Sensorimotor contingencies as a key drive of development: from babies to robots

Lisa Jacquey* (1, 2), Gianluca Baldassarre (3), Vieri Giuliano Santucci (3), & J. Kevin O'Regan (1)

(1) Integrative Neuroscience and Cognition Center - UMR 8002; CNRS & Université Paris Descartes; 45 rue des Sts. Pères; 75006 Paris; France

(2) Laboratoire Ethologie Cognition Développement - Université Paris Nanterre; 200 avenue de l'Université; 92000 Nanterre; France

(3) Laboratory of Computational Embodied Neuroscience; Istituto di Scienze e Tecnologie della Cognizione; Consiglio Nazionale delle Ricerche (LOCEN-ISTC-CNR); Via San Martino della Battaglia 44, I-00185 Roma; Italy

* Corresponding author: Email: lisa.jacquey@gmail.com

Abstract

Much current work in robotics focuses on the development of robots capable of autonomous unsupervised learning. An essential prerequisite for such learning to be possible is that the agent should be sensitive to the link between its actions and the consequences of its actions, called sensorimotor contingencies. This sensitivity, and more particularly its role as a key drive of development, has been widely studied by developmental psychologists. However, the results of these studies may not necessarily be accessible or intelligible to roboticians. In this paper, we review the main experimental data demonstrating the role of sensitivity to sensorimotor contingencies in infants' acquisition of four fundamental motor and cognitive abilities: body knowledge, memory, generalization and goal-directedness. We relate this data from development psychology to work in robotics, highlighting the links between these two domains of research. In the last part of the article we present a blueprint architecture demonstrating how exploitation of sensitivity to sensorimotor contingencies, combined with the notion of "goal", allows an agent to develop new sensorimotor skills. This architecture can be used to guide the design of specific computational models, and also to possibly envisage new empirical experiments.

Key words: Infant Development, sensorimotor development, sensorimotor contingencies, goals, Unsupervised learning

1. Introduction

Over recent years there has been growing interest in constructing robotic systems whose physical and cognitive competence evolves progressively in a way analogous to how biological systems, and in particular humans, develop in the womb and after birth. A variety of approaches to unsupervised learning are currently being investigated (cf. for ex. Cangelosi and Schlesinger, 2015), but all of them share an underlying assumption, which is that the agent's own action, often autonomously decided, must be used to progressively enrich the agent's knowledge by probing the environment. In biological organisms the most basic type of such probing must occur already at the very beginning of development. Thus, for example in human fetuses and infants, the organism must be sampling the sensory effects of its actions, making use of these "sensorimotor contingencies" to construct its understanding of its own body, its sensory systems, and of its proximal and distal environment, including the social environment. Ultimately, it is assumed, sensitivity to sensorimotor contingencies will contribute to the establishment of socio-cognitive and motor skills (Gergely and Watson, 1996; Tarabulsy et al., 1996; Nagai and Asada, 2015; Verschoor and Hommel, 2017) like language (e.g., Goldstein and Schwade, 2008; Schwartz et al., 2012), reaching and grasping (e.g., Corbetta, and Snapp-Childs, 2009) and walking (e.g., Adolph et al., 2000).

In robotics it is simply taken for granted that agents must be sensitive to sensorimotor contingencies, and research focuses on the details of how new cognitive, motor, and social skills are acquired. On the other hand, in developmental psychology, sensitivity to sensorimotor contingencies has itself been studied and documented. First, researchers have demonstrated that sensitivity to sensorimotor contingencies is already present in fetuses and newborns for stereotyped actions such as hand-mouth coordination (e.g., Myowa-Yamakoshi and Takeshita, 2006) or non-nutritive sucking (e.g., DeCasper and Spence, 1986), and then that it extends to other actions in the first months of life (e.g., Rovee and Rovee, 1969). Researchers have also

tried to describe how sensitivity to sensorimotor contingencies contributes to the development of infants' motor and cognitive functions. In this paper, we will focus on four examples (body knowledge, memory, generalization, and goal-directedness) which are central to development, and that we will address through the four following questions: Are infants able to identify the particular motor skills or body parts involved in a sensorimotor contingency? How long is a sensorimotor contingency retained in memory? To what extent are infants capable of generalizing a sensorimotor contingency to a new situation? Are infants able to use sensorimotor contingencies in order to attain desired states?

In what follows we will look at how these four questions have been experimentally addressed in developmental psychology — for reference, experimental methods used in the studies reviewed in this paper are summarized in Table 1. In each section, we will link infant data to the main computational and developmental robotics literature also reviewed in a prospective and critical way. Note that the "Approaches in robotics" sections do not claim to be exhaustive, but suggest links between infant study findings and current work in developmental robotics. We close the paper by presenting a blueprint architecture directed to integrate the key elements of contingency-based learning from a robotics perspective. This can be used to guide the design of specific computational models and also to possibly envisage new empirical experiments (the architecture was introduced in Baldassarre et al., 2018a).

We are aware that remarkable work has already linked data from developmental robotics and developmental psychology (e.g., Guerin et al., 2013; Caligiore et al., 2014; Cangelosi and Schlesinger, 2015; Jamone et al., 2016; Oudeyer et al., 2016). The novelty of our approach is to focus on the mechanism of sensitivity to sensorimotor contingency itself (since as such it has not been a main focus of developmental robotics) and to illustrate with four examples how this sensitivity contributes to development. In our survey, we shall concentrate on studies involving babies under one year of age, where presumably the mechanisms involved in

contingency detection will be easier to study, since they will be less likely to be masked by higher learning processes.

Procedures exploiting infants' sensitivity to sensorimotor contingencies mentioned in this paper

Operant conditioning procedure — Experiments in which an infant is exposed to a sensorimotor contingency, i.e., a particular action of the infant is reinforced by a "pleasant" sensory feedback. Experimenters measure the frequency of occurrence of the reinforced action before, during and after exposure to the sensorimotor contingency. There is a wide variety of operant conditioning procedures, those used in the studies cited in this article are:

Mobile paradigm — One limb (leg or arm) of the infant is attached with a ribbon to a mobile placed above the infant. Movements of the "connected" limb of the infant produce contingent movement of the mobile (audiovisual feedback). The infant's limb movements are measured. Such a procedure is called a "conjugate reinforcement" as movements of the mobile are proportional to the infant's movements. This paradigm was developed and intensively used with 3-4-month-old infants by Rovee-Collier and her collaborators.

Non-nutritive sucking — A non-nutritive pacifier is placed in the infant's mouth. Sucking of the pacifier produces a contingent feedback (visual or auditory). Different features of the sucking behaviour can be reinforced (e.g., high amplitude sucking). The infant's number of sucks and/or frequency of sucking are measured. This protocol has been mainly used with newborns and very young infants, since sucking is one of the first mastered motor skills.

Gaze contingency — One position of the infant's gaze on a screen is associated with the appearance of a picture on the screen (at the same position or elsewhere). The location of the infant's gaze on the screen is measured and the number of "reinforced gazes" is computed. The technique is effective because gaze movements are well controlled in infants from 4 months of age.

Button contingency (e.g., the train paradigm) — The infant pushes on a button to trigger an audiovisual feedback (e.g., motion of an electric train). The number of pushes on the button is

measured. This protocol is used in infants older than 6 months since younger infants cannot easily use buttons.

Means-end procedure — Experiments in which an action on an object triggers an effect on another part of the object or on another object (e.g., pulling a string to retrieve an out-of-reach toy). The purpose here is to measure the infant's ability to differentiate between the means used (e.g., grasping the string) and the goal achieved (e.g., attaining the toy).

Revaluation procedure — Experiments in which the valence of a contingent feedback is modulated. In a first phase, the infant learns one (or two) action-effect couple(s). Then, the value of (one of) the contingent effect(s) is changed. After this relative change in valuation of the stimulus, it is assumed that the infant will have a strong preference for (one of) the contingent effect(s). At the end, the infant has the opportunity to execute the previously learnt contingent action(s) to obtain the associated effect(s). It is important to highlight that the contingent feedback is not displayed at this phase of the experiment in order to ensure that the infant's choice of action is goal-directed (i.e., not stimulus-driven). The number of contingent actions realized by the infant is measured before and after the revaluation.

Buzzer procedure — Experiments in which a vibrotactile stimulation is applied on the infant's body for several seconds at different locations. Motor response of the infant is measured (e.g., limb activity, reaction times, reach/grasp of the buzzer).

Table 1: Overview of the procedures used in the infant experiments mentioned in this paper.

2. Body knowledge: Is the agent able to identify the particular motor skills or body parts involved in a sensorimotor contingency?

In a seminal paper, Oudeyer et al., (2005) imagined how a robotic agent (an AIBO robot), playing in the equivalent of a "baby gym", could discover both the functioning of its own body,

and also the actions that its immediate environment affords. This scenario is interesting because it can be applied both to robots and to human babies with respect to the question of the acquisition of body knowledge. Thus, one can imagine that the agent starts by doing random "motor babbling". At every moment it activates a variety of effectors and receives feedback through its sensors from the environment. For example, the agent might by chance touch a bell suspended near it, and perceive a resulting auditory input. Its "intrinsic motivation" might then prompt it to attempt to replicate the sound by repeating the action it performed. By progressively refining its actions, will the agent be able to precisely identify the motor command that produced the bell sound? Will this ultimately allow the agent to acquire that command as a new implicit skill, thereby contributing to better distinguish its body parts?

2.1. Evidence from developmental psychology

2.1.1. Motor skill refinement

Researchers have investigated to what extent exploration of sensorimotor contingencies could lead infants to refine their motor skills. To do so, they have tested infants in modified versions of the "mobile paradigm" (see description in Table 1) in which movements which were not in the infant's usual repertoire of actions generated the contingent feedback. For example, Angulo-Kinzler et al. (2002) tested 3-month-old infants and used knee extensions and knee flexions as triggering action. The authors showed that a few minutes after the beginning of exposure to the contingency, infants started to explore these non-usual leg movements. Interestingly, after a longer exposure, infants refined the necessary motor action and were able to make the actions in a more efficient manner. Similar observations were made in several other studies (e.g., Thelen, 1994; Angulo-Kinzler, 2001; Chen et al., 2002; Tiernan and Angulo-Barroso, 2008; Sargent et al., 2014; Sargent et al., 2015), suggesting that infants' exploration of sensorimotor contingencies might play a role in motor skill refinement. Developmental

psychologists have also measured the influence of contingent feedback on the development of emerging skills such as reaching. The best known example is the "sticky mittens" experiment (e.g., Needham et al., 2002) in which 3-month-old babies who are not yet capable of reaching/grasping received two weeks of daily training sessions during which they carried mittens with Velcro attached to their fingertips, allowing them to pick up objects. After this training, babies that had had this new sensorimotor experience showed more (visual and oral) object exploration compared to non-trained babies. In another experiment, Lee and Newell (2013) exposed 10 to 16-week-old infants biweekly to a contingency between movements of one arm and auditory feedback. They found that this longitudinal intervention did not influence the moment in development when reaching emerged. However, they observed that compared to a control group, infants exposed to the contingency made more exploratory and "goal-directed" movements. Compatible with this, Needham et al. (2014) showed that 3-month-old infants were able to increase their reaching-like movements when these movements were contingently reinforced (here through the mobile paradigm, c.f. Table 1). This increase in reaching-like movements was observed during contingency exposure but also directly afterwards, when infants participated in a free-play experiment with a rattle. These results have been confirmed in a recent study by William and Corbetta (2016). The authors trained 3-month-old infants to reach for an object every day for two weeks. They found that infants for whom reaching movements towards the object generated contingent movements of the object, contacted the object significantly more than the control groups for whom reaching movements did not produce the contingent feedback.

