From Saccades to Grasping

Law, James Alexander; Shaw, Patricia Hazel; Lee, Mark Howard; Sheldon, Michael Timothy

Published in:
IEEE Transactions on Autonomous Mental Development

DOI:
10.1109/TAMD.2014.2301934

Publication date:
2014

Citation for published version (APA):
Law, J. A., Shaw, P. H., Lee, M. H., & Sheldon, M. T. (2014). From Saccades to Grasping: A Model of Coordinated Reaching Through Simulated Development on a Humanoid Robot. *IEEE Transactions on Autonomous Mental Development*, *6*(2), 93-109. Article 6776396. https://doi.org/10.1109/TAMD.2014.2301934

From Saccades to Grasping: A model of Coordinated Reaching through Simulated Development on a Humanoid Robot

James Law *Member, IEEE,* Patricia Shaw, Mark Lee, and Michael Sheldon

***Abstract*—Infants demonstrate remarkable talents in learning to control their sensory and motor systems. In particular the ability to reach to objects using visual feedback requires overcoming several issues related to coordination, spatial transformations, redundancy, and complex learning spaces.**

This paper describes a model of longitudinal development that covers the full sequence from blind motor babbling to successful grasping of seen objects. This includes the learning of saccade control, gaze control, torso control, and visually elicited reaching and grasping in 3D space. The paper builds on and extends our prior investigations into the development of gaze control, eye-hand coordination, the use of constraints to shape learning, and a schema memory system for the learning of sensorimotor experience. New contributions include our application of the LWPR algorithm to learn how movements of the torso affect the robot's representation of space, and the first use of the schema framework to enable grasping and interaction with objects.

The results from our integration of these various components into an implementation of longitudinal development on an iCub robot show their ability to generate infant-like development, from a start point with zero coordination up to skilled spatial reaching in less than 3 hours.

I. Introduction

REACHING to grasp seen objects is a multifaceted skill that seems almost trivial to adult humans; yet its subtle complexity is dramatically exposed when attempts are made to implement such skills on humanoid robots. While it is possible to engineer robot grasping procedures directly, this produces fixed behaviours and adaptation to new objects or tasks has proven very limited [1]. A more fruitful approach is to recognise the progressive behaviour of very young infants, who have to learn to reach, and attempt to model this developmental learning for use as a method in robotics [2], [3]. This paper presents an implementation of an infant inspired model that takes an iCub robot [4] through a sequence of growth, similar to an infant, from uncoordinated motor babbling of the eyes and arms to accurate reaching for objects with coordinated involvement of eyes, head, torso, arms and hands. The results show this whole developmental process being run on an iCub robot in less than 3 hours. The full sequence behind this model is reported, although the involvement of body movement (the torso in this case) and a memory component (to allow the chaining of action fragments) are the main advances of interest here.

This research has received funds from the European Commission 7th Framework Programme (FP7/2007-2013), "Challenge 2 - Cognitive Systems, Interaction, Robotics", grant agreement No. ICT-IP-231722, project "IM-CLeVeR - Intrinsically Motivated Cumulative Learning Versatile Robots".

Authors are with the Department of Computer Science, Aberystwyth University, Ceredigion, SY23 2HW, UK, e-mail: {jxl|phs|mhl}@aber.ac.uk

Reaching in humans requires the coordination of several different muscle groups controlling the shoulder, elbow, and wrist. Each of these requires relationships to be established between the range of proprioceptively sensed joint positions and the muscle movements needed to reach those positions. Furthermore, reaching to seen objects requires the space of possible reach positions to be mapped onto the visual space perceived by the eye, however this is not straightforward as multiple arm poses may be available to reach each seen position [5].

These issues pose very significant problems for humanoid robots that have to learn to reach and control their arms. The main concerns are as follows:

- Robot arms often have multiple kinematically dependent joints. This means the joint space is high dimensional (>3: the strict minimum for three dimensional space) and this has two consequences: redundancy is introduced, and the data available for learning will be sparse [6].
- Visual- and joint-spaces are not topographically related, requiring some kind of transformation, and this will be non-linear [7].
- Noise involves more than normal measurement and actuator error because the environment is potentially entirely unconstrained. This means any data points for learning can be uncertain, conflicting or even bogus.
- Models of space under such circumstances can cause difficulties in generating smooth reaching trajectories without discontinuities [8].

There have been many studies and experiments on robot learning of reaching, using both neural models and AI based methods, e.g. [9], but very few perform hand/eye coordination learning on complex kinematics in real time without prior training, or *a priori* knowledge.

In this paper we present a model of learning to reach, inspired by early infant development. Our working hypothesis is that the closer we can model early infant behaviour and the more of its particular characteristics we can reproduce, the more powerful will be the algorithms produced for robotic learning. We identify several key factors that we consider important principles to be included in such models:

- *Stages in development*. All experimental psychologists report distinguishable stages in the acquisition of sensory-motor skill, the key source being Piaget [10]. These vary

in timing across individuals but are consistent in general sequencing [11]. Maturational constraints are believed to play a significant role in the emergence of stages, and reduce the complexity of learning [12]. With regard to reaching, particular attention has been paid to the staged release of constraints [13], their impact [14], and possible emergence [15].
- *Motor babbling*. This is spontaneous, internally motivated, action that generates sensorimotor data during infancy [16], and has been demonstrated as a powerful learning mechanism for robotics [17], [18], [19]. We show that this need not be a random activity but can be functional in relating previous action to current and new sensory-motor patterns. This has close links to the role of play behaviour.
- *Proprioception*. Proprioception develops during pre-natal growth and, along with motor babbling, facilitates the learning of muscle control [20]. Proprioception develops before vision and visual involvement in reaching, and is an essential component in learning and controlling reaching motions [21], [22].
- *Patterns of development*. Infant development follows cephalocaudal and proximal to distal patterns, with eye and head control appearing before arm and torsor [23]. Furthermore, upper arm control appears before forearm control and grasp learning [24], and this sequencing of activity has important ramifications for robot learning [25], [26], [27].
- *Coarse to fine development*. Infant abilities initially appear quite crude and coarse, but are gradually refined over time [28]. This relates to increasing resolution in the sensory and motor subsystems due to physiological maturity, as well as to the development of skills [29]. Such refinement has been shown to aid learning in robotic systems through the initial reduction in complexity [30], [31].

For further reading on these concepts in developmental psychology see the excellent texts [32] and [11]. For some of the origins see [10] and [29], and for infant behaviour see [23] and [33].

Investigations into psychologically grounded reaching have tended to focus on single periods, or described single algorithms which are reported as giving rise to particular behaviours seen in infant reaching. Particular focus is placed on early infant reaching, and the process of learning how to control the hand to reach to a seen target. Fagg et al. [34] explore the tuning of motor control parameters, through supervised learning, to drive a dynamic arm to a target. This is identified as being similar to the early reaching behaviour occurring in the infant at 4 months. Schlesinger et al. [35] also investigates early infant reaching, but the focus is on investigating the constraints that shape reaching development. Using an econet to simulate early reaching, they show how several constraints (muscle activation and stiffness, joint locking, stereotyped movements, and movement slowdown) emerge during the learning process. Dahl and Barige [36] demonstrate an algorithm for refining single pulse reaching movements, which could be seen as modelling the transition from neonatal swiping behaviour to the onset of controlled reaching.

We view developmental sequences as the key to skill learning, and various other works also acknowledge this. Grupen recognised the cephalocaudal progress of infant growth [26] and used this in skill development in robotics [25]; the RobotCub project has produced a robot development map [37] similar to [38]; and Asada and colleagues are researching into a range of robotic models with strong emphasis on human cognitive growth [3], including the earliest stage possible; fetal development [21]. Others report on developmental approaches to reaching, including staged release of maturational constraints [13], and experiments with proximo-distal maturation show that the application of developmental constraints produces more effective and efficient learning [14].

This paper extends our prior investigations into modelling infant development, which include the use of constraints to shape learning [38], the development of a biologically inspired architecture for gaze control [39], eye-hand coordination using an ego-centric "gaze space" [40], and a schema memory system for the learning of actions and action sequences [41]. A new contribution is our use of the LWPR algorithm [42] to learn a model of how movements of the torso affect the robot's representation of space. We find this can be used to reduce the complexity of the redundancy problem in early reaching, and provides a first step towards exploring the representation of movement by repositioning the robots workspace space in the world. We also report the first use of our schema framework on the iCub robot. This includes extensions to enable grasping and transport of objects. In concert we demonstrate how these mechanisms can be combined to create a model of the development of infant hand-eye coordination and object reaching over the first 6 months of life. Furthermore, we show how this model enables very fast on-line learning on an iCub robot.

In the following sections we first briefly outline the development of reaching in infancy, then we describe our implementation of reaching skill growth on our robotic platform, and then present some results and interpretations. The results (Section IV) follow the structure of the design (Section III) thus allowing each component of the model to be considered in turn.

II. The basis for a developmental model

We base our models on an analysis of behavioural development in the infant (as summarised in Table I).

In the first few months from birth infants orientate to sounds and attractive visual stimuli. They make ballistic attempts to reach towards stimulating targets but usually fail to make contact. This "pre-reaching" behaviour leads on to successful contact with objects at around 15 weeks [43], [44]. During this stage it seems that infants do not view the hand during reaching and vision is only used for target location [45]. This means that proprioception is important for arm guidance and it seems that proprioceptive development in the womb provides a more mature, although possibly incomplete, spatial framework by the time visual space is first experienced [20].

Limb movements are jerky for much of this early period. The cerebellum appears to be responsible for the production of smooth action but is very under-developed at birth. This is believed to be the cause of the marked under-damped oscillations of the arm, which gradually reduce as the cerebellum matures (over the relatively long period of 2 years).

Before 4 months there is no independent control over the fingers and grasps are formed only after contact as haptic experiences. Hand control for grasping develops later than reaching. This is an example of the cephalocaudal direction of development that is so prominent in infants [38]. It is also seen in early reaching, which involves trunk and shoulder movement, but with fingers locked. This principle of distal freezing of motor systems is an important feature and is a significant way of solving the problems associated with multiple and redundant degrees of freedom [46].

Only after 8 or 9 months does object size really affect approach and grasping. From this point the visually sensed object size modulates the hand aperture. Also at this age, the shift from proximal to distal control of reaching is started. It seems this is not solely due to maturational change but the trajectory of development depends heavily upon experience and patterns of behaviour [45], [47].

