverse and forward models (Dearden and Demiris, 2005; Demiris and Dearden, 2005) use these ideas to develop analogous skills for robots so that they can operate usefully among humans as robotic assistants. Conversely, we believe that work in robotics will also help sharpen ideas about the details of this developmental mechanism.

## Acknowledgments

This work has been supported in the UK (YD) by the Engineering and Physical Sciences Research Council (EPSRC) and the Royal Society, and in the USA (AM) by NIH (HD-22514) and NSF (SBE-0354453).

## References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4*, 267-278.

Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology, 53A*, 153-164.

Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: Comparing 12- and 18-month-olds. *Infant Behavior & Development, 22*, 277-282.

Breazeal C. and Scassellati B. (2002). Robots that imitate humans, *Trends in Cognitive Sciences, 6,* 481-487.

Breazeal C, Berlin M, Brooks A, Gray J., and Thomaz A. (2006). Using Perspective Taking to Learn from Ambiguous Demonstrations, *Robotics and Autonomous Systems, 54*, 385-393.

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

Brooks, R., & Meltzoff, A.N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology*, *38*, 958-966.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8,* 535-543.

Carey, D. P., Perrett, D. I., & Oram, M. W. (1997). Recognizing, understanding and reproducing action. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 11, pp. 111-129). Amsterdam: Elsevier Science.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63* (4, Serial No. 255).

Dawson, G., Meltzoff, A. N., Osterling, J., & Rinaldi, J. (1998). Neuropsychological correlates of early symptoms of autism. *Child Development, 69*, 1276-1285.

Dearden, A., & Demiris, Y. (2005). Learning forward models for robots. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 1440-1445). Edinburgh.

Demiris, Y. (2007). Using robots to study the mechanisms of imitation, in *Neuroconstructivism II – Perspectives and Prospects*, Mareschal, D, Sirois S., Westermann G., and Johnson M. (eds), (pp.159-178), Oxford: Oxford University Press.

Demiris, Y., & Dearden, A. (2005). From motor babbling to hierarchical learning by imitation: A robot developmental pathway. In L. Berthouze, et al. (Eds.), *Proceedings of the Fifth International Workshop on Epigenetic Robots (EPIROB): Modeling Cognitive Development in Robotic Systems* (pp. 31-37). Japan.

Demiris, J., Rougeaux, S., Hayes, G. M., Berthouze, L., & Kuniyoshi, Y. (1997). Deferred imitation of human head movements by an active stereo vision head. *Proceedings of*

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

*the 6th International Workshop on Robot and Human Communication (IEEE)*, 88-93, Sendai, Japan.

Demiris, Y., & Hayes, G. (2002). Imitation as a dual-route process featuring predictive and learning components: A biologically plausible computational model. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 327-361). Cambridge, MA: MIT Press.

Demiris, Y., & Johnson, M. (2003). Distributed, predictive perception of actions: A biologically inspired robotics architecture for imitation and learning. *Connection Science, 15*, 231-243.

Demiris, Y., & Khadhouri, B. (2006). Hierarchical attentive multiple models for execution and recognition of actions. *Robotics and Autonomous Systems, 54*, 361-369.

Demiris, Y., & Simmons, G. (2006). Perceiving the unusual: temporal properties of hierarchical motor representations for action perception. *Neural Networks, 19*, 272-284.

Flom, R., Lee, K., & Muir, D. (Eds.). (2006). *Gaze following: Its development and significance*. Mahwah, NJ: Erlbaum.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593-609.

Gleissner, B., Meltzoff, A. N., & Bekkering, H. (2000). Children's coding of human action: Cognitive factors influencing imitation in 3-year-olds. *Developmental Science, 3*, 405-414.

Goldenberg, G., & Hagmann, S. (1997). The meaning of meaningless gestures: A study in visuo-imitative apraxia. *Neuropsychologia, 35*, 333-341.

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (2001). *The scientist in the crib: What early learning tells us about the mind*. New York: Harper Paperbacks.