2.1.2. Body part differentiation

In the previous paragraph, we showed evidence for an accumulation of implicit motor knowledge acquired through exploration of sensorimotor contingencies. This accumulation of

implicit knowledge might lead infants to progressively differentiate between their body parts. Thus, one might expect that in the youngest infants, sensorimotor contingencies would not be specific to individual body parts, but rather that they would involve global body motions. However, as infants grow older one might expect that sensorimotor contingencies would narrow down to individual body parts. It is difficult to directly prove that there is such a causal link between accumulation of implicit knowledge provided by the detection of sensorimotor contingencies and body part differentiation. Below, we present what evidence there is more generally concerning the progressive differentiation of body responses over development.

Differentiation of body responses has been investigated in infants between 2 and 4 months of age by measuring their ability to narrow down a contingency to a specific limb, for instance to the limb connected to the mobile in the mobile paradigm (c.f. Table 1). Around 3 months of age, infants exposed to the mobile paradigm with one of their legs connected increase the movements of both of their legs and not the movements of their arms or even only the movements of their connected leg (e.g., Rovee-Collier et al., 1978; Heathcock et al., 2005). Going one step further, the team of Watanabe and Taga (Watanabe and Taga, 2006, 2009; Watanabe et al., 2011; Kato et al., 2013) investigated how differentiation of body responses evolves between 2 and 4 months of age, again using the mobile paradigm but connecting one arm of the infant to the mobile. Watanabe and Taga (2006) observed that 2-month-old infants increased activity of their four limbs equally, 3-month-old infants increased activity of both of their arms equally but did not increase the activity of their legs, and 4-month-old-infants increased the activity of their connected arm only. In another paper, Watanabe and Taga's team (Watanabe and Taga, 2009) showed that the moment when differentiation in body responses emerges differs depending on the body part involved. They found that 3-month-old infants with one arm connected increased activity of their arms only, but that infants of the same age with one leg connected increased the activity of their whole body. Interestingly, they proposed that

this difference in response specificity might be explained by developmental differences in inhibitory abilities between arm and leg motor systems (Kato et al., 2013). D'Souza et al. (2016) have added a further interesting result by testing older infants' motor specialization during unimanual actions. More precisely, they measured the number of extraneous movements made by 9- and 12-month-old infants during play with a rattle. Here again, they found evidence for a progressive refinement of body knowledge, since extraneous "overflow" movements decreased with age.

Differentiation of body responses has also been investigated by measuring infant's responses to vibrotactile stimulation (cf. Table 1) (Bremner et al., 2008; Begum Ali et al., 2015; Hoffman et al., 2017; Somogyi et al., 2018; Chinn et al., 2019). In these experiments, authors measured the ability of infants to localize a vibrating buzzer on a specific body part. More precisely, authors measured both infants' implicit localizations — which limb was moved first or which limb was more active during the stimulation — and explicit localizations — looks/touches/grasps towards the buzzer. The authors found a developmental pattern: at 4 months of age infants seemed able to implicitly localize the tactile stimulation; at 6 months they became able to touch and/or grasp it and only at 10 months did they start to look at it. More interestingly for our purpose, Somogyi et al. (2018) found that the youngest infants' reactions (3 months of age) involved global body motions — i.e. the youngest infants moved their whole bodies during stimulation — whereas older infants' reactions (5-6 months of age) involved limb-specific body motions — i.e. older infants moved the stimulated limb more.

2.1.3. Conclusion from developmental psychology

The studies reviewed above suggest that exploration of sensorimotor contingencies allows infants to refine motor skills and to differentiate their body parts. Thus, in the first two months of life, detecting contingent actions involve general motor activity of the infant. Then, the third

month of postnatal life appears to be a pivot in development: detected contingent actions become more integrated into the refinement of existing skills. This change at 3 months of age is also reflected in body differentiation: arms and legs seem more differentiated. One month later, body knowledge specialization seems to be even more effective, and limb differentiation seems to be achieved. Thus, body knowledge development, achieved through sensorimotor contingency detection, presumably brings infants to acquire a repertoire of well-defined actions, leading them to be ready at 3-4 months of age to develop reaching, grasping and later, tool-use abilities. Another point is worth noting: gradual maturation of inhibitory processes might also be involved in the progressive specialization of motor responses. The lack of inhibition in motor babbling may be a factor that prevents younger infants from easily making limb-specific actions[1]. It may be the increase in inhibitory abilities that leads older infants to discover limb-specific actions.

2.2. Approaches in robotics

The work from Oudeyer et al. (2005) confirms that a simple robot equipped with what has been called "intelligent adaptive curiosity" can indeed acquire information about its body, at least information of an implicit type -- that is, the agent gradually learns to use its body more effectively to explore its environment. Following the idea that understanding one's effects on the environment is crucial for the autonomous development of animals and humans (White 1959; Berlyne, 1960) a variety of work in robotics has focused on the autonomous learning of skills on the basis of the interactions between the body of the artificial agent and the environment, where robots are tested in "simple" reaching or avoidance scenarios (e.g.,

---

[1] Nonetheless, even newborns are able to display distinctive actions when they are in specific conditions (Grenier, 1981; van der Meer et al., 1995, 1996; van der Meer, 1997).

Santucci et al., 2014a; Hafez et al., 2017; Reinhart, 2017; Hester and Stone, 2017; Tannenberg et al., 2019) or in more complex tasks involving interactions between objects (Da Silva, 2014; Seepanomwan et al., 2017), tool use (Forestier and Oudeyer, 2016) or hierarchical skill learning (Forestier et al., 2017; Colas et al., 2018; Santucci et al., 2019), and even in imitation learning experiments (Duminy et al., 2018). When combined with the use of "goals", intended here as specific states or effects that a system is willing to obtain, curiosity and intrinsic motivation are able to properly guide task selection (Merrick, 2012; Santucci et al., 2016) and reduce the exploration space (Rolf et al., 2010; Baranes and Oudeyer, 2013).

On the other hand, the relation between sensorimotor contingencies and the development of early body knowledge per se is still poorly studied in robotics. Hoffmann and colleagues in different studies has underlined how the development of the body schema may improve the functioning of artificial agents (Hoffmann et al., 2010; Roncone et al. 2014), as well as how robots can be used to model early sensorimotor development (Hoffmann et al., 2017; Hoffmann and Pfeifer 2018). The way peripersonal motor maps can help robot learning has been studied in Jamone et al. (2014). Mannella et al. (2018) presented a specific hypothesis on the learning of sensorimotor contingencies in relation to the gradual acquisition of the knowledge of the agent's own body driven by intrinsic motivations. In particular, a simulated robotic model of a torso with two arms is free to move around with its effectors. Whenever a new contingency is discovered (here the self-touch generated by the hands colliding with the body) the robot systematically attempts to replicate the contingency, thus acquiring new motor skills and at the same time forming an internal map of its own body on the basis of the touch events that it perceives during the experiment. The blueprint architecture introduced below represents the theoretical framework within which this model was developed. Similarly, Juett and Kuipers (2019), using a Baxter robotic platform, study how visual sensorimotor contingencies are a key

element in developing the robot's knowledge of its own body and the way it relates to the environment.

3. Memory: How long is a sensorimotor contingency retained in memory?

Returning to Oudeyer et al.'s (2005) example of the agent in its baby gym, let us imagine that yesterday it discovered a new contingency: for example, touching a bell to generate a sound. When returning to the baby gym today, will the agent remember what it learned yesterday, or must it relearn the contingency from scratch? And how does this depend on the delay since first encountering the contingency? How does practice in the interim affect recall?

3.1. Evidence from developmental psychology

Rovee-Collier and her team are the authors who have contributed extensively to the question of memory of sensorimotor contingencies. Their studies involved infants from 2 to 18 months of age in an experimental framework consisting in training infants for two or three days with a sensorimotor contingency that involved 6 to 9 minutes of exposure on each training session, and then in testing infants' retrieval of this contingency after different delays: from one day to a few months after training ended. In addition, in some of their studies they reactivated the learning between training and testing through a brief exposure to the contingent stimulus or to the context in which learning occurred. The paradigm they used was the mobile paradigm with infants aged less than 6 months, and the "train" paradigm with older infants (c.f. Table 1).

3.1.1. Retention

One of Rovee-Collier and her colleagues' lines of research was the retention of sensorimotor contingencies (e.g., Sullivan et al., 1979). Rovee-Collier and her colleagues showed that one day after training with the mobile, infants of all ages were able to retrieve the contingency.

However, for longer delays, the retrieval was dependent on the age of the infant: the duration of retention increased progressively from one day to 3 months, as the age of the infant increased from 2 to 18 months of age (see Rovee-Collier, 1999 for a review). It is worth noting that infants' retention can be modulated by the training schedule. For example, for the same amount of training — 18 minutes — 2-month-old infants' retrieval of a contingency increased from one day to two weeks when the training was spread over three days instead of two days (Rovee-Collier, 1999).

3.1.2. Reactivation

In another series of experiments, Rovee-Collier and her colleagues introduced a "reactivation treatment" (priming) one week (or two) after the infant had forgotten the contingency (e.g., Fagen and Rovee-Collier, 1983). This reactivation consisted in three minutes of exposure to the original mobile/train or to the original environment used previously (e.g., testing room, color of the infant's crib). They found that both these modes of reactivation enabled infants aged 3 months and older to again retrieve the forgotten contingency. The authors found that the temporal course of forgetting after reactivation was similar to the temporal course after training: forgetting increases with delay. Moreover, the reactivation treatment becomes effective more quickly as age increases: while young infants need around 24 hours for the reactivation treatment to become effective, it becomes effective immediately for infants that are 12-months-old or older (Rovee-Collier, 1999). The delay needed to obtain this reactivation increases logarithmically over the first year of life. Also, the time of exposure needed to obtain reactivation decreases, going from 2 minutes at 3 months of age to 1.8 seconds at 12 months (Rovee-Collier, 1999).