Another contribution to the mastery of arm control is the use of stereotypical motor patterns that have the effect of reducing the number of degrees of freedom during the early stages. By close coupling groups of muscles it is possible to reduce the number of control variables while producing a set of effective space covering actions [48]. It has also been observed that humans have a tendency to avoid extremes in arm configurations, probably because such positions considerably reduce the options for the next move. Similarly, it has been shown that older children and adults adapt their initial pose and grasp for the final arm configuration in an action task [49]. For example, subjects will choose a grasping configuration on a handle such that their hand ends up in a non-awkward position when releasing or using the object. These considerations imply that any constraining or cost function intended to reduce the redundancy problem caused by superfluous degrees of freedom should be applied to the final configuration, not the starting configuration.

From these findings we note that a view of the hand is much less important when reaching than when grasping objects or during other manipulations. Grasp learning can be seen as a separate skill from reaching because it follows on from successful reaching and involves learning new skills covering object properties (affordances), finger control, tactile and other related experiences. This observation, based on earlier infants, who don't have much grasp control (i.e. use of fingers), suggests that proprioception may provide enough information for the earliest reaching actions, and therefore we only need to consider visually-elicited, not visually-guided, behaviours in our model.

Drawing on these considerations, we have performed a longitudinal experiment that models infant development from birth to six months, that is, from a start point with no coordination between any of the sensorimotor systems up to skilled spatial reaching and simple object interactions. This is to test our hypothesis that the key factors listed in Section I should produce a learning process displaying cumulative and staged growth of competence, which (a) is very effective, rapid and flexible, and (b) shows increasing richness in motor behaviour concomitant with the growth of complexity in experience. As behaviour is the outcome of interest it will often be appropriate to measure qualitative changes in patterns of behaviours and their duration, for example, the latter point (b) will be satisfied if motor babbling behaviours evolve into distinct stages or skill levels.

TABLE I
INFANT DEVELOPMENT OBSERVATIONS

Age (months)	Observed sensorimotor behaviours
prenatal	Grasp reflex [32]
	Arm babbling in the womb [16]
1	Sufficient muscle tone to support brief head movements [43]
	Eyes and head move to targets [33]
	Saccades are few in number [23]
	Hand-mouth movements [50]
	Directed (to the hemifield in which a target appears), but unsuccessful, hand movements [51], [24]
	Initial reaching is goal directed, and triggered by a visual stimulus, but visual feedback is not used to correct movements mid-reach [11, p.38]
2	More saccades [23] and improved control [43]
	Head only contributes to larger gaze shifts due to lack of muscle tone [52]
	Involuntary grasp release [43]
3	Head contributes to small gaze shifts 25% of the time, and always to large gaze shifts [52]
	Reach and miss [53] with some contacts [43]
	Hand regard and hands to mouth [43]
	Clasps and unclasps hands [33]
	Infants often move their hand to a pre-reaching position near the head before starting a reach [54], which then follows the line of sight [29, p.44]
	Infants engaged in early reaching maintained a constant hand-body distance by locking the elbow, and instead used torso movements to alter the distance to targets [54]
	Appearance of successful reaching [53], [43], [54], [55]
	Gaze still focused on the target and not the hand [56], [57], [58], [32]
4	Good eye and head control [43]
	Beginning thumb opposition [28]
	Infants begin to use visual feedback to refine the movement of the hand [59]
	As infants age their reaching becomes straighter, with the hand following the shortest path [60]
5	Rotation in upper trunk [43]
	Palmar grasp [43]
6	Successful reach and grasp [33]
7	Thumb opposition complete [28]
8	Pincer grasp, bilateral, unilateral, transfer [43]
	Crude voluntary release of objects [43]
9	Leans forward without losing balance [33]

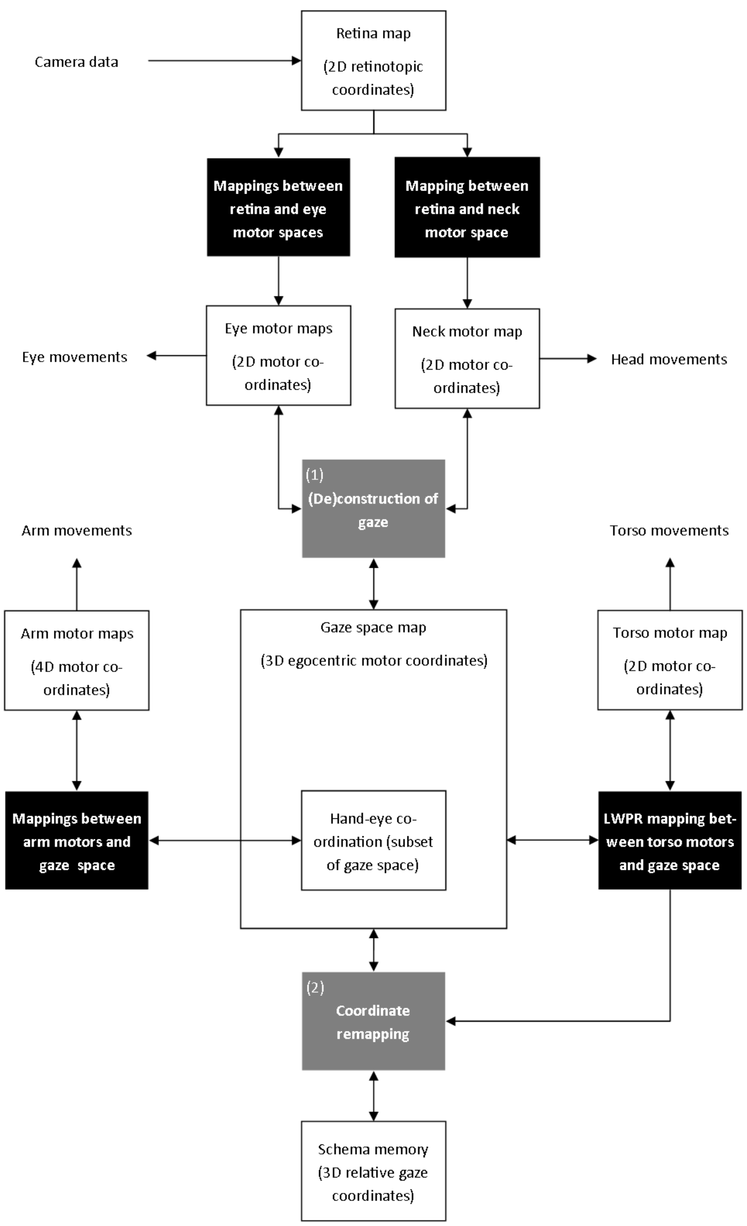


Fig. 1. System architecture showing the elements used to learn the coordination of behaviours described in these experiments. Elements at the top correspond to gaze control; those to the left to reach control; those on the right to torso control; those at the bottom to the schema memory; all interacting through the common gaze space. White boxes indicate data structures (topographic field maps representing sensor and motor spaces, and schemas for action memory), whilst black boxes indicate the learnt mappings between them (all mappings use the mapping framework described in [61], except the torso mapping, which uses LWPR [42]). Arrows indicate the direction of data flow, consisting mainly of spatial coordinates or motor values. Grey boxes indicate additional computational mechanisms, described in the text, that are required to perform coordination not covered by the sensorimotor mappings. See main text for further details.

III. Experimental design

Figure 1 shows the sensorimotor architecture underlying the experiment presented here. It includes motor and proprioceptive spaces for the eyes, neck, torso, and upper arm and elbow; the visual space of the retinas; a schema memory; and an egocentric spatial representation constructed from a combination of the aforementioned systems.

The architecture utilises our mapping framework [61], [62]. Sensorimotor spaces are represented as arrays, or *maps*, of overlapping circular *fields*. These structures are analogous to topographic sheets in the brain [63], [64], which are ordered by collections of receptive neurons. In our simple representation, fields partition the sensorimotor space into areas of equivalence: any point on the map that is stimulated is measured as occurring at the centre of the encompassing field(s). Hence, the arrangement of fields is a crucial consideration (see [62] for a detailed analysis of overlapping field structures). Links, or *mappings*, connect fields on corresponding maps enabling coordination of sensor and motor acts. For example, a stimulated field on the retina map will be mapped to a fixating saccade vector in motor space, whilst an arm pose in joint space will be mapped to a hand position in the egocentric space.

Mappings are not pre-structured, but learnt by the robot through action in the environment. If two fields on different maps are co-stimulated, for example if the arm moves the hand into view thus stimulating both visual fields on the retina and motor fields for the arm joints, then an explicit link will be created between the relevant fields. Thus, a fully developed mapping will consist of a bundle of links between two sensory or motor maps. These mappings also contain a weighting indicating their reliability.

A simple measure of novelty drives the robot to investigate these relationships. Any new stimulus (e.g. the first appearance of a sensory value) is given a high excitation level, which declines with both time and repetition. So, a movement of the arm may result in a previously unexperienced stimulation of a visual field. If this is a new event the stimulus will have high excitation and this will cause the robot to repeat the movement in order to reproduce the relationship. If it is successful, i.e. can be repeated, the weighting between the activated fields will increase, and vice versa. Figure 2 shows the algorithm for this process. A simple winner-takes-all decision is used to select the highest excitation and so the robot may become distracted towards the newest events or discovered behaviours.

When there is a lack of novelty to drive action, which includes at the outset of the experiment, we use spontaneous action to stimulate learning. A global measure of excitation across a map gives an indication of engagement/boredom and this can trigger spontaneous action when there are no stimuli of interest. We have experimented with various excitation and habituation functions for stimuli [65] but find that the parameters are not critical; they simply affect the level of focus on, or distraction by, new events. Sufficiently interesting behaviour is produced by simply selecting the most excited stimuli (above a general background of decaying excitation) and an effective threshold for the algorithm in Figure 2 is $h = 0.6$, i.e. 60% of maximum possible excitation. Initially these spontaneous actions will appear to be random, though they are structured by existing constraints on joints and movements. As learning progresses, actions are further shaped by experience, giving rise to infant-like motor babbling. This is particularly evident in the repetitive novelty-driven exploration behaviour.

Our architecture consists of 4 main components relating to gaze control, arm control, torso control, and long term memory. All of these interact through the *gaze space*, an egocentric map of the space surrounding the robot and it's actions within it. Learning is conducted through our developmental framework using constraints to restrict learning of sensorimotor mappings [61].