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

Johnson, M., & Demiris, Y. (2005). Perceptual perspective taking and action recognition. *International Journal of Advanced Robotic Systems, 2*, 301-308.

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2006). Neural circuits involved in imitation and perspective-taking. *NeuroImage, 31*, 429-439.

Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology, 24*, 470-476.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838-850.

Meltzoff, A. N. (2005). Imitation and other minds: The "Like Me" hypothesis. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science* (Vol. 2, pp. 55-77). Cambridge, MA: MIT Press.

Meltzoff, A. N. (2007a). 'Like me': a foundation for social cognition. *Developmental Science, 10,* 126-134.

Meltzoff, A. N. (2007b). The 'like me' framework for recognizing and becoming an intentional agent. *Acta Psychologica*, *124*, 26-43.

Meltzoff, A. N., & Brooks, R. (2004, May). Developmental changes in social cognition with an eye towards gaze following. In M. Carpenter & M. Tomasello (Chair), *Action-based measures of infants' understanding of others' intentions and attention*. Symposium conducted at the Biennial meeting of the International Conference on Infant Studies, Chicago, Illinois.

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development, 54*, 702-709.

Meltzoff, A. N., & Moore, M. K. (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology, 25*, 954-962.

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting, 6*, 179-192.

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks, 9*, 1265-1279.

Nadel, J. (2002). Imitation and imitation recognition: Functional use in preverbal infants and nonverbal children with autism. In A. N. Meltzoff & W. Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases* (pp. 42-62). Cambridge: Cambridge University Press.

Narendra K.S. and Balakrishnan J., Adaptive Control using Multiple Models, *IEEE Transactions on Automatic Control*, *42,* 171-187.

Nehaniv C. and Dautenhahn K. (2007), *Imitation and Social Learning in Robots, Humans and Animals,* Cambridge University Press, 2007.

Piaget, J. (1952). *The origins of intelligence in children* (M. Cook, Trans.). New York: International Universities Press.

Rao, R. P. N., Shon, A. P., & Meltzoff, A. N. (2007). A Bayesian model of imitation in infants and robots. In C. L. Nehaniv & K. Dautenhahn (Eds.), *Imitation and social learning in robots, humans, and animals: Behavioural, social and communicative dimensions* (pp. 217-247). Cambridge: Cambridge University Press.

Rogers, S. J. (1999). An examination of the imitation deficit in autism. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy* (pp. 254-283). Cambridge: Cambridge University Press.

Schaal, S. (1999), Is Imitation learning the route to humanoid robots?, *Trends in Cognitive Sciences, 3*, 233-242,.

Schaal, S., Ijspeert, A., & Billard, A. (2003). Computational approaches to motor learning by imitation. *Philosophical Transactions of the Royal Society of London. Series B,*

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

*Biological Sciences, 358*, 537-547.

Tani J. and Nolfi S., (1999), Learning to perceive the world as articulated: an approach for hierarchical learning in sensory motor systems, *Neural Networks*, *12*, 1131-1141.

Trafton J., Cassimatis N., Bugajska M., Brock D., Mintz F. and A. Schultz, Enabling Effective Human-Robot Interaction using Perspective Taking in Robots, *IEEE Transactions on Systems, Man and Cybernetics Part A: Systems and Humans*, *35,* 460-470.

Wolpert D. M. and Kawato M. (1998), Multiple paired forward and inverse models for motor control, *Neural Networks, 11,* 1317:1329.

Wolpert D. M., Doya K., Kawato M. (2003) A unifying computational framework for motor control and social interaction, *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences,358, 593-602.*

Williamson, R. A., & Markman, E. M. (2006). Precision of imitation as a function of preschoolers' understanding of the goal of the demonstration. *Developmental Psychology, 42*, 723-731.