3.1.3. Conclusion from developmental psychology

Take home messages from these studies are that (i) from very early in life (2 months of age), babies are able to remember a contingency one day after exposure, (ii) the duration of retention increases gradually to as much as 3 months after exposure in the first year of life, and (iii) long-term retention of a contingency is enhanced by as little as a few minutes of reactivation of the contingency. Interestingly such reactivation does not require practicing the contingency, but only perception of the stimulus involved in the contingency (e.g., seeing the mobile moving in a non-contingent manner) (e.g., Rovee-Collier et al., 1980). Thus, one might suggest that young infants create something like an "inverse model" of contingencies they are exposed to, since they are able to activate an entire contingency from its stimulus only. Moreover, it is worth noting that retention abilities of infants are enhanced when learning is spread over several days, suggesting that maybe, even in very young babies, sleep might help in consolidating sensorimotor learning (see Gómez and Edgin, 2015 for a review).

3.2. Approaches in robotics

Studies in artificial neural networks and robotics may shed some light on the memory processes illustrated above. First, modelling through artificial neural networks allows the specification of the mechanisms behind the memorisation processes. Neural-network studies on learning processes has shown that the ways in which neural networks can "memorise" experience are multiple and can be grouped into three broad categories depending on the source of the output to be associated to the (perceptual) input: supervised learning, unsupervised learning and reinforcement learning. In supervised learning the output is provided to the agent by an external teacher or by another internal component (Bengio et al., 2017). In unsupervised learning the output is the result of algorithms that identify different possible statistical regularities in the sensory input (e.g., as in *self-organising maps*, Kohonen, 2001). Finally, in reinforcement

learning the output is actively searched for by the agent using active exploration of the environment while attempting to maximise a reward measure having either an external or an internal origin. Among the three classes, reinforcement learning is clearly closer to contingency-based learning since it captures the fact that learning relies on free exploration of the environment and since contingent events would seem to be eligible to produce the reward signals needed to guide the learning process. These features are indeed at the core of the blueprint architecture presented below.

Once a memory is formed, it may undergo forgetting processes. One might wonder if the memory decay observed in infants is detrimental or advantageous for cognitive processing. Computational studies offer hypotheses for both cases. One explanation of memory loss might be that when new experiences are acquired, these might impair the existing memories from previous experiences ("catastrophic interference", French, 1999). The fact that robustness of infants' memories differs at different ages might reflect the fact that initially their memories are stored in a broader overlapping unspecific way that makes them more vulnerable to interference from new experiences (we are not aware of any model studying this). On the other hand, memory loss might be beneficial in facing the very important problem of "overfitting", i.e. the tendency to memorize irrelevant aspects of experience that have no informative value. Interestingly, some of the most important solutions to this problem within the computational literature, called "regularization" mechanisms, rely on some form of forgetting (Bengio et al., 2017). Intuitively, regularisation mechanisms allow the system to release memories that correspond to rare aspects of experiences (irrelevant noise is usually of this type) and to retain the important "structure" of information that tends to repeat across different experiences.

4. Generalization: To what extent is the agent capable of generalizing a sensorimotor contingency to a new situation?

Imagine now that someone had changed the carpet of the baby gym from red to green. How is the agent's behaviour affected by this change in the background context? Is the agent able to generalize its prior learning? And what if the objects themselves in the gym had been changed, instead of the background color?

4.1. Evidence from developmental psychology

Generalization within the framework of sensorimotor contingencies would be the ability to transfer a contingency learned in one, original, set of conditions to a new situation. Experiments designed to test this have considered cases when the new situation differs from the original situation by a change in the stimulus and/or in the context.

4.1.1. Changes in stimulus

The effect of changing the stimulus was extensively investigated during the late 60's and 70's using the non-nutritive sucking paradigm (See Table 1). At that time, researchers were interested in infants' discrimination abilities. They noticed that in non-nutritive sucking experiments, infants progressively lost interest in the visual or auditory stimulus generated by their sucking (this is known as habituation, see for ex. Aslin, 2007 for a recent review on visual habituation). Interestingly, if the stimulus was then changed to a new stimulus, infants' interest increased, as reflected by an increase in their sucking rate. Results from these experiments suggest that infants are able to generalize their learning of contingency to a new stimulus when the change occurs immediately after learning (e.g., Siqueland and DeLucia, 1969).

The effect of changing the stimulus was also explored by Rovee-Collier and her collaborators in 2- to 12-month-old infants using the mobile paradigm (see Table 1) (see

16

Hartshorn et al., 1998 or Rovee-Collier, 1999 for reviews). They observed that young infants — 2 to 6 months of age — could not transfer their learning of a contingency to a novel stimulus (i.e. a different mobile) when they were tested one day after training with this new stimulus. In contrast, older infants — 9 and 12 months of age — could easily transfer the learning of a contingency between a lever and the movements of a toy train to a novel train when they were tested on the next day and up to 2 weeks later but not for longer delays (3 to 8 weeks). Rovee-Collier's team (Butler and Rovee-Collier, 1989) also tested generalization abilities of infants when both stimulus and context were changed simultaneously. They found that in this case infants were not able to generalize their learning even when they were tested only one day after training. This result is in line with what had been found previously — i.e. changes in stimulus strongly impair generalization abilities. Moreover, Rovee-Collier and her collaborators (Fagen and Rovee-Collier, 1976; Mast et al., 1980) explored the ability of 3-month-old infants to generalize their learning when quantitative aspects of the contingency were changed. They trained infants with one mobile and then switched to another mobile that provided weaker or stronger visual feedback (they did this by changing the number of components of the mobile). They found that even though infants noticed the change and reacted negatively, they were still able to generalize their learning and even to adapt their kicking rate to maintain the original amount of visual feedback.

These studies show that younger infants have greater difficulty transferring their learning to a novel situation compared to older infants. But the question then arises of exactly what degree of difference makes for a "novel" situation? It seems that for young infants, very little change produces a "novel" situation. For instance, in the mobile paradigm, young infants were not able to generalize their learning as soon as a few components (one to three out of five) or the symbols painted on the mobile were changed (e.g., Fagen et al., 1976; Hayne et al., 1986; Adler and Rovee-Collier, 1994). Nevertheless, Rovee-Collier and her collaborators found four

specific situations in which infants were able to generalize their learning to a novel stimulus (see Hartshorn et al., 1998 for a review). The first situation is when the novel and the original mobile share some visual features to which the infant selectively attends. The second situation is when the infant has forgotten the specific details of the original mobile. The third situation is when the infant is trained with a series of different mobiles. The fourth situation is when the infant is exposed to a novel object in motion after training is completed.

4.1.2. Changes in context

Regarding the effect of changing the environment (e.g., testing room, color of the infant's crib), Rovee-Collier and her collaborators (see Hartshorn et al., 1998 or Rovee-Collier, 1999 for reviews) studied this on infants aged 3 to 18 months of age. They showed that when tested one day after training, both young infants — 3 to 6 months of age — and older infants — 9 to 12 months of age — were able to transfer their learning to a new context. This ability remains stable for delays longer than one day, and up to one month for older infants (9-12 months) but not for young infants (3-6 months). Thus, it seems that young infants are strongly sensitive to the context in which they learned a contingency. Moreover, what is considered as context in the infant's learning is not only nearby and very salient cues but also the more global environment. For example, Hayne et al. (1991) trained 3-month-old infants in a portable crib in their bedroom, and then tested the infants in the same crib with the same mobile but in a different room. They found that infants did not show retrieval of the contingency when they were tested in a different room. Hartshorn and Rovee-Collier (1997) found the same results with 6-month-old infants using the train task. Nonetheless, as for a change in mobile, the failure in generalizing learning to a new context can be overcome by training in multiple contexts (e.g., Rovee-Collier and Dufault, 1991).

4.1.3. Conclusion from developmental psychology

It appears that over the course of development infants' ability to generalize from one contingency to another increases gradually. While at first learning of sensorimotor contingencies is strongly impacted by the specific features of the contingent stimulus and by the context, learning becomes progressively more tolerant to stimulus and context changes during infants' first year of life. This tolerance does not correspond to a decrease in infants' ability to discriminate between the original and the new stimulus/context. Rather, it may reflect an increase in infants' ability to recognize two situations as functionally equivalent and then to apply what they have previously learned in one situation to another one. Interestingly, this ability seems to emerge at the time when infants become self-locomoting (around 9 - 12 months of age with crawling and then walking), suggesting that infants become able to separate what has been learnt from the associated details/context only when they become able to spatially explore the environment by themselves (Rovee-Collier, 1996; Herbert et al., 2007).

4.2. Approaches in robotics

Generalisation, the ability to respond to patterns and situations different from those previously experienced, is an appealing feature of neural-network models (McClelland et al., 1986), in particular when compared to "symbolic systems" (Russell and Norvig, 2010). Generalisation ranges from the ability to keep responding when the main stimulus is slightly novel to the ability to keep responding when both the main stimulus and the context are very different. Regarding generalisation over stimuli, research with computational models explains some of the observations of Rovee-Collier and colleagues on infants. In particular, the fact that the exposure to a large number of stimuli facilitates generalisation is in line with one of the most important principles of neural networks, according to which increasing the size of the dataset used to train networks greatly benefits their generalisation capabilities (Bengio et al., 2017).

The reason for this is that important information tends to be present in all stimuli and so to be encoded, whereas irrelevant information due to noise or secondary features is much more variable and so tends to change connection weights in opposing directions and thus to not be encoded. With few experiences of stimuli, this mechanism is impaired as both important and irrelevant information is encoded thus resulting in poor generalisation. This process might also explain the result for which infants can generalise based on the fact that they have forgotten the specific details of the original mobile (e.g., Rovee-Collier and Sullivan, 1980).

The design of neural models having generalisation capabilities that are robust with respect to substantial stimulus/context changes is an important investigation known as *transfer learning*. This can involve either perceptual processes (Pan and Yang, 2010) or motor processes (Taylor and Stone, 2009; Tommasino et al., 2016). Transfer problems involve situations where the system learns some "source tasks" and has the opportunity to use the gained knowledge to solve new "target tasks". The possibility of knowledge transfer depends on the similarity between the context and stimuli of the source and target tasks, and higher similarity enhances transfer. Such "similarity" however depends on the level of abstraction with which the context and stimuli are internally represented, in particular higher abstraction implies higher similarity. This might explain why older infants have an enhanced capacity to reuse acquired knowledge in novel situations: younger infants perceive reality through low-level features, so new situations appear to them as completely novel. On the other hand, older infants perceive reality through higher-level, more abstract features that make situations memorable with already acquired knowledge and skills.