A. Gaze control

Following earlier experiments [66], [67], [40], we allow eye saccading and eye/head coordination to develop independently of the construction of a proprioceptive mapping of limb space. This follows from observations that (a) proprioception emerges, in the womb, before vision, and (b) eye movement is the only reliable motor system for the newborn.

The gaze component consists of the eyes and head, and learns how to direct both to fixate on a visual stimulus appearing on the retina. The motor maps are 2D pan-tilt spaces corresponding to joints in the eye and neck, and these are mapped to x-y positions on a 2D retinotopic map corresponding to the camera sensor. Reflexes in the human gaze system prevent the eye remaining stationary as the head moves, which prevents direct mapping of head movements to changes in visual positions. To overcome this an ocular reflex keeps the eye fixated on the target as the head moves, and the impact of head movement on gaze direction is determined from the corresponding eye mapping [39].

The eye and neck maps are incorporated into a model, based on biological data, which generates gaze shifts with stereotypical contributions from both systems. For small gaze shifts, G, the head makes no contribution, and the total gaze shift is performed by the eye. That is the case when $G \leq \alpha + p$, where α is the maximum eye movement without head contribution, and p is the initial eye position contralaterally to the direction of the gaze shift.

For large gaze shifts ($G > \alpha + p$) the gaze decomposition is given in Equation (1). The gaze shift, G, is made up of contributions by the eye, $E(G, p)$, and head, $H(G, p)$, which are both functions of the gaze shift. Here, D is the maximal gaze amplitude (dictated by the detectable angle on the retina), and β is the maximum eye offset during large gaze shifts.

$$G = E(G, p) + H(G, p) \tag{1}$$

where,

$$E(G, p) = \alpha + p + \frac{\beta - \alpha}{D - p - \alpha} \cdot (G - (\alpha + p))$$
$$H(G, p) = \frac{D - p - \beta}{D - p - \alpha} \cdot (G - (\alpha + p))$$

This composition process is performed in grey box (1) in figure 1 and full description of the decomposition is presented in [68].

Begin action
  If Global Excitation = low (no fields excited above threshold h)
    Motor-values := **Select**(random motor values);
      or (extract from any previously learnt {sensor, motor} pairing)
    goto Perform action

  If Global Excitation = high (some fields are excited above threshold h)
    Sensory-values := Select(stimuli with highest excitation)
    Motor-values := Retrieve from {sensor, motor} pair in mapping
    goto Perform action

  Perform action
    Use Motor-values to perform a motor act
    Receive all resultant sensory inputs
    Search for novel states or values (previously undetected values in fields or mappings)
    If Novelty detected
      Set high excitation value for novel stimulus
      Record {sensor, motor} pair with initial weighting in mappings
    Else
      If retrieved action **then** increment Hebbian weight of stored {sensor, motor} pair
Repeat (goto Begin action)

Fig. 2. Algorithm for novelty-driven action selection (derived from experiments in [38])

The gaze direction is of utmost importance in our architecture, as it is this that defines the *gaze space* - the 3D egocentric model of space used to coordinate the robot's actions. Azimuth and elevation are provided by the sum of the pan and tilt angles of the eye and neck joints, while depth is provided by the vergence angle between the two eyes.

B. Arm control

The arm component relates movements in the shoulder and elbow joints to movements of the hand in the visual space. We note that the eyes of the fetus do not open until 26 weeks after conception, and that any vision is likely to be very limited. However, arm movements and tactile perception appear at 7-9 weeks, and there is the possibility for early learning through proprioception and tactile feedback. We are currently working on a model of reach development that reflects the gradual mapping of proprioception to vision through several stages identified in the infant literature [69], but in this experiment we use a simple 2-stage model of reach development: firstly a series of arm movements are generated with proprioceptive information to model movements learnt prenatally; secondly, vision is activated, and motor babbling is used to explore these stored movements, resulting in their mapping to the gaze space.

We note that in early reaching infants will lock their elbow, and reaches will often start and end with the hand near the head or mouth [54]. Therefore, for the growth of the proprioceptive reach space we arranged that the arm would have restricted movement on the distal joints (elbow and below), and a "rest" position was defined, with the arm retracted and the hand near the head. A reach action consisted of a movement from the rest location to a spatial position in front of the robot, and a range of target locations were generated for the volume of space around the robot by motor babbling (to facilitate grasping, an artificial constraint limits learning of arm movements to those that result in the hand in a horizontal orientation). To prevent damage to the robot this learning is currently conducted in simulation and potential reaches are then transferred to the robot.

The developmental constraints used in the experiment are listed in sequence in table II. As mentioned above, the 2-stage reach constraint allows the arms to move first (in their proprioceptive stage) before vision is activated and hand-eye coordination can begin. Arm movements with vision facilitates the interaction of hand and eye behaviour known as "hand regard" activity [70]. This behaviour helps to coordinate the visual gaze space with the proprioceptive space by matching data pairs between the seen location and the "felt" location of the hand. Up to this point progress has been very similar to our previously described experiments [66], [67], [40], [39]. The other constraints in table II are discussed in the following sections.

The criteria for lifting constraints is always based on the level of maturity reached by the current stage. These can vary in form as they need to be appropriate for the modalities involved. We have previously used the concept of *saturation* as a trigger for constraint lifting [61]. For example, for the mappings, a measure of field density can provide an indication of how complete the mappings have become, but for grasping objects it is necessary to monitor the schemas and determine when sufficient generalisation has occurred. It is possible to automatically connect such saturation measures to the lifting of constraints but in this experiment the release times were chosen by the experimenter by observing the level of behaviour and noting performance data such as map densities. The times when the constraints were released are also shown

TABLE II
CONSTRAINTS USED TO STRUCTURE BEHAVIOURAL STAGES ON THE ROBOT

Constraint	Effect	Removal trigger	Release time (mins)
Eye motor	Prevents eye motion	Start of experiment	0
Neck motor	Prevents head motion	Threshold on eye control	20
Arm motors	Prevents arm movement	Threshold on gaze control	60
Hand-eye coordination	Limits visual coordination for reaching	Threshold on proprioceptive reach map	90
Reflex grasp	Causes hand to close on tactile stimuli	Active until reaching threshold attained	100
Controlled grasp	Prevents voluntary grasping of objects	Threshold on hand-eye coordination	110
Torso motor	Prevents motion at waist	Threshold on hand-eye coordination	120
Schema memory active	Prevents temporal memories	Threshold on grasp control	130

in table II.

C. *Torso control*

At this stage of development the robot is able to reach to a gaze point and look at a hand position. But notice that the gaze space is much larger than the reach space. This is mainly because the maximum reach is determined by the arm length which is much less than the visual range. Another important point is that the reach and gaze geometry are closely coupled in the sense that they are both grounded or referential to a point on the body centre-line somewhere near the neck. This means that, regardless of the configuration of the rest of the body below the shoulders, the mapping between eyes and hands will remain constant. If a stimulus is seen to be within the reachable range of the gaze/reach mappings then it can be reached. Conversely, if a stimulus is unreachable (i.e. seen but has no mapping into reach space) it may be possible to bring it within reach by repositioning the upper body. This effect can also stimulate the recruitment of locomotion to achieve distant desirable goals. However, as locomotion is not yet available, we use torso movement (which develops early, [43]) to extend the reach space.

Of course, a stimulus may be within the gaze space but beyond any possible reach. In this case the torso mapping process will not find a solution and, after learning, the system will cease attempting to reach to such points. This relates to experiments that show that humans develop clear distinctions between their peripersonal and extrapersonal space [71]. Interestingly, the failed reaches to extrapersonal space during map learning appear to an observer as pointing actions [72].

On the iCub, we find that the joints in the eyes and neck are all close enough to the retina that they can be mapped independently, without incurring noticeable errors. However, this is not the case with joints at the waist. Here, there is a noticeable dependency between the effect of tilt (forward) and rotation (about the body centre line) movements on gaze shift. Consequently, our standard mapping technique is not suitable for representing this relationship. For this reason we have used an implementation of the LWPR algorithm, which is well suited to learning the inverse kinematic mapping [42].

LWPR is an algorithm for incremental learning in high dimensional spaces with redundant, sparse data. It is based on the observation that complex high dimensional data sets often contain locally low-dimensional structure. The method uses regression techniques to learn locally linear models in low-dimensional areas of the input space, with a weighted gating mechanism to select model contributions to the output. The output is, therefore, effectively built up from piecewise contributions about local space.

Although not biologically plausible, due to the extensive tuning or prior knowledge needed, the idea of learning localised models to control a higher-dimensional system resonates with our field-based approach to sensorimotor control [61]. Our investigations require methods with low computational overheads and with rapid learning from sparse data, and therefore have some compatibility with the theory behind LWPR.

We use the LWPR algorithm to learn the relative torso rotation and tilt required to move a target in the gaze space from one location to another. To train the model we position a target in front of the iCub and perform motor babbling at the waist joints. The gaze controller is used to keep track of the target and report its position. Data relating to the different waist joint positions and gaze coordinates are used to train the LWPR model. This is repeated with the target in several different locations.

We use the MATLAB implementation of the LWPR algorithm[1], using the model parameters given in Table III. Inputs to the model are in the form of an 8-value array whose elements correspond to: the position of a target in the gaze space (pan, tilt, vergence), the position of the torso (rotation and tilt), and the desired position of the target in the gaze space (pan, tilt, vergence). Outputs are in the form of data pairs corresponding to the change in torso rotation and tilt required to present the target at the desired location. To generate the values in Table III we measured the variance of the input data taken from the robot (in radians), and tuned other parameters by trial and error.

When the torso mapping has been developed, it is possible to reach to a target in a two step process: use the torso map to bring the target into a reachable location in gaze space; then use the gaze/reach mapping to generate a reach action.

An additional feature has to be added to the system when torso movement is involved. Any remembered references to object locations (in the schema memory) must be in terms of some body framework, and we use the upright body centre

[1] Available from http://wcms.inf.ed.ac.uk/ipab/slmc/research/software-lwpr

TABLE III
MATLAB PARAMETERS USED TO INITIALISE THE LWPR MODEL FOR TORSO CONTROL. I_n IS THE IDENTITY MATRIX AND J_n IS THE UNIT MATRIX.