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

## FIGURES

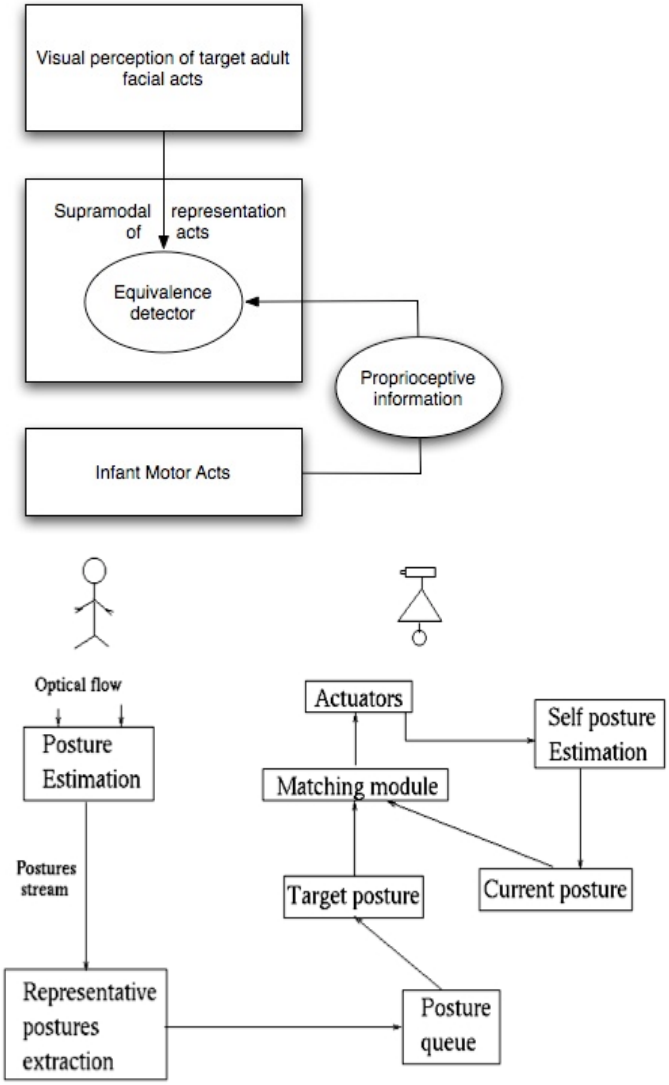



Figure 1: The theory of Active Intermodal Matching (Meltzoff & Moore 1997) and the outline of the robotic implementation from (Demiris et al 1997)

Preprint version; final version available at onlinelibrary.wiley.com
Infant and Child Development (2008), vol: 17(1), pp: 43-53
DOI: 10.1002/icd.543

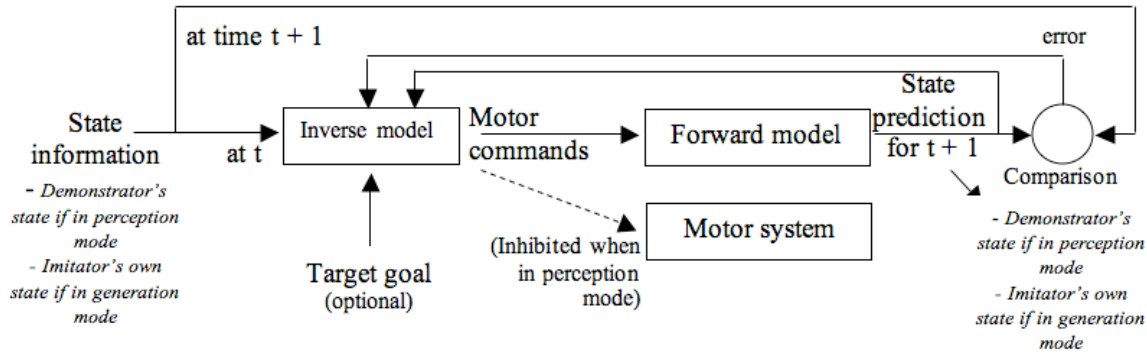



*Figure 2: HAMMER's basic building block, an inverse model paired with a forward model (from Demiris & Hayes 2002, Demiris and Johnson 2003)*