A trend in the study of artificial neural networks involves the proposal of systems having increasing ability to disentangle elements of the scene (e.g., objects) from the context. Work on these systems offers possible interpretations of the fact that infants' capacity to recall knowledge on contingencies in different contexts increases with age. Generally, neural

networks and robots engaged in recognizing objects process the whole scene. Currently the most advanced systems that do this, convolutional neural networks, automatically parse the image into regular sub-parts (local fields) and repeat the same processing on each of them (Bengio et al., 2017). Recently, the literature has proposed "regional convolutional neural networks" that do not process all parts of the image regularly, but actively focus on a small number of them that might provide richer information, thus leading to improved performance (see Maiettini et al., 2018 for a review). This process (e.g., for object recognition) is analogous to overt attention that in humans directs the gaze on salient areas of the scene (Wang et al., 2012; Marraffa et al., 2012). Here generalisation is explained by the fact that irrelevant aspects of the scene are ignored altogether (Ognibene and Baldassarre, 2015). Compatible with these ideas, it could be that, at an early age, infants process the whole scene without being able to isolate single elements of it. With development, they might then develop the capacity to focus on limited portions of the scene so that the perception and hence memorisation of contingencies is more effective (in parallel, memory structures might themselves become more robust).

5. Goal-directedness: Is the agent able to use sensorimotor contingencies in order to attain desired states?

When the agent (robot or baby) is exploring its baby gym, how does it decide what to do next? Does the agent *have in mind* a nice sound that it would like to hear? If yes, does this representation of the sound prompt the agent to touch the bell? Or on the contrary, does the agent need to actually see the bell again in order to touch it and produce the sound? In other words, are the agent's actions goal-directed?

5.1. Evidence from developmental psychology

The first issue in answering these questions resides in the definition of goal-directedness (see Kenward et al., 2009 for more detail on different views of goal-directedness). In developmental psychology, goal-directed actions are frequently described with respect to motor planning. An action is considered as goal-directed if the motor sequence is organized with respect to an object, an agent, a gesture or a posture (von Hofsten, 2004). With this definition, goal-directed actions would seem to be already present in very young infants (see for ex. von Hofsten, 2004; Rochat and Striano, 2000). In what follows here, we will characterize goal-directedness with respect to *decision-making* (Dolan and Dayan, 2013). More precisely, we will adopt the framework developed by Baldassarre and his collaborators (Santucci et al., 2016), where a *goal* is defined as a specific state of the world, and a goal-directed action is defined as an action made by an agent to attain this specific state of the world. Hence, to make a goal-directed action, the agent needs to possess two abilities: the ability to select a goal (i.e. a desired state of the world) and the ability to make the action that gives rise to the selected state of the world in the current context. This second ability implies that the agent has previously learned the contingent link between the selected goal and the action. One might say that the agent must have learned the "inverse model" of the contingency with regard to the context. Tackling the question of goal-directedness with respect to decision-making in infancy is challenging because cognitively demanding paradigms from the decision-making literature cannot be used (see Dolan and Dayan, 2013 for a review of these paradigms and Decker et al., 2016 for an attempt of implementation in a child experiment). Nevertheless, developmental psychologists have developed several strategies to explore goal-directedness: (i) means-end behaviours, (ii) anticipation abilities, (iii) bidirectionality of contingency, and (iv) "representation" of the action-outcome coupling. Exceptionally in this section we will discuss data involving infants older than 1 year as a reference for comparison to goal-directed behaviour in younger infants.

5.1.1. Means-end behaviours

A first way researchers have studied goal-directed actions in infants was by looking at infants' ability to act on an object when this generates an effect on a second object — for example, pulling a support/string to retrieve an out-of-reach object or pushing a button to activate a device (e.g., Munakata et al., 1997). Such abilities are referred to as "means-end behaviours". Infants are for example reported to solve the simplest tasks, such as selecting a lever, around 5-6 months of age (e.g., Lobo and Galloway, 2008) and to discover the pulling effect around 7-9 months of age (e.g., Munakata et al., 1997; Willatts, 1999; Babik et al., 2018). Many developmental psychologists (e.g., Piaget, 1936/1952; Diamond, 1991; Willatts, 1999) consider these early means-end behaviours as evidence for "intentionality" and goal-directedness in young infants. Nevertheless, as mentioned previously, certain apparent means-end behaviours could actually simply be stimulus-driven, meaning that the infants associate the view of the object with the pulling/pushing action without the notion of goal being involved (reinforcement learning). One method used to overcome this ambiguity is to offer the infant a choice between a set of potential "means" objects, among which only one allows the "end" object to be retrieved (e.g., a set of strings in which only one is attached to the out-of-reach object). With this protocol, it seems that young infants have great difficulties in choosing the effective "means" object, and if one considers that solving such a multiple means-end problem is a proof of goal-directedness, it seems that goal-directedness is only acquired during the second year of life (e.g., Rat-Fischer et al., 2014; Paulus et al., 2015). The acquisition of goal-directedness has also been studied through investigation of the baby's understanding of other people's means-end behaviours (e.g., Carpenter et al., 2005; Elsner, 2007). Some authors have argued that babies have the ability to perceive other people's actions as *intentional* from an early age, and that they are able to use this early *teleological reasoning* to selectively imitate

other people's actions (e. g. Gergely et al., 2002). However, some authors have questioned the precocity of babies' ability to perceive the actions of others as *intentional* (e.g., Paulus et al., 2013) and have argued that this ability develops gradually during the first years of life as a function of the baby's sensorimotor experience (e.g., Sommerville et al., 2005) — see Paulus and Király, 2013 for a review.

5.1.2. Anticipation abilities

Going a step further in the study of goal-directedness, developmental psychologists have investigated infants' ability to anticipate the appearance of a contingent feedback that follows after realizing an action. Wang et al. (2012) explored this question on 6- and 8-month-old infants using a gaze contingency set-up. In their experiments, infants could activate the appearance of a picture in the middle of the screen by looking at a red disc located on one side of the screen. The authors measured the extent to which the infant anticipated the appearance of the picture — namely if the infant started to look at the middle of the screen directly after looking at the contingent disc but before the appearance of the picture. They found that at both ages, infants were able to learn the contingency to anticipate the contingent feedback. Along the same lines, Kenward (2010) found that 10-month-old infants were able to anticipate the appearance of a video generated by pushing a button. Adopting the assumption that anticipation of an action's effect means that this action was goal-directed, one might consider that infants are able to make goal-directed actions from at least 6 months of age. Nevertheless, anticipation paradigms do not guarantee that infants are capable of *goal selection*. Indeed, one might argue that when an infant anticipates the consequences of its actions, this does not mean that the infant has selected this action to obtain a desired goal (contingent feedback). Instead, the infant could have learnt the contingency by reinforcement learning without any use of the notion of goal. With the aim of obviating this weakness of anticipation paradigms, developmental psychologists have used a paradigm coming from the animal literature and more effectively

testing goal-directedness: the revaluation paradigm (see later). But before addressing this issue, a simpler question to ask is: to what extent is infants' sensitivity to sensorimotor contingencies bidirectional?

5.1.3. *Bidirectionality of contingency learning*

Infants' ability to identify an action from its effects has mainly been investigated by Verschoor and his colleagues (Verschoor et al., 2010; Verschoor et al., 2013). In Verschoor et al. (2010), the authors first exposed 9-, 12- and 18-month-old infants to two sounds: one sound was self-produced by the infant by touching a panel and the other sound was independent of the infant's behaviour. Then, the authors measured infants' propensity to realize the contingent action (i.e. touch the panel) after presentation of the contingent or the non-contingent sound. The authors found that infants at all ages were faster to produce the contingent action in response to the contingent sound than to the non-contingent sound, suggesting that from 9 months of age, infants are able to identify an action from its effects. In their second study (Verschoor et al., 2013), the authors exposed 7-, 12-month-old infants (and adults) to two contingencies, each associating respectively a right and a left saccade to a different sound. Then, they measured participants propensity to make the "good" contingent action (right or left saccade) after being exposed to its previously learnt effect (one of the two sounds). They found that the older infants (12 months of age) and adult were able to identify an action from its effect whereas younger infants (7 months of age) were not able to do so. Interestingly, they found that both 7- and 12-month-old infants were surprised when they made the "wrong" action, suggesting that 7-month-old infants formed bidirectional knowledge of the contingencies but were not able to use this knowledge in action selection.

5.1.4. "Representation" of the action-outcome coupling

The studies presented above suggested that after 9 months of age, infants are able to "predict" the consequences of their actions and to identify an action from its effects. However, one question remained open: Are infants able to represent a desired effect *in their mind*, and to make an action in order to obtain this effect? To investigate this question, developmental psychologists have used the revaluation paradigm (see Table 1 for a description). Klossek et al. (2008) tested to what extent infants' sensitivity to the revaluation paradigm evolves in the second and third year of life. They tested three groups of infants: 16 to 27, 27 to 37 and 37 to 48 months of age. All age groups were able to learn two contingencies in an acquisition phase and were sensitive to revaluation (i.e. inducing a preference for one of the stimuli generated by the contingent actions). But, interestingly, in a test phase, only the two older groups of infants, from 27 to 48 months of age, were able to show their preference in stimulus (goal selection) by making the correct action (action selection). This last result means that the infants younger than 2 years of age were not able to make goal-directed actions, and remained predominantly stimulus-driven in their behaviours. To ensure that this inability to make goal-directed actions in younger infants was not due to the complexity of the task, Kenward et al. (2009) developed a simplified version of the revaluation paradigm. In their setup, the infants had to learn only one contingency in the acquisition phase. Then, for one group of infants (the revaluation group) the stimulus used as contingent feedback in the acquisition phase was increased in value, while a control group were occupied by an independent irrelevant task. In the test phase, infants' propensity to repeat the contingent action in the absence of feedback was compared between the revaluation and control groups. The authors tested infants of three ages 14, 19 and 24 months. Here again, the authors found that infants at all ages were able to learn the contingency in the acquisition phase and were sensitive to the revaluation. But the authors observed that only at the oldest age (24 months), did infants trained in the "revaluation group" show a higher

number of contingent actions during test compared to the control group. If one accepts that the revaluation paradigm is a test of goal-directedness, it seems that infants only become able to use goals to select their actions after 2 years of age.