Parameter	Value
nIn	8
nOut	2
init_D	$I_8 * 25$
norm_in	$[2.47\ 1.72\ 0.56\ 1.41\ 0.93\ 2.47\ 1.72\ 0.56]^\top$
init_alpha	$J_8 * 250$
diag_only	0
w_gen	0.2
w_prune	0.7
meta	0
meta_rate	250
update_D	1
kernel	'Gaussian'

line. Thus, any remembered coordinates are based on 3D egocentric coordinates with the torso centred. However, the torso may be off-centre and a coordinate transform is then necessary to relate back to the reference frame. This coordinate remapping process is performed in box (2) shown in figure 1 and uses values from the LWPR torso mapping to adjust the coordinates such that they remain relative to the upright body centre line in the gaze space, regardless of the current torso orientation.

D. Manipulation

As described in section II, very early reaching behaviour arises before any hand control has been established and so we set the hand to be normally open with the fingers flat. If the front of the hand makes good contact with an object then an automatic finger close is executed. This provides a kind of grasp reflex which is maintained, even while the iCub performs other actions, and is only released by removal of the object, either by accident or external interaction. Unlike object contact, the release is not a significant sensed event.

As a result of the earlier hand regard behaviour the system is able to spatially correlate visual stimuli with hand positions and vice versa. Thus, when an object is presented for the first time it is likely to be detected in periphery vision and a saccade will bring the object to fixation. This fixation location in gaze space will stimulate a corresponding target for a reaching action and a reach will be initiated. At this early stage it would be expected that some reaches would miss the object and others would contact it. Some of those that make contact will also grasp the object through the grasp reflex. In accord with infant stereotypical motor patterns [73] the reach actions are completed by a return of the arm to the "rest" location near the body. (A canonical example of such quiescent "home" positions is the hand at the mouth, and mouthing is almost a default behaviour for any object acquired by the hand [74].)

After a period of early reaching, the system can be expected to have experienced "disturbing a stimulus" (by moving it or knocking it completely out of the environment) and "holding" (with kinaesthetic and possibly tactile signals). The next constraint to be lifted is the reflexive grasp and we do this by allowing the fingers to close to a given aperture and by activating a "hand empty" sensor (effectively the inverse of tactile contact). Also the release of a grasped object now becomes an experienced event and so this allows objects to be dropped deliberately and thus the sophisticated skill of moving an object from one place to another is now available to learn. We note that more hand control is now possible — small movements of the fingers can be related to properties of objects and better grasps can be produced by matching the aperture to objects — but these are outside the scope of our experiment.

E. Memory

The egocentric gaze space shares a reference point (the body centreline) with the reach space. This supports the cross-modal mapping of reaches to fixation points and the gaze space becomes the central reference for spatial operations, as seen in Figure 1. The gaze map has an additional function: it also provides a space in which to represent events and experiences in the robot's environment. Each field in the gaze map is able to store the values of tactile or visual features that have been sensed at that location. Thus as objects (or other environmental features) are seen or touched, simple local features (colour, contact) are entered in the maps. Consequently, the mappings collectively hold the currently perceived world state and this effectively includes the existence and location of objects. However, these are short-term memory traces because they have saliency functions which cause them to decay as described in Section III. Thus objects can be forgotten, and if they appear again after a long period they will be treated as "new" objects. The mapping system hence acts as a short-term memory and can be likened to infra-cortical mechanisms, where attention selection, action selection, salience, and priority functions are performed.

In contrast, long-term memory of experience is provided for by our schema mechanism [41], [75]. This operates on all the available current sensorimotor data from the robot's environment and short-term memory, which effectively provide a "world model". The mechanism is inspired by Piaget's notions of schemas [76], and stores action representations as triples consisting of: the pre-conditions that existed before the action; the action performed; and the resulting post-conditions. The pre-conditions are a representation of the world state before an action is performed, and are made up of a set of *observations*. For example, a visual observation may include the spatial position of an object and its colour. Similarly, the post-conditions are a set of observations made after an action has been performed. The action itself is constructed by following mappings.

The mappings learn the relationships between sensory modalities, whilst the schema memory learns the temporal effects of actions. These actions can range from simply reaching to an object in the environment, through to building a stack of objects. The information on how to *perform* the action, such as the reach or gaze is stored in the underlying mappings, and referenced from the schemas for the current conditions. For example, a new object may be placed in the workspace. The camera data feeds into the gaze control system and coordinates in gaze space, along with visual data about the object, are passed to the schema memory, adjusting for any coordinate offset caused by the torso (see Figure 1). The schema memory identifies the object as new, which it is stimulated to explore due to it's high level of novelty. This starts by triggering a gaze action that employs the gaze control subsystem to fixate on the target using the eye and head sensorimotor maps. A reach action is then triggered by the schema memory, which uses the coordination between the gaze space and reach space in the sensorimotor mappings to select an appropriate reach configuration, moving the hand to the target. Further basic actions such as press or grasp can then be triggered by the schema memory based on currently stimulated schemas.

Our schema implementation has a number of important features including: creation algorithms that produce new schemas from the merger of past and current experience; an excitation device to modulate schema selection; a generalisation algorithm that generates abstract schemas from a small number of examples through parameterisation; and a schema chaining mechanism for formulating and executing multi-part actions.

Rather than having explicit goals that the robot tries to achieve, we use an intrinsic motivation device to drive execution of actions relevant to novel experiences in an attempt to learn new ways of interacting with the world. The schema system attempts to select the schema which has the most in common with the novel aspects of the environment currently being experienced, in effect being reminded of actions that it had previously used in a similar, but not identical, context. This is achieved by schemas being excited according to a combination of the novelty of the current experience and the similarity to past experiences; as given by partial or complete matches. Thus the level of excitation is determined by the novelty of the sensation combined with the similarity to a remembered sensation. Together with the cumulative learning approach this intrinsic motivation system often results in the selection of actions learned in earlier stages of development that have a high likelihood of being relevant to the aspects of the environment that had previously not been experienced.

The generalisation algorithm is a significant feature of the schema system. Beyond simply determining which aspects of the schema may be interchangeable with other values, as many existing schema systems do, this mechanism attempts to find generalisable relationships between the preconditions, the action, and the post-conditions of a schema. This allows the generalised schemas which are produced as a result of this process to represent the agent's hypotheses about how an interaction may work at a more abstract level. Generalisation allows the creation of "target actions" which consist of a list of observations that should be achieved by any schema implementing that action.

Schema chaining is the process of linking pre-conditions and post-conditions from different schemas, and this creates a traversable network representing different world states and the actions required to move between them. Without schema chaining the robot's interest in unreachable objects would decrease as it failed to reach them. Schema chaining allows for cases in which the feedback of an action is not instantaneous to still be recognised as being useful. This allows an entire series of actions to be interesting to the robot. Dijkstra's algorithm is used to resolve alternate paths in the network of candidate schemas by taking account of the path weights that record the probability of past success, thus allowing the system to determine the shortest chain of actions required to achieve a goal in the manner most likely to succeed. Thus selection in a chain is driven by a combination of a schema's ability to contribute towards the current goal and its reliability judged by past experience.

During the execution of a chain the robot evaluates the success of the predictions made by each schema in the chain after the action component of that schema has been executed. If the execution of one of the schemas results in an unexpected outcome which is not compatible with the next schema's preconditions then the chaining process begins anew, finding a chain of schemas from the new world state to the target. Continual evaluation of the environment during the execution of a chain allows the system to adapt to both incorrect or unreliable learning and to other agents acting upon or disturbing the environment.

The dynamic nature of our system, allowing major changes during interactions between schemas and features of the world model, during generalisation, and during schema chaining, represents a significant difference between our system and other implementations of the schema idea.

IV. Experimental Results

A. Gaze control

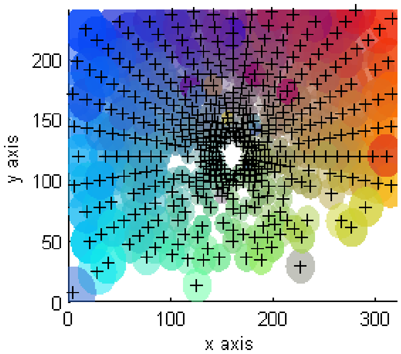
Following the cephalocaudal development of the infant, the robot begins by learning the eye movements required to saccade to a visual target. The first constraint is that only the eyes have sufficient muscle tone to perform reliable actions. Fig. 3 shows the learnt mappings between sensor and motor spaces for making eye saccades, built up by a process of motor babbling. When a stimulus is received on the retina, the mapping provides the associated motor movement parameters, triggering a saccade that fixates on the stimulus.

Next, a constraint is released enabling the learning of neck control. This could be a physical constraint, such as the lack of sufficient torque in the neck, or an emergent constraint, such as the prerequisite for accurate eye saccades as a basis for learning head movements [77]. Fig. 5 shows the learnt mapping between neck muscles and the impact of these on the visual space[2]. Figure 4 shows the iCub robot using learned mappings from both eye and head to coordinate a fixation.

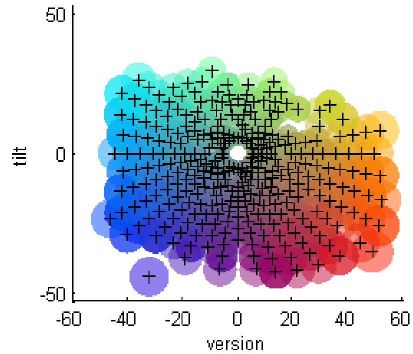
B. Reach control

Reaching movements are mapped onto the gaze space using a combination of motor babbling and hand regard behaviours. Following the literature on early infant reaching, constraints are imposed on the type of reaches possible. In the early stages, the elbow joint is fixed, and "swiping" movements are

[2]A video of the iCub learning eye-head gaze control can be seen at http://youtu.be/DIGpCIzAF_E

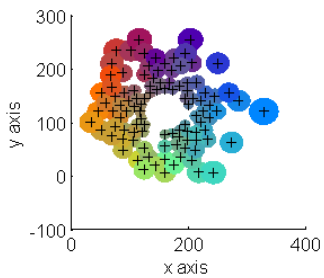


(a) Retinal map

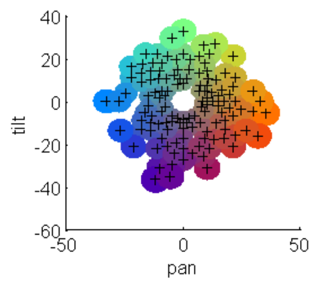


(b) Eye motor map

Fig. 3. Maps for eye saccade control. Coloured circles indicate learnt fields over which stimuli or actions (in the case of motor maps) are considered identical. Matched colours indicate the explicit links between maps (i.e. that comprise the mappings). A stimulus in a visual field will trigger the associated motor movement, allowing the eye to saccade to the stimuli.