5.1.5. *Conclusion from developmental psychology*

Taken together these studies suggest that infants' ability to make goal-directed actions emerges progressively during the first three years of life. Young infants (before approximately 9 months of age) seem sensitive to sensorimotor contingencies, and even seem able to predict the consequences of their actions, but without being able to use this knowledge in order to select their actions. Thus, they are capable of goal selection (since they can develop a preference for a stimulus) but not of action selection. Then, at the end of their first year of life, infants become able to select their actions in a stimulus-driven fashion, since they can recall an action when they first perceive the stimulus (e.g., move to make a static mobile move) but not when they have to cause presentation of the desired stimulus (e.g., obtain a specific sound). Finally, at the end of their second year of life, infants become able to select an action in order to attain a desired (and not necessarily currently perceived) state of the world. Thus, one might conclude that it is only at this age that infants become really capable of model-based decision making. Note that this progressive pattern of acquisition might possibly be related to the progressive acquisition of inhibitory abilities in infancy.

5.2. Approaches in robotics

Computational modelling confirms that observing anticipatory behaviour like that observed in the gaze contingency experiments of Wang et al. (2012) does not imply that the infants are goal-directed. In this respect, Marraffa et al. (2012) presented a model that mimics the behaviour of the infants when tested in a simulated version of the set-up used in the experiment of Wang and colleagues. Like the infants, the model learns by reinforcement learning to gaze

at salient parts of a (simulated) screen to cause the appearance of a rewarding image. The model incoerporates bottom-up and top-down attention mechanisms that allow it to learn as fast as real infants. The model also exhibits the same anticipatory gazing behaviour as the infants. This is explained by the fact that with learning as soon as the simulated infant sees the cue associated with the appearance of the image, it immediately looks at the place where the image will appear as this is the place to look at to receive the reward. Of course, this only shows that reactive behaviour and reinforcement learning are sufficient to produce the anticipatory behaviour seen in infants, but does not rule out the possibility that they rely on goal-directed mechanisms. Regarding the experiments of Verschoor and colleagues (Verschoor et al., 2010; Verschoor et al., 2013) on the bidirectionality of contingency learning, in future work it would be interesting to check which mechanisms are needed to reproduce the observed behaviours and whether they represent a precursor of more sophisticated goal-directed processes.

While goals are a fundamental element in classic artificial intelligence planning systems (Russell and Norvig, 2010), their use in robotic and machine learning neural-network systems has been limited by the fact that the focus is usually on the solution of only one user assigned task. In this context, goals have been used only recently to face situations involving multiple tasks where robots need to decide on which of them to focus (Hart and Grupen, 2011; Santucci et al. 2012). In particular, in developmental robotics goals are crucial to support the learning of autonomous agents as they allow them to decide on which task to focus on the basis of intrinsic motivations (Santucci et al., 2012). Moreover, goals are also important to explore the task space (Cartoni and Baldassarre, 2018; Nair et al., 2018) and, as described in the framework of goal babbling (Rolf et al., 2010), they can possibly contribute to reduce such space thereby avoiding the "curse of dimensionality" problem (Bishop, 1995). On this basis, different artificial agents have been proposed that are able to autonomously discover and select their own goals and use them to drive the learning process (e.g., Baranes and Oudeyer, 2013; Seepanomwan et al.,

2017). While it can be complicated to identify goal-directed processes in very young babies, robotic systems such as GRAIL (Santucci et al. 2016) are endowed with the possibility of storing and explicitly selecting interesting possible environment states, thus turning them into goals that can determine action selection and skill learning. These robotic models suggest that the possibility of performing goal-directed behaviour should go hand-in-hand with the possibility of forming representations of objects/states/events that can be reactivated in the absence of corresponding entities in the world. This allows an agent to act so as to increase the chance of obtaining "desired outcomes" (goals). This ability has to be paired to a "goal-matching" ability that compares the actual state of the world with the pursued goal so as to detect when it is achieved: this mechanism is a possible way to computationally operationalise how contingency detection might guide the acquisition of skills, as illustrated in detail in the following section.

6. Blueprint architecture: a theoretical integration of the elements of contingency-based learning

In the previous sections we have considered multiple elements involved in contingency-based learning. We now consider a "blueprint architecture" representing a theoretical scheme that integrates these elements in a whole framework (Figure 1). A first version of the architecture was proposed in Baldassarre et al. (2018). The architecture proposes a possible specification of the contingency-based learning mechanisms and processes considered this far, and also adds to them further mechanisms needed to have a fully autonomous agent. The value of the architecture is that it proposes a possible way to integrate such processes, thus offering a coherent view of how they could work together. The architecture can also inform the construction of specific computational models, as was done in Mannella et al. (2018).
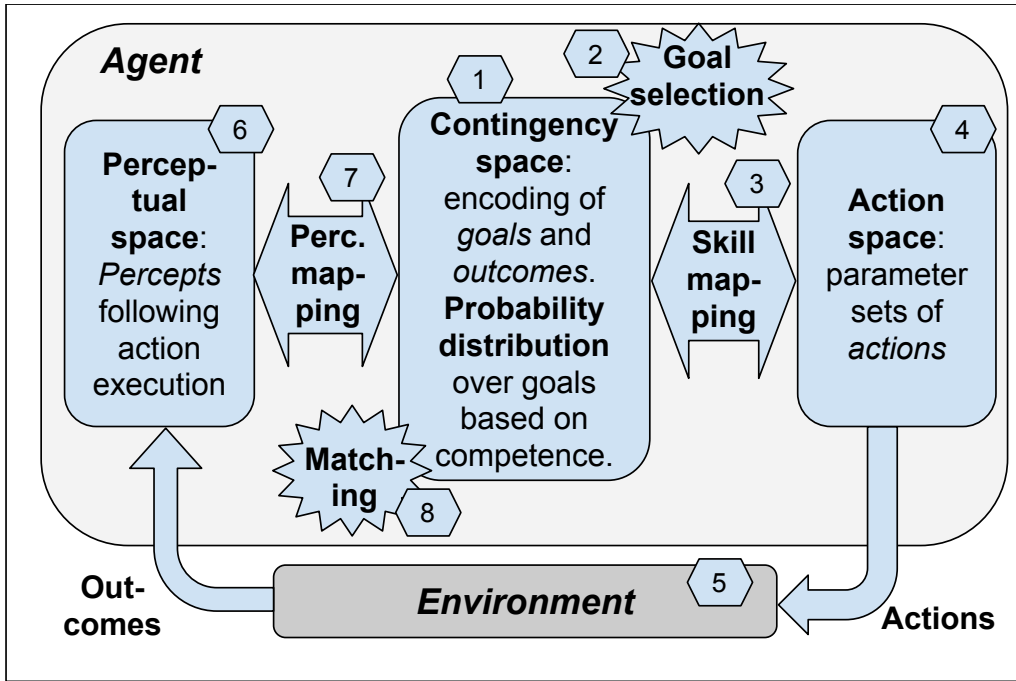
Figure 1: The blueprint architecture incorporating our hypothesis about the key elements underlying open-ended learning of multiple skills. Boxes: the components of the architecture. Numbers: sequence of processes happening in one trial of functioning of the system.

The architecture is based on the hypothesis that contingency-based learning encompasses general mechanisms able to support the self-generation of goals, the encoding of the perceptual consequences of actions, and the learning of actions of any level of complexity, e.g., from moving a single finger to performing grasping with the whole hand. To illustrate the architecture, we will focus on a case involving learning to reach parts of the agent's own body, a relevant form of knowledge that infants acquire soon after birth. However, we hypothesise that analogous contingency-based learning processes might support the acquisition of more sophisticated classes of goals and actions.

A key idea at the core of the architecture is that contingency-based learning relies on a specific type of events that are *internal* to the agent and reflect the sensorimotor contingencies happening in the environment, i.e. the motor action performed by the body and the physical consequences that it causes in the environment. These events, called "matching events",

correspond to a match between the selected goal encoding (encoding of the desired perceptual effect of the action), and the outcome encoding (encoding of the actual perceptual effect of the action). For example, intrinsic motivation might lead an infant to select the goal of touching a part of her chest. The goal activation might trigger the performance of an action directed to accomplish it. The performed action leads to a perceived effect in the environment (e.g., feeling of body touch). The perceived effect is internally represented in the brain and may or may not match the representation of the pursued goal. In case of "success" (the action produced the desired outcome), the matching event plays a crucial role in the architecture as it contributes (a) to modify the intrinsic motivation that leads to goal-selection, and (b) to guide the learning of perceptual mapping, the learning of the skill mapping, and a learning process leading to align the goal encoding and the outcome encoding within the contingency space. We now review these processes in more detail.

The architecture is formed by the following main components:

(a) *Contingency space*: a set of vectors (e.g., related to the possible activations of a group of neurons; the vectors are the points forming the space) each of which can represent two things: (1) a goal encoding the desired perceptual effect of the action triggered by the goal (where the "perceptual effect" can involve proprioception, perception of the outer world, or both); (2) the encoding of the actual effect resulting from the performance of an action. Importantly, as discussed below, with learning the system aims to progressively drive each goal (encoded in the contingency space) closer to the outcome (as encoded in the contingency space) of the action triggered by the goal. The contingency space can be either discrete or continuous.

(b) *Goals*: points of the contingency space that represent *desired* perceptual effects of actions, and can be internally selected and activated to trigger the related action.

(c) *Goal probability distribution*: a probability distribution over goals, used to select the goal to pursue, and that depends on intrinsic motivations.

(d) *Skill mapping*: mapping from the selected goal within the contingency space, to the action directed to pursue it.

(e) *Action space*: set of action parameters where one parameter combination represents one action that can be selected by the skill mapping.

(f) *Perception space*: set of sensor encodings of the possible effects (outcomes) caused by actions in the environment.

(g) *Perception mapping*: mapping from the percepts resulting from action performance to their encoding within the contingency space.

The architecture works and learns in trials, and for each trial the following processes take place (see Figure 1, numbers in hexagons):

(1) Initially the skill mapping and the perception mapping are randomly initialised so the points in the contingency space have random correspondences to the actions encoded in the action space and to percepts encoded in the perceptual space.

(2) At the beginning of each trial the agent selects and activates a goal in the contingency space on the basis of a certain probability distribution explained below.

(3) The skill mapping generates an action parameter set encoded in the action space.

(4) The agent performs in the environment the action corresponding to the selected parameter set.

(5) As an effect of the action, the environment changes into a certain state (outcome).

(6) The outcome is encoded as a point in the perceptual space.

(7) The point in the perceptual space corresponding to the action outcome is encoded into a point in the contingency space that constitutes the internal representation of the outcome.

(8) The internal outcome representation can match the goal or not (if the contingency space is discrete), or it can have a certain degree of similarity with it (if the contingency space is continuous). This allows the computation of a matching signal: this signal has a discrete value in $\{0, 1\}$ in the case the contingency space is discrete, or a continuous value within $[0, 1]$ in the case the contingency space is continuous (in this case it is computed on the basis of a suitable outcome/goal metric).

The matching value is used to determine the probability distribution over the goals that supports the selection of the goal to pursue (number 2 in the figure). The probability distribution is generated on the basis of intrinsic motivations, for example related to competence (e.g., the goals with a lower competence, or with a higher competence-improvement, have a higher probability of selection; Santucci et al., 2013).

The matching value is also used as a learning signal to drive the learning processes involving the perceptual mapping and the skill mapping. In particular, it is used as a reward to: (a) find the action parameters by reinforcement learning, and use them to update the skill mapping; (b) update the perceptual mapping so that the percepts that are similar in the perceptual mapping connect to similar points in the contingency space; (c) at the same time, update the perceptual mapping and the skill mapping so that the internal outcome representation matches, or becomes more similar, to the goal representation. Through these learning processes, the internal outcome representation and the goal representation tend to align, i.e. each goal tends to produce an action that causes an outcome in the world that once perceived is encoded in a point in the perceptual space that matches the goal. Through these processes, the agent discovers the possible outcomes that it can produce in the environment, learns to encode them in the contingency space, and acquires the motor capabilities needed to accomplish such outcomes when desired.

An issue not considered this far involves the agent's capacity to understand that a certain effect observed in the environment is caused by the agent's own action and not by the spontaneous evolution of the environment or by other agents. In this respect, we have so far assumed that all the effects that the agent observes are caused by its own actions. This is a good assumption in some conditions (imagine an infant exploring her body while in the cradle) but not always. The capacity we are discussing is linked to *causality* and can be detected as the difference between the probability that a certain outcome happens after a certain action is performed and the probability that the outcome happens if the action is not performed (Gopnik et al., 2004). To check the effectiveness of actions, the architecture discussed above should be enhanced with suitable mechanisms to observe the environment multiple times after having executed an action or not, and to compare the probability of achieving the outcome in the two cases (for some computational models of this, see: Pitti et al., 2009; Butko and Movellan, 2010; Sperati and Baldassarre, 2017).

Another very important issue not considered this far concerns the fact that the actions produced in the environment can produce different effects on the body and the environment and so the agent should be endowed with a mechanism to decide which effects (action-effect contingencies) are `interesting' and so worth triggering the learning processes illustrated above. One possibility is that the agent considers as interesting all `effects' corresponding to a *change* in the environment; the rationale of this is that all actions causing a change in the body/environment might be useful in the future (Barto et al., 2004; Santucci et al., 2016; Sperati and Baldassarre, 2017). Another possibility is to use intrinsic motivations such as *novelty* or *surprise*, i.e. to consider as interesting those effects of actions that respectively have never/rarely been experienced before, or that violate expectations (Barto et al., 2013). A last possibility is that the agent considers as interesting those action-effects that might be useful to achieve other effects if used to form action sequences (various techniques can be used to this

purpose, e.g., see McGovern and Barto, 2001). Overall, this is an important open issue at the core of development and open-ended learning in robotics that deserves further exploration in the future.

7. Conclusion and perspectives

The sensitivity of an agent to the links between its actions and their consequences is central in the theories of psychology that roboticians use in their work, such as the sensorimotor contingency theory (e.g., O'Regan, 2011), the Gibsonian theory of affordances (Gibson, 1977) or the predictive coding theory (e.g., Clark, 2013). Sensitivity to sensorimotor contingencies is particularly important in developmental robotics because it provides a simple way to equip an agent with the ability to learn to interact with the world using self-organized exploration of its environment. In this paper, we have strengthened this statement by demonstrating how the use of this sensitivity by babies allows them to acquire and/or mature new motor and cognitive abilities such as body knowledge, memory, generalisation and goal-directedness. In addition, our literature review highlights the links between conclusions obtained in developmental psychology and current work in robotics, and thus offers readers from robotics a concrete and direct application to their field of study.

The model proposed at the end of the article shows how exploitation of sensitivity to sensorimotor contingencies, combined with the notion of "goal" (i.e. a desired state of the world), allows an agent to organize its exploration of the world and develop new sensorimotor skills, in the present case the skills that underlie understanding of the structure of its body. This model is conceived to be suitable for all types of sensorimotor learning, and we hope that in the future it will be adapted both to better understand baby development and to build robots capable of more effective autonomous learning.

This article is also intended to highlight the value of close collaboration between developmental psychology and developmental robotics, as such a collaboration will (i) fill the

lack of mechanistic approaches in development psychology, and (ii) provide new ideas and concrete data that roboticians can use to develop open-ended learning. We hope that our article contributes to the emergence of a common language between the two fields that will facilitate future fruitful collaboration.

Acknowledgments

References

Adler, S. A., & Rovee-Collier, C. (1994). The memorability and discriminability of primitive perceptual units in infancy. *Vision Research*, *34*(4), 449–459.

Adolph, K. E., Eppler, M. A., Marin, L., Weise, I. B., & Wechsler Clearfield, M. (2000). Exploration in the service of prospective control. *Infant Behavior and Development*, *23*(3–4), 441–460. https://doi.org/10.1016/S0163-6383(01)00052-2

Angulo-Kinzler, R. M. (2001). Exploration and Selection of Intralimb Coordination Patterns in 3-Month-Old Infants. *Journal of Motor Behavior*, *33*(4), 363–376. https://doi.org/10.1080/00222890109601920

Angulo-Kinzler, R. M., Ulrich, B., & Thelen, E. (2002). Three-Month-Old Infants Can Select Specific Leg Motor Solutions. *Motor Control*, *6*, 52–68.

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48–53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Babik, I., Cunha, A. B., Ross, S. M., Logan, S. W., Galloway, J. C., & Lobo, M. A. (2018). Means-end problem solving in infancy: Development, emergence of intentionality, and transfer of knowledge. *Developmental Psychobiology*, *61*(2), 191–202. https://doi.org/10.1002/dev.21798

Baldassarre, G., Mannella, F., Santucci, V. G., Somogyi, E., Jacquey, Hamilton, M. D. and O'Regan, J.K. (2018a), Action-Outcome Contingencies as the Engine of Open-Ended Learning: Computational Models and Developmental Experiments. In *Development and Learning and Epigenetic Robotics (ICDL), 2018 IEEE International Conference*. IEEE.

Baranes, A., and Oudeyer, P. Y. (2013). Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robotics and Autonomous Systems*, *61*(1), 49-73.

Barto, A. G., Singh, S., and Chentanez, N. (2004). Intrinsically motivated learning of hierarchical collections of skills. In *Proceedings of the 3rd International Conference on Development and Learning* (pp. 112-19).

Barto, A., Mirolli, M. and Baldassarre, G. (2013). Novelty or surprise?. *Frontiers in Psychology*, *4* (907).

Begum Ali, J., Bremner, A. J., & Spence, C. (2015). Human infants' ability to perceive touch in external space develops postnatally. *Current Biology*, *25*, R978–R979.

Bengio, Y., Goodfellow, I. J. and Courville, A. (2017). *Deep Learning*. Boston, MA: The MIT Press.

Berlyne, D. E. "*Conflict, arousal, and curiosity.*" (1960), McGraw Hill.

Bishop, C. M. (1995), *Neural networks for pattern recognition*, Oxford University Press.

Bremner, A. J., Mareschal, D., Lloyd-Fox, S., and Spence, C. (2008). Spatial localization of touch in the first year of life: Early influence of a visual spatial code and the development of remapping across changes in limb position. *Journal of Experimental Psychology: General, 137*(1), 149–162. https://doi.org/10.1037/0096-3445.137.1.149

Butko, N. J. and Movellan, J. R. (2010). Detecting contingencies: an infomax approach. *Neural Networks, 23*, 973-984.

Butler, J., and Rovee-Collier, C. K. (1989). Contextual gating of memory retrieval. *Developmental Psychobiology, 22*(6), 533–552.

Caligiore, D.; Parisi, D. & Baldassarre, G. (2014). Integrating reinforcement learning, equilibrium points, and minimum variance to understand the development of reaching: a computational model. *Psychological Review*, *121*, 389-421.

Cangelosi, A., and Schlesinger, M. (2015). *Developmental robotics: from babies to robots. Cambridge*, MA, USA: The MIT Press.

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science*, *8*(1), F13–F20. https://doi.org/10.1111/j.1467-7687.2004.00385.x

Cartoni, E., and Baldassarre, G. (2018). Autonomous discovery of the goal space to learn a parameterized skill. *arXiv [cs.AI] 1805.07547v1*.

Chen, Y.-P., Fetters, L., Holt, K. G., and Saltzman, E. (2002). Making the mobile move: Constraining task and environment. *Infant Behavior and Development, 25*(2), 195–220.

Chinn, L. K., Noonan, C. F., Hoffmann, M., and Lockman, J. J. (2019). Development of Infant Reaching Strategies to Tactile Targets on the Face. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.00009

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Colas, C., Sigaud, O., and Oudeyer, P.-Y. (2018). CURIOUS: Intrinsically Motivated Multi-Task, Multi-Goal Reinforcement Learning. *arXiv preprint arXiv:1810.06284*.

Corbetta, D., and Snapp-Childs, W. (2009). Seeing and touching: The role of sensory-motor experience on the development of infant reaching. *Infant Behavior and Development*, *32*(1), 44–58. https://doi.org/10.1016/j.infbeh.2008.10.004

Da Silva, B. C., Baldassarre, G., Konidaris, G., and Barto, A. (2014, May). Learning parameterized motor skills on a humanoid robot. In *2014 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 5239-5244). IEEE.

DeCasper, A. J., and Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, *9*(2), 133–150.

Decker, J. H., Otto, A. R., Daw, N. D., and Hartley, C. A. (2016). From Creatures of Habit to Goal-Directed Learners: Tracking the Developmental Emergence of Model-Based Reinforcement Learning. *Psychological Science*, *27*(6), 848–858. https://doi.org/10.1177/0956797616639301

Diamond, A. (1991). Neuropsychological insights into the meaning of object concept development. In S. Carey & R. Gelman (Eds.), *The Jean Piaget Symposium series. The epigenesis of mind: Essays on biology and cognition* (pp. 67-110). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Dolan, R. J., and Dayan, P. (2013). Goals and Habits in the Brain. *Neuron, 80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

D'Souza, H., Cowie, D., Karmiloff-Smith, A., & Bremner, A. J. (2017). Specialization of the motor system in infancy: from broad tuning to selectively specialized purposeful actions. *Developmental science, 20*(4), e12409.

Duminy, N., Nguyen, S. M., and Duhaut, D. (2018). Learning a set of interrelated tasks by using a succession of motor policies for a socially guided intrinsically motivated learner. *Frontiers in neurorobotics*, *12*.