(a) Retinal map



(b) Head motor map

Fig. 5. Maps for head contributions to gaze shift. Motor movements are mapped to corresponding shifts in the visual space

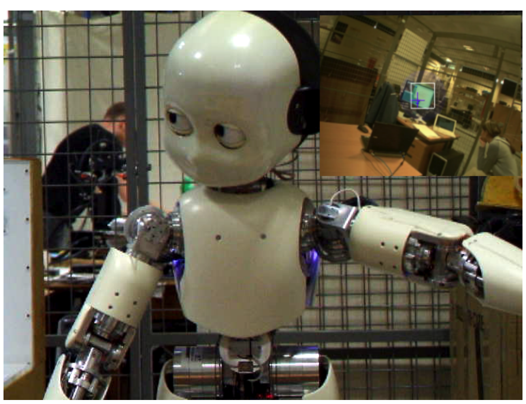


Fig. 4. A gaze involving coordinated eye-head motor action to fixate a target.

made using the joints in the shoulder. Reaches are initiated from the rest position near the head. We note that this “pre-reaching” pose enables the robot to reach to objects on a line similar to the gaze direction, and limits collisions with other objects. The resultant gaze-reach mapping is between two 3-dimensional spaces corresponding to pan, tilt and depth in the gaze space and the proprioceptive space formed by shoulder joints 0, 1 and 2. Fig. 6 shows a 2-dimensional projection of this mapping[3].

C. Torso control

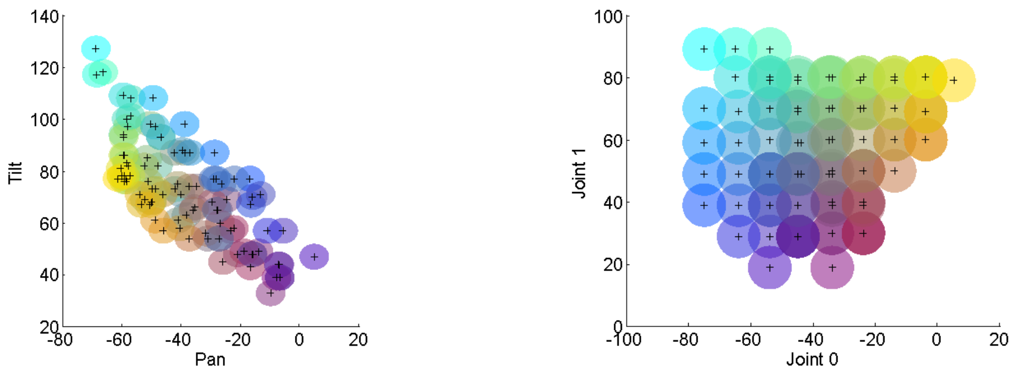
With restricted elbow movement, the range of reach distances is very limited. The infant overcomes this by using movements of the torso to bring objects into range. By rotating and tilting at the torso, the shoulder can be moved closer to, or further from, objects and thus altering the distance for reaching. An added benefit is that movement of the torso has no impact on the eye-hand coordination already established because the kinematic coupling between eye and hand is via the head and shoulders and so this subsystem can be moved as a unit without disturbing its internal structure. This removes additional potential degrees of freedom from the eye-hand coordination problem, making torso learning less complex.

Although we have used an implementation of the LWPR algorithm successfully on the iCub[4], the results in this section have been generated in simulation. In order to best reflect performance on the iCub, learning has been limited to amounts of training that can be easily performed on the iCub, and initialisation parameters have been based on iCub data.

The LWPR model demonstrated here was learnt in 10 cumulative stages, corresponding to learning targets positioned at 10 different locations. For each target the robot performed babbling movements of the waist and recorded the shift in position of the target in gaze space. 55 unique torso movement/gaze shift pairs were generated for each target, and then

[3]A video of the iCub learning the reaching movements described here can be seen at http://youtu.be/_ZIkbU8FZbU

[4]A video of the robotic implementation can be seen at http://youtu.be/OhWeKIyNcj8



(a) Gaze map

(b) Motor map

Fig. 6. Two dimensional projections of reach maps in gaze space (pan, tilt, depth) and motor space (the 3 joints in the shoulder). Joints 0 and 1 refer to the first two joints in the shoulder of the iCub, going from proximal to distal

used to train the LWPR model. After each learning phase the model was tested against 30 known torso movements.

Fig. 8 shows the performance of the model after each training phase. The graphs show the error in fixation point measured in each gaze dimension. The improvement in performance can also be seen in terms of the decrease in the normalised mean squared error, which is shown in Fig. 9. Data relating to the learnt model at each stage is given in Table IV, and shows the number of target positions used in training, the number of postures trained from, and the number of local models learnt in each output dimension. Most improvement occurs on the first few training targets, and we find that the learnt models are sufficient for our needs after just 10 training phases.

When the iCub selects to reach to an object that is out of reach, it first attempts to move the target into the centre of the reach space by a movement of the torso. The robot calls the LWPR model with the current torso and target positions, and the desired end position of the target. Often such a movement will not be possible, and the model will produce a movement that only brings the target part way towards the goal. However, this is sufficient provided the target ends somewhere within the reachable space, as the robot selects a reach based on its final position. If the target is still out of reach the robot will instead point towards its location. Figure 7 shows the iCub robot torso mapping being used to extend the reach space so that the learned gaze mappings can be employed in reaching for a seen object.

D. Memory

By this stage, the robot is capable of gazing to objects, orientating itself to bring the objects into reaching distance, and making reaching motions toward them from the rest position. Using the schema learning mechanism it now starts to build composite actions from these beginnings.

When the robot sees an object it checks for schemas excited by that stimulus and finds that the most excited schema is one in which it remembers seeing its own hand in the location the object now occupies (Fig. 10a). Upon executing this schema

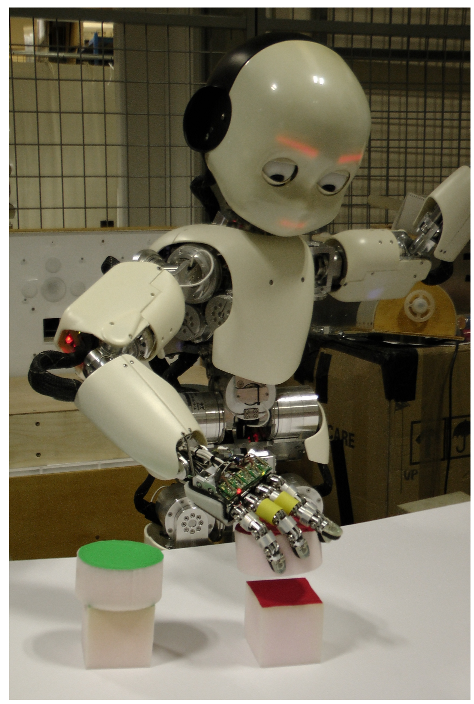
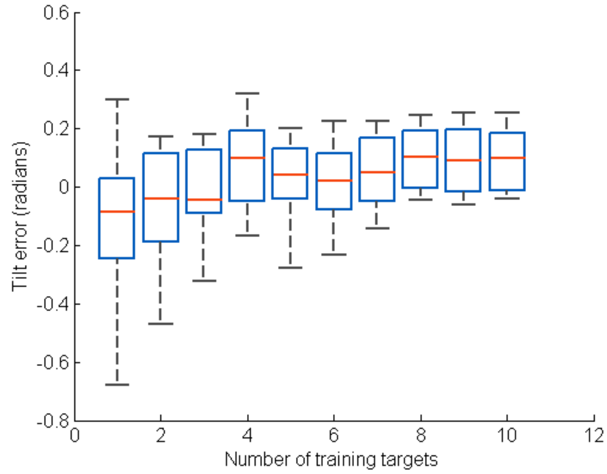
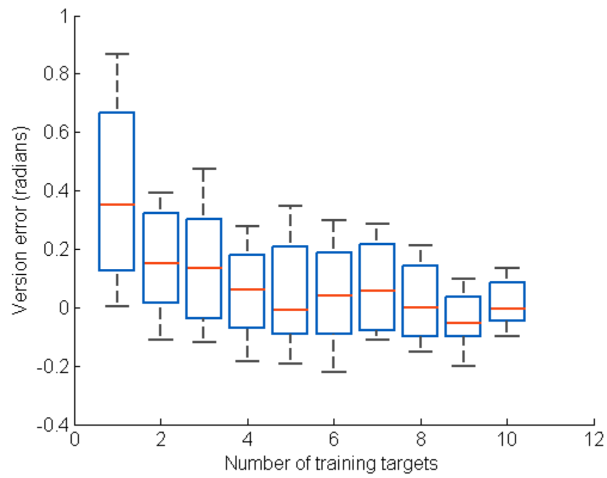


Fig. 7. The iCub robot leaning its torso forward in order to extend the gaze/reach space.

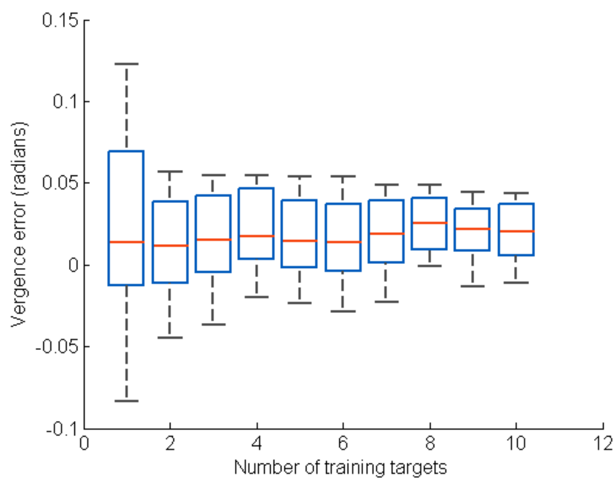
the robot finds that when an object is present in the target location it receives an unexpected touch sensation. A new schema is then formed to represent this knowledge (Fig. 10b), which can then be generalised into a form which represents reaching towards and touching objects in any position (Fig. 10c). The $ symbols identify the bindings between the variables and are created by the generalised schema building algorithm [75]. When generalised schemas are being instantiated the first assignment to a given $ variable is then bound to all other occurrences of that variable. Note that this can happen in any order in schemas. Thus, the schema memory can operate in various ways: it can retrieve candidate actions for possible contexts; it can suggest actions that might lead to desired results; or it can predict the outcome of a proposed action.