Elsner, B. (2007). Infants' imitation of goal-directed actions: The role of movements and action effects. *Acta Psychologica*, *124*(1), 44–59. https://doi.org/10.1016/j.actpsy.2006.09.006

Fagen, J. W., and Rovee, C. K. (1976). Effects of Quantitative Shifts in a Visual Reinforcer on the Instrumental Response of Infants. *Journal of Experimental Child Psychology, 21*, 349–360.

Fagen, J. W., Rovee, C. K., and Kaplan, M. G. (1976). Psychophysical Scaling of Stimulus Similarity in 3-Month-Old Infants and Adults. *Journal of Experimental Child Psychology, 22*, 272–281.

Fagen, J. W., and Rovee-Collier, C. K. (1983). Memory Retrieval: A Time-Locked Process in Infancy. *Science, 222*, 1349–1351.

Forestier, S., and Oudeyer, P. Y. (2016, August). Curiosity-driven development of tool use precursors: a computational model. In *38th annual conference of the cognitive science society (cogsci 2016)* (pp. 1859-1864).

Forestier, S., Mollard, Y., and Oudeyer, P. Y. (2017). Intrinsically motivated goal exploration processes with automatic curriculum learning. *arXiv preprint arXiv:1708.02190.*

French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends in cognitive sciences, 3*(4), 128-135.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755–755. https://doi.org/10.1038/415755a

Gergely, G., and Watson, J. S. (1996). The Social Biofeedback theory of parental affect-mirroring: the development of emotional self-awareness and self-control in infancy. *Journal of Psycho-Analysis*, *77*, 1181–1211.

Gibson, J. J. (1977) The theory of affordances. in R. Shaw and J. Bransford, Eds. *Perceiving, Acting, and Knowing: Toward an ecological psychology.* (pp. 67–82). Hillsdale, NJ: Erlbaum.

Goldstein, M. H., and Schwade, J. A. (2008). Social Feedback to Infants' Babbling Facilitates Rapid Phonological Learning. *Psychological Science, 19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Gómez, R. L., and Edgin, J. O. (2015). Sleep as a Window Into Early Neural Development: Shifts in Sleep-Dependent Learning Effects Across Early Childhood. *Child Development Perspectives, 9*(3), 183–189. https://doi.org/10.1111/cdep.12130

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T. and Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review, 111*, 3.

Guerin, F., Kruger, N., and Kraft, D. (2013). A Survey of the Ontogeny of Tool Use: From Sensorimotor Experience to Planning. *IEEE Transactions on Autonomous Mental Development, 5*(1), 18–45. https://doi.org/10.1109/TAMD.2012.2209879

Hafez, M.B., Weber, C., and Wermter, S. (2017) "Curiosity-driven exploration enhances motor skills of continuous actor-critic learner," in *Proceedings of the 7th Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*.

Hart, S. and Grupen, R. (2011). Learning generalizable control programs. *IEEE Transactions on Autonomous Mental Development, 3*, 216-231.

Hartshorn, K., and Rovee-Collier, C. K. (1997). Infant learning and long-term memory at 6 months: A confirming analysis. *Developmental Psychobiology, 30*(1), 71–85.

Hartshorn, K., Rovee-Collier, C. K., Gerhardstein, P., Bhatt, R. S., Klein, P. J., Aaron, F., … Wurtzel, N. (1998). Developmental Changes in the Specificity of Memory over the First Year of Life. *Developmental Psychobiology, 33*, 61–78.

Hayne, H., Greco, C., Earley, L. A., Griesler, P. C., and Rovee-Collier, C. K. (1986). Ontogeny of Early Event Memory: II. Encoding and Retrieval by 2- and 3-Month-Olds. *Infant Behavior and Development, 9*, 461–472.

Hayne, H., Rovee-Collier, C. K., and Borza, M. A. (1991). Infant memory for place information. *Memory and Cognition, 19*(4), 378–386. https://doi.org/10.3758/BF03197142

Heathcock, J. C., Bhat, A. N., Lobo, M. A., and Galloway, J. C. (2005). The Relative Kicking Frequency of Infants Born Full-term and Preterm During Learning and Short-term and Long-term Memory Periods of the Mobile Paradigm. *Physical Therapy, 85,* 8–18.

Herbert, J., Gross, J., & Hayne, H. (2007). Crawling is associated with more flexible memory retrieval by 9-month-old infants. *Developmental Science, 10*(2), 183-189.

Hester, T., and Stone, P. (2017). Intrinsically motivated model learning for developing curious robots. *Artificial Intelligence*, *247*, 170-186.

Hoffmann, M., Chinn, L., Somogyi, E., Heed, T., Fagard, J., Lockman, J. J., and O'Regan, J. K. (2017). Development of reaching to the body in early infancy: From experiments to robotic models. 2017 *Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 112–119. https://doi.org/10.1109/DEVLRN.2017.8329795

Hoffmann, M., Marques, H., Arieta, A., Sumioka, H., Lungarella, M., & Pfeifer, R. (2010). Body Schema in Robotics: A Review. *IEEE Transactions on Autonomous Mental Development, 2*(4), 304–324. https://doi.org/10.1109/TAMD.2010.2086454

Hoffmann, M., and Pfeifer, R. (2018). Robots as powerful allies for the study of embodied cognition from the bottom up. *arXiv preprint arXiv:1801.04819*.

Jacquey, L., Popescu, T. S., Vergne, J., Fagard, J., Esseily, R. & O'Regan, J. K. (*submitted*) Development of body knowledge as measured by arm differentiation in infants: from global to local?

Jamone, L., Bradao M., Natale L., Hashimoto K., Sandini G., Takanishi A. (2014). Autonomous online generation of a motor representation of the workspace for intelligent whole-body reaching. *Robot. Auton. Syst. 6*2, 556–567.

Jamone, L., Ugur, E., Cangelosi, A., Fadiga, L., Bernardino, A., Piater, J., and Santos-Victor, J. (2016). Affordances in psychology, neuroscience and robotics: a survey. *IEEE Transactions on Cognitive and Developmental Systems, 1*–1. https://doi.org/10.1109/TCDS.2016.2594134

Juett, J., and Kuipers, B. (2019). Learning and Acting in Peripersonal Space: Moving, Reaching, and Grasping. *Frontiers in Neurorobotics*, *13*.

Kato, M., Watanabe, H., and Taga, G. (2013). Diversity and changeability of infant movements in a novel environment. *Journal of Motor Learning and Development, 1*, 79–88.

Kenward, B. (2010). 10-Month-Olds Visually Anticipate an Outcome Contingent on Their Own Action. *Infancy, 15*(4), 337–361. https://doi.org/10.1111/j.1532-7078.2009.00018.x

Kenward, B., Folke, S., Holmberg, J., Johansson, A., and Gredebäck, G. (2009). Goal directedness and decision making in infants. *Developmental Psychology, 45*(3), 809.

Klossek, U. M. H., Russell, J., and Dickinson, A. (2008). The control of instrumental action following outcome devaluation in young children aged between 1 and 4 years. *Journal of Experimental Psychology: General, 137*(1), 39–51.

Kohonen, T. (2001). *Self-organizing maps*. Berlin: Springer-Verlag.

Lee, M.-H., and Newell, K. M. (2013). Contingent auditory feedback of arm movement facilitates reaching behavior in infancy. *Infant Behavior and Development, 36*(4), 817–824. https://doi.org/10.1016/j.infbeh.2013.09.006

Lobo, M. A., and Galloway, J. C. (2008). Postural and Object-Oriented Experiences Advance Early Reaching, Object Exploration, and Means–End Behavior. *Child Development, 79*(6), 1869–1890.

Maiettini, E., Pasquale, G., Rosasco, L., and Natale, L. (2018). Speeding-up Object Detection Training for Robotics with FALKON. *arXiv preprint: 1803.08740*

Mannella, F., Santucci, V.G., Somogy,i E., Jacquey, L., O'Regan, K.J., and Baldassarre, G. (2018). Know Your Body Through Intrinsic Goals, *Frontiers in Neurorobotics* 12:30.

Marraffa, R., Sperati, V., Caligiore, D., Triesch, J., and Baldassarre, G. (2012, November). A bio-inspired attention model of anticipation in gaze-contingency experiments with infants. In

Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on (pp. 1-8). IEEE.

Mast, V. K., Fagen, J. W., Rovee-Collier, C. K., and Sullivan, M. W. (1980). Immediate and Long-Term Memory for Reinforcement Context: The Development of Learned Expectancies in Early Infancy. *Child Development, 51*(3), 700. https://doi.org/10.2307/1129455

McClelland, J.L., Rumelhart, D. E., PDP Research Group (1986). *Parallel distributed processing: explorations in the microstructure of cognition.* Boston, MIT Press.

McGovern, A. and Barto, A.G. (2001). Automatic discovery of subgoals in reinforcement learning using diverse density. *International Conference on Machine Learning (ICML 2001)*.

Merrick, K. E. (2012). Intrinsic motivation and introspection in reinforcement learning. *IEEE Transactions on Autonomous Mental Development*, *4*(4), 315-329.

Munakata, Y., McClelland, J. L., Johnson, M. H., and Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 10*4(4), 686–713. https://doi.org/10.1037/0033-295X.104.4.686

Myowa-Yamakoshi, M., and Takeshita, H. (2006). Do Human Fetuses Anticipate Self-Oriented Actions? A Study by Four-Dimensional (4D) Ultrasonography. *Infancy, 10*(3), 289–301.

Nagai, Y., and Asada, M. (2015). Predictive learning of sensorimotor information as a key for cognitive development. *Proc. of the IROS 2015 Workshop on Sensorimotor Contingencies for Robotics.*

Nair, A., Pong, V., Dalal, M., Bahl, S., Lin, S., and Levine, S. (2018). Visual Reinforcement Learning with Imagined Goals. *The Second Lifelong Learning: A Reinforcement Learning Approach Workshop (LLRLA 2018 at FAIM 2018).*

Needham, A., Barrett, T., & Peterman, K. (2002). A pick-me-up for infants' exploratory skills: Early simulated experiences reaching for objects using 'sticky mittens' enhances young infants' object exploration skills. *Infant Behavior and Development*, *25*(3), 279–295. https://doi.org/10.1016/S0163-6383(02)00097-8

Needham, A., Joh, A. S., Wiesen, S. E., and Williams, N. (2014). Effects of Contingent Reinforcement of Actions on Infants' Object-Directed Reaching. *Infancy, 19*(5), 496–517. https://doi.org/10.1111/infa.12058

Ognibene, D., & Baldassarre, G. (2014). Ecological active vision: Four bioinspired principles to integrate bottom–up and adaptive top–down attention tested with a simple camera-arm robot. *IEEE transactions on autonomous mental development, 7*(1), 3-25.