(a) Tilt error



(b) Version error



(c) Vergence error

Fig. 8. Error in fixation point at each stage of training the LWPR model for torso control. Each training target corresponds to an additional learning phase, after which the model is evaluated against 30 sets of test data. On each box the central mark represents the median error and the edges correspond to the 25th and 75th percentiles

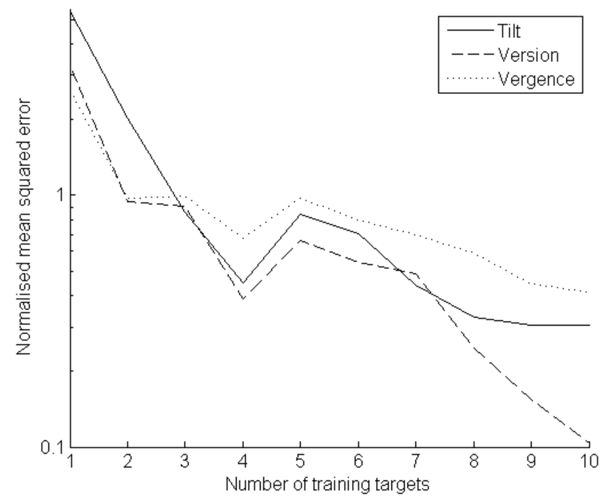


Fig. 9. Graph showing the reduction in normalised mean squared error in each of the gaze dimensions against the number of training targets for the LWPR torso model

TABLE IV
DETAILS OF THE LEARNT LWPR MODELS, SEE TEXT FOR DETAILS

No. target positions	No. training samples	No. local rotation relations	No. local tilt relations
1	55	12	12
2	110	25	25
3	165	26	26
4	220	32	31
5	275	42	41
6	330	43	45
7	385	63	63
8	440	86	87
9	495	117	117
10	550	122	122

Pre-conditions	Action	Post-conditions
	Reach to 35,-66	Hand at 35,-66

(a) Initially excited schema

Pre-conditions	Action	Post-conditions
Obj. 1 at 35,-66	Reach to 35,-66	Obj. 1 at 35,-66 Hand at 35,-66 Touching obj. 1

(b) Extended schema with new information

Pre-conditions	Action	Post-conditions
Obj. $a at $x,$y	Reach to $x,$y	Obj. $a at $x,$y Hand at $x,$y Touching obj. $a

(c) Generalised schema

Fig. 10. The creation of a schema representing the act of touching objects

The new touching schema is then likely to be executed a number of times due to the novelty of the experiences involved. However after a short while the excitation drops below that of the next most excited schema, which in this case is a grasping schema. The grasping schema is excited by the

Pre-conditions	Action	Post-conditions
	Grasp	Touching hand

(a) Initially excited schema

Pre-conditions	Action	Post-conditions
Obj. 1 at 35,-66 Touching obj. 1	Grasp	Obj. 1 at 35,-66 Hand at 35,-66 Holding obj. 1

(b) Extended schema with new information

Pre-conditions	Action	Post-conditions
Obj. $a at $x,$y Touching obj. $a	Grasp	Obj. $a at $x,$y Hand at $x,$y Holding obj. $a

(c) Generalised schema

Fig. 11. The schema memory learns to go from touching to grasping objects

Pre-conditions	Action	Post-conditions
Obj. $a at $x,$y	Reach to $x,$y	Obj. $a at $x,$y Hand at $x,$y Touching obj. $a

Pre-conditions	Action	Post-conditions
Obj. $a at $x,$y Touching obj. $a	Grasp	Obj. $a at $x,$y Hand at $x,$y Holding obj. $a

Fig. 12. Chaining of touching and grasping schemas

memory of the tactile sensation caused by the robot closing its own hand when performing a grasp with no objects present, which it is reminded of by the touch sensation it receives from the object it has reached towards (Fig. 11a).

Executing the grasp schema whilst touching an object results in the robot successfully grasping the object and receiving the sensation of holding an object. A new schema is then created to represent this new information (Fig. 11b). As with the new touching schema this grasping representation can also be generalised as shown in Fig. 11c, which represents the act of grasping an object in any location.

With schemas for reaching to, touching, and grasping an object, chaining can be used to form a plan of action which allows the robot to reach towards and then grasp an object at any location (Fig. 12).

Table V charts discovery of the above schemas during a learning session[5]. Initially the robot has access to the primitive sensorimotor actions contained in the learnt mappings, which include gazing and reaching. It also has a preprogrammed reflex grasp. The requirements for generalisation have been set to the minimum necessary for the simple setup to enable fast learning.

The experiment begins with the robot being presented with a green object. The object is the most salient stimuli, and so the robot selects the available gaze and reach actions to perform on it. At 0:18, the robot has reached to the object and receives tactile feedback. This new event results in the generation of a schema for touch, which represents reaching and touching green objects at that location. Excitation of this new schema causes repetition of the action, but noise causes some variation leading to the generation of a generalised touch schema at 0:50. This schema applies to touching any coloured object at any location, and becomes the most excited option for further exploration. By 1:45 the excitation of the touch schemas have dwindled, and the grasp action becomes most salient (this is due to the similarity between the existing touch sensation and the recollection of the touch sensation triggered by closing the hand on itself). At 1:56 the robot generates a new schema for grasping a green object at that location, and this is quickly followed by the generalised version due to the similarity with the existing generalised touching schema. The robot cannot re-grasp, and so reaching again becomes the most excited action. At 2:19 the robot has moved its hand to a new position whilst still holding the object. This creates a new schema for moving an object that, following more repetition, becomes generalised at 2:36. At 3:32, after further repetition, the most excited option becomes the press action. This is particularly interesting as the robot is still holding the object, and provides the opportunity for learning basic tool use. However, in this instance the action does not cause a change in the world state, so no schema is generated. Finally, the release action becomes most exciting, and so the robot drops the object, learning the "release" schema.

This completes the process of attaining visually elicited reaching. Learning is driven by novelty in the early stages, giving way to goal directed behaviour only when suitable goals have been found through exploratory motor action, or "play". Table VI outlines typical times at which each behaviour first appears, with stricter generalisation requirements suitable for more complex learning. The durations of the behavioural stages is also shown and it can be seen that the behaviours involving more degrees of freedom (e.g. the arms) need more learning time than those with fewer (e.g. the 2D retina and the 2D eye motors). The experimental sequence shows cumulative learning of skills from primitive sensorimotor coordination to action planning. Fig. 13 shows the iCub completing a successful reach and grasp, and (shortened) videos of all the developmental learning can be seen at http://www.aber.ac.uk/en/cs/research/ir/robots/icub/dev-icub/. A key indication of the power of this approach is that the whole developmental sequence described here can be run on the iCub robot in just under 3 hours.

Just as the timing of behavioural stages varies between infants, so does it vary between experiments on the iCub. In the early stages, where sensor and motor activity is being coordinated, learning is affected by variation and noise in stimuli and motor babbling. In the later stages, the trajectory of schema development is dependent on the learnt primitive actions, the initial excitation of schemas, and the environment. Therefore trajectories can, and do, vary in their appearance.

A critical issue is the scheduling of the release of constraints. In connected work we have investigated how the timing of constraint release impacts on learning of gaze control [39], [77]. Those results show a trade off between timing of constraint release and the rate of learning. If there are no sequencing constraints, then sub-systems are allowed to learn in parallel and learning is found to be slow, due to

[5] A video of this sequence can be seen at http://youtu.be/3zb88qYmxMw

TABLE V
SCHEMA DISCOVERY ON THE ICUB

Time (mm:ss)	Preconditions	Action	Postconditions	Description
00:18	Green object at (17.5, 72.4)	Reach to (17.5, 72.4)	Hand at (17.5, 72.4) Green object at (17.5, 72.4) Touch sensation	New touch schema
00:50	$z colour object at ($x,$y)	Reach to ($x,$y)	Hand at ($x,$y) $z colour object at ($x,$y) Touch sensation	Generalised touch schema
01:56	Green object at (17.5, 72.4) Touch sensation	Grasp	Hand at (17.5, 72.4) Green object at (17.5, 72.4) Holding object	New grasping schema
02:01	$z colour object at ($x,$y) Touch sensation	Grasp	Hand at ($x,$y) $z colour object at ($x,$y) Holding object	Generalised grasp schema
02:19	Hand at (17.5, 72.4) Green object at (17.5, 72.4) Holding object	Reach to (8.8, 62.6)	Hand at (8.8, 62.6) Green object at (8.8, 62.6) Holding object	New transport schema
02:36	Hand at ($x,$y) Green object at ($x,$y) Holding object	Reach to ($u,$v)	Hand at ($u,$v) Green object at ($u,$v) Holding object	Generalised transport schema
03:42	Hand at (17.5, 72.4) Green object at (17.5, 72.4) Holding object	Release	Hand at (17.5, 72.4) Green object at (17.5, 72.4) Touch sensation	New release schema

TABLE VI
BEHAVIOUR DEVELOPMENT ON THE ICUB

Behaviour	Description	Time of first appearance (mins)	Duration (mins)	Duration of stage (mins)
Saccading	Accurate, direct moves to peripheral stimuli	3	17	20
Gazing	Coordinated eye and head moves to fixate on stimuli	35	25	40
Arm movements	Hand regard - gazes at hand	62	58	60
Torso movement	Moves at waist whilst accurate gaze at objects	125	15	20
Object play	Reaches to, touches, grasps, and moves objects	143	27	30
Learning ends	Experiment ends	170		

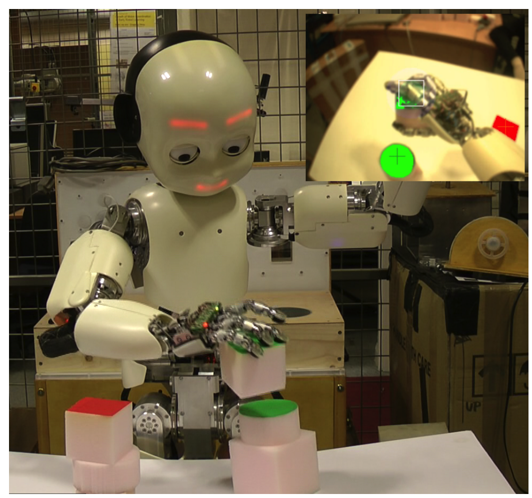


Fig. 13. The iCub humanoid robot completing a grasp (inset showing the robot's view)

added physical and computational complexity. As a result, connectivity between maps is sparse. Alternatively, if constraints remain in place for a prolonged period, learning of the unconstrained system is initially fast and connectivity is high, but at the expense of improvement in the constrained system. By releasing the constraint on a sub-system at an intermediate time, the learning rates of mappings in both systems are increased. Preliminary results suggest that the optimal time to release a constraint to maximise learning depends on the interaction of the codependent learning rates of the systems involved. This is a matter for further investigation.

V. CONCLUSIONS

A major challenge in robotics is the achievement of true cognitive autonomy. That is, the ability to learn, from novel experiences, new behaviour that is useful for achieving new goals and skills. The challenge is in being able to learn without either supervision, *a priori* knowledge, tuning, extensive training, or other forms of preprogramming. Learning must also be cumulative and incremental, with primitive actions being used as building blocks for more advanced behaviours. Additionally, it must be driven by intrinsic motivation because formative experience is gained through autonomous activity, even in the absence of extrinsic goals or tasks. Our experiments take inspiration from early infant development where these conditions apply.

The work described here extends our previous investigations by combining components involving gaze control, eye-hand coordination, the use of constraints to shape learning, and a sensorimotor schema memory system, with new implementations of torso learning and the first application of the schema framework to enable grasping and movement of objects.

The experiment shows how a mapping between the gaze space and the waist joints can modulate the gaze space to shift the reachable space of the robot thus bringing nearby objects into reaching range. This can be seen as an early precursor to locomotion whereby the (local) egocentric gaze space is maintained yet becomes translated into new regions [78]. Infants use torso movements in this way, in combination with freezing of the elbow, which reduces the demands on hand-eye coordination by removing degrees of freedom from the eye-hand chain. In this work we have used the LWPR algorithm, and shown that a suitable mapping can be learnt from a small amount of training data, making it suitable for deployment on a real robot system. Errors in the learnt relationship are not critical, as arm movements can compensate for object position, provided it appears within the reachable space.

We have also demonstrated the deployment of the schema system on the iCub robot. This has enabled the acquisition of new competencies relating to object grasping and interaction. We have described how these schemas are created based on the sensorimotor abilities of the iCub, and shown how they appear as part of a novelty-driven sequence of exploration on the robot. The schemas described here were generated in less than 4 minutes, showing the suitability of this approach to online robot learning.

Our experiment gives a full demonstration of longitudinal development on an iCub humanoid robot. It shows how constraints can structure infant-like behavioural stages, advancing from uncontrolled motor babbling through eye/head control and crude arm control, to eventual skilled hand/eye integrated reaching, grasping and transport of objects. The results display qualitative stages in behaviour and support our hypothesis that constraining the order of development to the cephalocaudal and proximo-distal maturational patterns should produce very effective, cumulative growth of competence. This approach is now being advocated by several researchers (see [79] for a detailed discussion of these constraints and their role in robotic models).

Our other hypothesis concerns the role of motor babbling and its relation to more advanced behaviour. The change from early motor babbling into more purposive looking behaviour is due to the algorithm, which emphasises repetition and coordination. This has been called "goal babbling" by some authors [80], [81] and indicates that motor babbling is not necessarily random action but can be modulated and structured by prior and immediate behaviour. The repeating testing of fortuitous yet correlated events is the difference between goal babbling and random motor action.

The experiment also gives evidence that very simple novelty-directed attention is sufficient to drive early sensory-motor coordination and skill acquisition. Our novelty detection method contains a very general principle that can apply to many levels of babbling and play. This approach to intrinsic motivation is related to the work of the Flowers laboratory [82], [83] who show that mechanisms that attempt to maximize competence cause motivation to focus on increasing complexity and are more efficient than random motor babbling methods.

The results from our integration of these various components into an implementation of longitudinal development on an iCub robot shows their ability to generate infant-like development and object interaction in real time. The results also show very fast on-line learning with successful reaching and manipulation behaviours being produced from a start point with zero coordination in less than 3 hours.

Using these lessons from human development we have built a robotic system that learns to reach in a way that overcomes many of the hurdles to humanoid reaching. Our experiments can be seen to expose some of the "logic" that appears to be behind the infant's development in early sensory-motor learning. We believe this continued approach will offer further valuable models and solutions for robotics.

References

[1] A. Bicchi and V. Kumar, "Robotic grasping and contact: A review," in *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, vol. 1. IEEE, 2000, pp. 348–353.

[2] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: a survey," *Connection Science*, vol. 15, no. 4, pp. 151–190, 2003.

[3] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, "Cognitive developmental robotics: a survey," *Autonomous Mental Development, IEEE Transactions on*, vol. 1, no. 1, pp. 12–34, 2009.

[4] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, "The icub humanoid robot: an open platform for research in embodied cognition," in *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems*. Washington DC, USA: ACM, 2008, pp. 50–56.

[5] O. Sporns and G. Edelman, "Solving Bernstein's Problem: A Proposal for the Development of Coordinated Movement by Selection," *Child Development*, vol. 64, no. 4, pp. 960–981, 1993.

[6] A. D'Souza, S. Vijayakumar, and S. Schaal, "Learning inverse kinematics," in *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on*, vol. 1, 2001, pp. 298–303 vol.1.

[7] C. A. Buneo, M. R. Jarvis, A. P. Batista, and R. A. Andersen, "Direct visuomotor transformations for reaching," *Nature*, vol. 416, no. 6881, pp. 632–636, 2002.

[8] N. A. Melchior and R. Simmons, "Graph-based trajectory planning through programming by demonstration," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 1929–1936.

[9] F. Nori, L. Natale, G. Sandini, and G. Metta, "Autonomous learning of 3D reaching in a humanoid robot," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. IEEE, 2007, pp. 1142–1147.

[10] J. Piaget, *The Child's Conception of the World*. London: Paladin, 1973.

[11] J. G. Bremner, *Infancy*. Cambridge, MA: Blackwell, 1994.

[12] A. Baranes and P.-Y. Oudeyer, "The interaction of maturational constraints and intrinsic motivations in active motor development," in *Development and Learning (ICDL), 2011 IEEE International Conference on*, vol. 2, 2011, pp. 1–8.

[13] P. Savastano and S. Nolfi, "Incremental learning in a 14 dof simulated icub robot: Modeling infant reach/grasp development," *Biomimetic and Biohybrid Systems*, pp. 250–261, 2012.

[14] S. Ramırez-Contla, A. Cangelosi, and D. Marocco, "Developing motor skills for reaching by progressively unlocking degrees of freedom on the icub humanoid robot," in *Proceedings of the Post-Graduate Conference on Robotics and Development of Cognition*, 2012.

[15] F. Stulp and P.-Y. Oudeyer, "Emergent proximo-distal maturation through adaptive exploration," in *Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on*, 2012, pp. 1–6.

[16] J. I. P. De Vries, G. H. A. Visser, and H. F. R. Prechtl, "Fetal motility in the first half of pregnancy," in *Continuity of neural function from prenatal to postnatal life*, H. F. R. Prechtl, Ed. London: Spastics International Medical Publications, 1984, pp. 46–64.

[17] D. Caligiore, T. Ferrauto, D. Parisi, N. Accornero, M. Capozza, and G. Baldassarre, "Using motor babbling and hebb rules for modeling the development of reaching with obstacles and grasping," in *Proceedings of International Conference on Cognitive Systems COGSYS 2008*, R. Dillmann, C. Maloney, G. Sandini, T. Asfour, G. Cheng, G. Metta, and A. Ude, Eds. Karlsruhe, Germany: Springer, 2008, pp. E1–8.

[18] R. Saegusa, G. Metta, G. Sandini, and S. Sakka, "Active motor babbling for sensorimotor learning," in *Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference on*. IEEE, 2009, pp. 794–799.

[19] Y. Kuniyoshi and S. Sangawa, "Early motor development from partially ordered neural-body dynamics: experiments with a cortico-spinal-musculo-skeletal model," *Biol. Cybern.*, vol. 95, no. 6, pp. 589–605, 2006.

[20] B. Hopkins, *Prenatal development of postnatal functions*. Praeger Pub Text, 2005.

[21] H. Mori and Y. Kuniyoshi, "A human fetus development simulation: Self-organization of behaviors through tactile sensation," in *Development and Learning (ICDL), 2010 IEEE 9th International Conference on*. IEEE, 2010, pp. 82–87.

[22] Y. Yamada and Y. Kuniyoshi, "Embodiment guides motor and spinal circuit development in vertebrate embryo and fetus," in *Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on*. IEEE, 2012, pp. 1–6.

[23] D. Maurer and C. Maurer, *The World of the Newborn*. New York: Basic Books, 1988.

[24] C. von Hofsten and L. Rönnqvist, "The structuring of neonatal arm movements," *Child Development*, vol. 64, no. 4, pp. 1046–1057, 1993.

[25] J. Coelho, J. Piater, and R. Grupen, "Developing haptic and visual perceptual categories for reaching and grasping with a humanoid robot," *Robotics and Autonomous Systems*, vol. 37, no. 2, pp. 195–218, 2001.

[26] R. Grupen, "A developmental organization for robot behavior," Dept Of Computer Science, Massachusetts University, Amherst., DTIC Document ADA439115, 2005.

[27] S. Hart and R. Grupen, "Intrinsically motivated affordance discovery and modeling," in *Intrinsically Motivated Learning in Natural and Artificial Systems*. Springer, 2013, pp. 279–300.

[28] N. Bayley, "The development of motor abilities during the first three years: A study of sixty-one infants tested repeatedly," *Monographs of the Society for Research in Child Development*, vol. 1, no. 1, pp. 1–26, 1936.

[29] J. S. Bruner, *Processes of cognitive growth: infancy*. Worcester, Mass: Clark University Press, 1968.

[30] J. L. Elman, "Learning and development in neural networks: The importance of starting small," *Cognition*, vol. 48, no. 1, pp. 71–99, 1993.

[31] Y. Nagai, Y. Kawai, and M. Asada, "Emergence of mirror neuron system: Immature vision leads to self-other correspondence," in *Development and Learning (ICDL), 2011 IEEE International Conference on*, vol. 2, 2011, pp. 1–6.

[32] G. Butterworth and M. Harris, *Principles of Developmental Psychology*. Hove: Lawrence Erlbaum Associates, 1994.

[33] M. D. Sheridan, *From Birth to Five Years*. Windsor, USA: NFER Publishing, 1973.