Oudeyer, P.-Y., Kaplan, F., Hafner, V. V., and Whyte, A. (2005). The Playground Experiment: Task-Independent Development of a Curious Robot. *Proceedings of the AAAI Spring Symposium on Developmental Robotics,* 42–47.

Oudeyer, P.-Y., Gottlieb, J., and Lopes, M. (2016). Intrinsic motivation, curiosity, and learning. In *Progress in Brain Research* (Vol. 229, pp. 257–284).

O'Regan, J. K. (2011). *Why red doesn't sound like a bell: Understanding the feel of consciousness*. Oxford University Press.

Pan, S. J., and Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering, 22*, 1345-1359.

Paulus, M., Hunnius, S., & Bekkering, H. (2013). Examining functional mechanisms of imitative learning in infancy: Does teleological reasoning affect infants' imitation beyond motor resonance? *Journal of Experimental Child Psychology, 116*(2), 487–498. https://doi.org/10.1016/j.jecp.2012.10.009

Paulus, M., & Király, I. (2013). Early rationality in action perception and production? A theoretical exposition. *Journal of Experimental Child Psychology, 116*(2), 407–414. https://doi.org/10.1016/j.jecp.2013.01.001

Paulus, M., & Sodian, B. (2015). Which way to take? Infants select an efficient path to their goal. *Journal of Experimental Child Psychology, 137*, 111–124. https://doi.org/10.1016/j.jecp.2015.03.014

Piaget, J. (1936/1952). The origins of intelligence in children [La naissance de l'intelligence chez l'enfant] (M. Cook, Trans.). New York: Basic Books

Pitti, A., Mori, H., Kouzuma, S. and Kuniyoshi, Y. (2009). Contingency perception and agency measure in visuo-motor spiking neural networks. *IEEE Transactions on Autonomous Mental Development, 1*, 86-97.

Rat-Fischer, L., Fagard, J., and O'Regan, J. K. (2014). Comparison of active and purely visual performance in a multiple-string means-end task in infants. *Cognition, 133*(1), 304–316. https://doi.org/10.1016/j.cognition.2014.06.005

Reinhart, R. F. (2017). Autonomous exploration of motor skills by skill babbling. *Autonomous Robots*, *41*(7), 1521-1537.

Rochat, P., and Striano, T. (2000). Perceived self in infancy. *Infant Behavior and Development, 23*(3), 513–530.

Rolf, M., Steil, J. J. and Gienger, M. (2010), Goal babbling permits direct learning of inverse kinematics, *Autonomous Mental Development, IEEE Transactions on* 2(3), 216–229.

Roncone, A., Hoffmann, M., Pattacini, U., and Metta, G. (2014, May). Automatic kinematic chain calibration using artificial skin: self-touch in the icub humanoid robot. In *2014 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2305-2312). IEEE.

Rovee, C. K., and Rovee, D. T. (1969). Conjugate reinforcement of infant exploratory behavior. *Journal of Experimental Child Psychology, 8*(1), 33–39.

Rovee-Collier, C. K. (1996). Shifting the focus from what to why. *Infant Behavior and Development, 19*(4), 385–400.

Rovee-Collier, C. K. (1999). The Development of Infant Memory. *Current Directions in Psychological Science, 8*(3), 6.

Rovee-Collier, C. K., and Dufault, D. (1991). Multiple contexts and memory retrieval at three months. *Developmental Psychobiology, 24*(1), 39–49.

Rovee-Collier, C. K., Morrongiello, B. A., Aron, M., and Kupersmidt, J. (1978). Topographical Response Differentiation and Reversal in 3-Month-Old Infants. *Infant Behavior and Development, 1*, 323–333.

Rovee-Collier, C. K., & Sullivan, M. W. (1980). Organization of Infant Memory. *Journal of Experimental Psychology: Human Learning and Memory, 6*(6), 798–807.

Rovee-Collier, C. K., Sullivan, M. W., Enright, M., Lucas, D., and Fagen, J. (1980). Reactivation of Infant Memory. *Science, 208*(4448), 1159–1161.

Russell, S. J. and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach.* New York, Pearson Education.

Santucci, V.G., Baldassarre, G., and Cartoni, E. (*2019*). Autonomous Reinforcement Learning of Multiple Interrelated Tasks. In *2019 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*

Santucci, V. G., Baldassarre, G., & Mirolli, M. (2012, November). Intrinsic motivation mechanisms for competence acquisition. In 2012 *IEEE International Conference on Development and Learning and Epigenetic Robotics* (ICDL) (pp. 1-6). IEEE.

Santucci, V. G., Baldassarre, G., and Mirolli, M. (2013). Which is the best intrinsic motivation signal for learning multiple skills?. *Frontiers in neurorobotics*, *7*, 22.

Santucci, V. G., Baldassarre, G., and Mirolli, M. (2014a). Cumulative learning through intrinsic reinforcements. In *Evolution, Complexity and Artificial Life* (pp. 107-122). Springer, Berlin, Heidelberg.

Santucci, V. G., Baldassarre, G., and Mirolli, M. (2016). GRAIL: A goal-discovering robotic architecture for intrinsically-motivated learning. *Cognitive and Developmental Systems*, *IEEE Transactions on, 8*(3), 214-231.

Sargent, B., Reimann, H., Kubo, M., and Fetters, L. (2015). Quantifying Learning in Young Infants: Tracking Leg Actions During a Discovery-learning Task. J*ournal of Visualized Experiments,* (100). https://doi.org/10.3791/52841

Sargent, B., Schweighofer, N., Kubo, M., and Fetters, L. (2014). Infant exploratory learning: Influence on leg joint coordination. *PloS One, 9*(3), e91500.

Schwartz, J.-L., Basirat, A., Ménard, L., and Sato, M. (2012). The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics, 25*(5), 336–354. https://doi.org/10.1016/j.jneuroling.2009.12.004

Seepanomwan, K., Santucci, V. G., and Baldassarre, G. (2017). Intrinsically Motivated Discovered Outcomes Boost User's Goals Achievement in a Humanoid Robot. In *2017 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (pp. 178-183).

Siqueland, E. R., and DeLucia, C. A. (1969). Visual Reinforcement of Nonnutritive Sucking in Human Infants. *Science, 165*, 1144–1146.

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*(1), B1–B11. https://doi.org/10.1016/j.cognition.2004.07.004

Somogyi, E., Jacquey, L., Heed, T., Hoffmann, M., Lockman, J. J., Granjon, L., … O'Regan, J. K. (2018). Which limb is it? Responses to vibrotactile stimulation in early infancy. *British Journal of Developmental Psychology, 36*(3), 384–401. https://doi.org/10.1111/bjdp.12224

Sperati, V. and Baldassarre, G.(2017). A bio-inspired model learning visual goals and attention skills through contingencies and intrinsic motivations. *IEEE Transactions on Cognitive and Developmental Systems, 10*(2), 326-344.

Sullivan, M. W., Rovee-Collier, C. K., and Tynes, D. M. (1979). A Conditioning Analysis of Infant Long-Term Memory. *Child Development, 50*(1), 152. https://doi.org/10.2307/1129051

Tanneberg, D., Peters, J., and Rueckert, E. (2019). Intrinsic motivation and mental replay enable efficient online adaptation in stochastic recurrent networks. *Neural Networks*, *109*, 67-80.

Tarabulsy, G. M., Tessier, R., and Kappas, A. (1996). Contingency Detection and the Contingent Organization of Behavior in Interactions: Implications for Socioemotional Development in Infancy. *Psychological Bulletin, 120*(1), 25–41.

Taylor, M., and Stone, P. (2009). Transfer learning for reinforcement learning domains: A survey. *The Journal of Machine Learning Research, 10*, 1633-1685.

Thelen, E. (1994). Three-month-old infants can learn task-specific patterns of interlimb coordination. *Psychological Science, 5*(5), 280–285.

Tiernan, C. W., and Angulo-Barroso, R. M. (2008). Constrained Motor-Perceptual Task in Infancy: Effects of Sensory Modality. *Journal of Motor Behavior, 40*(2), 133–142. https://doi.org/10.3200/JMBR.40.2.133-142

Tommasino, P., Caligiore, D., Mirolli, M., and Baldassarre, G. (2016). A Reinforcement Learning Architecture that Transfers Knowledge between Skills when Solving Multiple Tasks. *IEEE Transactions on Cognitive and Developmental Systems*

Verschoor, S. A., and Hommel, B. (2017). Self-By-Doing: The Role of Action for Self-Acquisition. *Social Cognition, 35*(2), 127–145. https://doi.org/10.1521/soco.2017.35.2.127

Verschoor, S. A., Spapé, M., Biro, S., & Hommel, B. (2013). From outcome prediction to action selection: developmental change in the role of action–effect bindings. *Developmental Science, 16*(6), 801-814.

Verschoor, S., Weidema, M., Biro, S., and Hommel, B. (2010). Where Do Action Goals Come from? Evidence for Spontaneous Action–Effect Binding in Infants. *Frontiers in Psychology, 1*, 201. https://doi.org/10.3389/fpsyg.2010.00201

von Hofsten, C. (2004). An action perspective on motor development. *Trends in Cognitive Sciences, 8*(6), 266–272. https://doi.org/10.1016/j.tics.2004.04.002

Wang, Q., Bolhuis, J., Rothkopf, C. A., Kolling, T., Knopf, M., and Triesch, J. (2012). Infants in Control: Rapid Anticipation of Action Outcomes in a Gaze-Contingent Paradigm. *PLoS ONE, 7*(2), e30884. https://doi.org/10.1371/journal.pone.0030884

Watanabe, H., Homae, F., and Taga, G. (2011). Developmental emergence of self-referential and inhibition mechanisms of body movements underling felicitous behaviors. *Journal of Experimental Psychology: Human Perception and Performance, 37*(4), 1157.

Watanabe, H., and Taga, G. (2006). General to specific development of movement patterns and memory for contingency between actions and events in young infants. *Infant Behavior and Development, 29*(3), 402–422. https://doi.org/10.1016/j.infbeh.2006.02.001

Watanabe, H., and Taga, G. (2009). Flexibility in infant actions during arm- and leg-based learning in a mobile paradigm. *Infant Behavior and Development, 32*(1), 79–90. https://doi.org/10.1016/j.infbeh.2008.10.003

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*(5), 297.

Willatts, P. (1999). Development of means–end behavior in young infants: Pulling a support to retrieve a distant object. *Developmental Psychology, 35*(3), 651–667. https://doi.org/10.1037/0012-1649.35.3.651

Williams, J. L., and Corbetta, D. (2016). Assessing the Impact of Movement Consequences on the Development of Early Reaching in Infancy. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00587