[34] A. H. Fagg, A. G. Barto, and J. C. Houk, "Learning to reach via corrective movements," in *In Proceedings of the Tenth Yale Workshop on Adaptive and Learning Systems*, 1998, pp. 179–185.

[35] M. Schlesinger, D. Parisi, and J. Langer, "Learning to reach by constraining the movement search space," *Developmental Science*, vol. 3, pp. 67–80, 2000.

[36] T. S. Dahl and S. Barige, "Crude corrective movements for humanoid reaching learning," in *Towards Autonomous Robotic Systems (TAROS'09)*, Londonderry, UK, 2009, pp. 223–228.

[37] D. Vernon, C. von Hofsten, and L. Fadiga, *A roadmap for cognitive development in humanoid robots*, ser. Cognitive Systems Monographs (COSMOS). Springer, 2010.

[38] J. Law, M. Lee, M. Hülse, and A. Tomassetti, "The infant development timeline and its application to robot shaping," *Adaptive Behaviour*, vol. 19, pp. 335–358, 2011.

[39] J. Law, P. Shaw, and M. Lee, "A biologically constrained architecture for developmental learning of eye-head gaze control on a humanoid robot," *Autonomous Robots*, vol. 35, no. 1, pp. 77–92, July 2013.

[40] M. Hülse, S. McBride, J. Law, and M. Lee, "Integration of active vision and reaching from a developmental robotics perspective," *IEEE Transactions on Autonomous Mental Development*, vol. 4, no. 2, pp. 355–367, 2010.

[41] M. T. Sheldon and M. Lee, "PSchema: A developmental schema learning framework for embodied agents," in *Proceedings of the IEEE Joint International Conference on Development and Learning, and Epigenetic Robotics*. Frankfurt, Germany: IEEE, 2011.

[42] S. Vijayakumar, A. D'souza, and S. Schaal, "Incremental online learning in high dimensions," *Neural Comput.*, vol. 17, no. 12, pp. 2602–2634, Dec. 2005.

[43] M. R. Fiorentino, *A Basis for Sensorimotor Development, Normal and Abnormal: The Influence of Primitive, Postural Reflexes on the Development and Distribution of Tone*. Springfield, Ill: Charles C Thomas, 1981.

[44] C. Von Hofsten, "Prospective control: A basic aspect of action development," *Human Development*, vol. 36, pp. 253–270, Sep. 1993.

[45] N. E. Berthier and R. L. Carrico, "Visual information and object size in infant reaching," *Infant Behavior and Development*, vol. 33, no. 4, pp. 555 – 566, 2010.

[46] G. Gómez, M. Lungarella, P. Eggenberger Hotz, K. Matsushita, and R. Pfeifer, "Simulating development in a real robot: on the concurrent increase of sensory, motor, and neural complexity," in *Proc. 4th Int. Workshop on Epigenetic Robotics*. Lund University Cognitive Studies, 2004, p. 119122.

[47] C. von Hofsten and L. Rönnqvist, "Preparation for grasping an object: A developmental study," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 14, pp. 610–621, 1988.

[48] M. Latash, *Synergy*. Oxford University Press, USA, 2008.

[49] D. Rosenbaum, E. Halloran, and R. Cohen, "Grasping movement plans," *Psychonomic bulletin & review*, vol. 13, no. 5, pp. 918–922, 2006.

[50] P. Rochat, "Hand-mouth coordination in the newborn: Morphology, determinants, and early development of a basic act," in *The Development of Coordination in Infancy*, ser. Advances in Psychology, G. J. Savelsbergh, Ed. North-Holland, 1993, vol. 97, pp. 265–288.

[51] K. Ennouri and H. Bloch, "Visual control of hand approach movements in new-borns," *British Journal of Developmental Psychology*, vol. 14, no. 3, pp. 327–338, 1996.

[52] F. Goodkin, "The development of mature patterns of head –eye coordination in the human infant," *Early Human Development*, vol. 4, pp. 373–386, 1980.

[53] M. M. Shirley, *The First Two Years - A Study of Twenty-five Babies*, ser. Intellectual Development. Minneapolis: University of Minnesota Press, 1933, vol. 2.

[54] N. E. Berthier, R. K. Clifton, D. D. McCall, and D. J. Robin, "Proximodistal structure of early reaching in human infants," *Experimental Brain Research*, vol. 127, pp. 259–269, 1999.

[55] N. E. Berthier and R. Keen, "Development of reaching in infancy," *Experimental Brain Research*, vol. 169, pp. 507–518, 2006.

[56] N. E. Berthier and R. L. Carrico, "Visual information and object size in infant reaching," *Infant Behavior and Development*, vol. 33, no. 4, pp. 555–566, 2010.

[57] R. K. Clifton, D. W. Muir, D. H. Ashmead, and M. G. Clarkson, "Is visually guided reaching in early infancy a myth?" *Child Development*, vol. 64, no. 4, pp. 1099–1110, 1993.

[58] R. K. Clifton, P. Rochat, D. J. Robin, and N. E. Bertheir, "Multimodal perception in the control of infant reaching," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 20, pp. 876–886, 1994.

[59] B. L. White, P. Castle, and R. Held, "Observations on the development of visually-directed reaching," *Child Development*, vol. 35, no. 2, pp. 349–364, 1964.

[60] R. Carvalho, E. Tudella, and G. Savelsbergh, "Spatio-temporal parameters in infant's reaching movements are influenced by body orientation," *Infant Behavior and Development*, vol. 30, no. 1, pp. 26 – 35, 2007.

[61] M. Lee, Q. Meng, and F. Chao, "Staged competence learning in developmental robotics," *Adaptive Behaviour*, vol. 15, no. 3, pp. 241–255, 2007.

[62] K. Earland, J. Law, P. Shaw, and M. Lee, "Overlapping structures in sensory-motor mappings," *Plos One*, 2013.

[63] V. Braitenberg and A. Schüz, *Anatomy of the cortex: Statistics and geometry*. Springer-Verlag Berlin, 1991.

[64] H. Mallot, W. Von Seelen, and F. Giannakopoulos, "Neural mapping and space-variant image processing," *Neural Networks*, vol. 3, no. 3, pp. 245–263, 1990.

[65] Q. Meng and M. H. Lee, "Novelty and habituation: the driving forces in early stage learning for developmental robotics," in *Biomimetic Neural Learning for Intelligent Robots*, s. Wermter, P. G., and E. M., Eds. Springer, 2005, pp. 315–332.

[66] M. H. Lee, Q. Meng, and F. Chao, "Developmental learning for autonomous robots," *Robotics and Autonomous Systems*, vol. 55, no. 9, pp. 750–759, 2007.

[67] F. Chao, M. Lee, and J. Lee, "A developmental algorithm for ocular-motor coordination," *Robotics and Autonomous Systems*, vol. 58, pp. 239–248, 2010.

[68] X. Wang and J. Jin, "A quantitive analysis for decomposing visual signal of the gaze displacement," in *Proceedings of the Pan-Sydney area workshop on Visual information processing - Volume 11 (VIP '01)*, D. D. Feng, J. Jin, P. Eades, and H. Yan, Eds., vol. 11. Darlinghurst, Australia: Australian Computer Society, Inc., 2001, pp. 153–159.

[69] J. Law, P. Shaw, K. Earland, M. Sheldon, and M. Lee, "A psychology based approach for longitudinal development in cognitive robotics," *Frountiers in neurorobotics*, submitted.

[70] B. White, P. Castle, and R. Held, "Observations on the development of visually-directed reaching," *Child Development*, vol. 35, no. 2, pp. 349–364, 1964.

[71] M. Costantini, E. Ambrosini, G. Tieri, C. Sinigaglia, and G. Committeri, "Where does an object trigger an action? An investigation about affordances in space," *Experimental Brain Research*, vol. 207, no. 1-2, pp. 95–103, 2010.

[72] V. V. Hafner and G. Schillaci, "From field of view to field of reach-could pointing emerge from the development of grasping," in *Proceedings of the IEEE Conference on Development and Learning and Epigenetic Robotics (IEEE ICDL-EPIROB 2011)*, 2011.

[73] N. E. Berthier, R. K. Clifton, D. D. McCall, and D. J. Robin, "Proximodistal structure of early reaching in human infants," *Experimental Brain Research*, vol. 127, pp. 259–269, 1999.

[74] P. Rochat, "Hand-mouth coordination in the newborn: Morphology, determinants, and early development of a basic act," in *The Development of Coordination in Infancy*, ser. Advances in Psychology, G. J. Savelsbergh, Ed. North-Holland, 1993, vol. 97, pp. 265 – 288.

[75] M. T. Sheldon, "Intrinsically motivated developmental learning of communication in robotic agents," Ph.D. dissertation, Aberystwyth University, September 2012.

[76] J. Piaget and M. Cook, *The origins of intelligence in children.* WW Norton & Co, 1952.

[77] P. Shaw, J. Law, and M. Lee, "An evaluation of environmental constraints for biologically constrained development of gaze control on an iCub robot," *Paladyn*, vol. 3, no. 3, pp. 147–155, 2012.

[78] J. J. Campos, D. I. Anderson, M. A. Barbu-Roth, E. M. Hubbard, M. J. Hertenstein, and D. Witherington, "Travel broadens the mind," *Infancy*, vol. 1, no. 2, pp. 149–219, 2000.

[79] P.-Y. Oudeyer, A. Baranes, and F. Kaplan, "Intrinsically motivated learning of real-world sensorimotor skills with developmental constraints," in *Intrinsically Motivated Learning in Natural and Artificial Systems*, G. Baldassarre and M. Mirolli, Eds. Springer, 2013, pp. 303–365.

[80] M. Rolf and J. J. Steil, "Goal babbling: a new concept for early sensorimotor exploration," in *Humanoids 2012 Workshop on Developmental Robotics: Can developmental robotics yield human-like cognitive abilities?*, 2012.

[81] S. Nguyen and P.-Y. Oudeyer, "Active choice of teachers, learning strategies and goals for a socially guided intrinsic motivation learner," *Paladyn*, vol. 3, no. 3, pp. 136–146, 2012.

[82] A. Baranes and P.-Y. Oudeyer, "Active learning of inverse models with intrinsically motivated goal exploration in robots," *Robotics and Autonomous Systems*, vol. 61, no. 1, pp. 49–73, 2013.

[83] C. Moulin-Frier, P.-Y. Oudeyer *et al.*, "Exploration strategies in developmental robotics: a unified probabilistic framework," in *ICDL-Epirob-International Conference on Development and Learning, Epirob*, 2013.