Intrinsic Motivations and Planning to Explain Tool-Use Development: A Study With a Simulated Robot Model

Kristsana Seepanomwan, Daniele Caligiore, Kevin J. O'Regan, and Gianluca Baldassarre

***Abstract*—Developmental psychology experiments on tool use show that infants' capacity to use a rake-like tool to retrieve a toy arises quite suddenly around 18 months. We use a developmental-robotics model to propose and test two alternative hypotheses to explain this conundrum. Both hypotheses rely on the assumptions that tool use involves goal-directed behavior processes guided by the goal of retrieving the toy, and that "understanding how to use a tool" means acquiring the capacity to assemble a sequence of actions to accomplish the goal (e.g., to "hook" and then "retrieve" the toy). The first hypothesis is that the tool-use ability emerges when the infant develops enough planning capabilities. The second hypothesis is that the ability emerges when the infant's intrinsic motivation system develops and makes playing with a couple of objects interesting enough so that the infant plays with objects similar to the tool at home and thus acquires the actions needed to retrieve the toy in the lab. These hypotheses are tested through a neural-network architecture controlling a simulated humanoid robot tested with the tool-rake task. Given the assumptions made in the model, the results show that both hypotheses can reproduce the average behavior of infants but only the intrinsic-motivation hypothesis can reproduce the sudden tool-use improvement.**

***Index Terms*—Affordances and planning, development of tool use, dynamic movement primitives (DMPs), embodied cognitive development, goal generation and intrinsic motivations (IMs), affordances and planning, neural networks, reinforcement learning, simulated iCub humanoid robot.**

Manuscript received May 24, 2019; revised October 16, 2019; accepted February 8, 2020. Date of publication May 28, 2020; date of current version March 11, 2022. This work was supported by the European Commission through the 8th Framework Programme H2020 (FP8/2014-2020), Topic H2020-EU.1.2.1.-FET Open, Project "GOAL-Robots—Goal-Based Open-Ended Autonomous Learning Robots" under Grant 713010. *(Corresponding author: Gianluca Baldassarre.)*

Kristsana Seepanomwan is with the Department of Computing, Faculty of Science, Silpakorn University, Nakhon Pathom 73000, Thailand (e-mail: seepanomwan_k@su.ac.th).

Daniele Caligiore and Gianluca Baldassarre are with the Laboratory of Computational Embodied Neuroscience, Institute of Cognitive Sciences and Technologies, National Research Council, 00185 Rome, Italy (e-mail: daniele.caligiore@istc.cnr.it; gianluca.baldassarre@istc.cnr.it).

Kevin J. O'Regan is with the Laboratoire Psychologie de la Perception, Centre National de Recherche Scientifique, Université de Paris, 75270 Paris, France.

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TCDS.2020.2986411.

Digital Object Identifier 10.1109/TCDS.2020.2986411

I. Introduction

FOR AN adult, it is relatively easy to retrieve an out-of-reach object through a tool, such as a stick or a rake. However, this task represents a challenging problem for infants. They can indeed interact with objects from an early age but they start to use objects to interact with other objects around eight months of age [1] and become able to solve tasks involving tools only during the second year of life when they have developed a good ability to combine actions [2]. The tool-use ability might develop in infants from object manipulation experience extending to multiple objects [3]–[5] and relying on trial-and-error mechanisms that sculpt the initial exploratory motor behavior into progressively more sophisticated skills [6]. Tool use requires knowledge about object–object interactions and the ability to mentally manipulate such knowledge [2], [5], [7].

Rat-Fischer *et al.* [8], [9] performed two developmental psychology tool-use experiments to investigate these issues. The results of these experiments are used to validate the model proposed here. The experiment reported in [8] used a cross section experimental design involving 60 infants aged 14, 16, 18, 20, and 22 months; instead, the experiment reported in [9], involving the same task, used a longitudinal experimental design focused on five infants. In these experiments, infants were presented with an out-of-reach toy and a rake-like tool in five conditions ranging from a condition where the toy and tool tip were physically connected to a condition where they were separated by a spatial gap. Even the youngest infants could retrieve the toy connected to the tool tip but when the tool and toy were separated by a spatial gap the first successes were observed only around the age of 18 months. For the single infant, this capacity seemed to emerge quite suddenly between two succeeding experimental sessions [9]. Currently, there is no explanation of the "puzzle" posed by the experiment of Fagard *et al.*, in particular, it is not clear why single infants develop the tool-use capacity quite abruptly around the age of 18 months [9]. In this respect, further experiments addressed this problem by varying the conditions of the test to focus on specific possible causes, such as specific motivations, social interactions, and other cognitive limitations [10]–[12], but so far the phenomenon has not yet received a satisfying explanation. The reason of this difficulty might be that infants develop cognitive processes and have experiences in real life that tend to progressively involve an increasing number of conditions: when these conditions become similar to the conditions studied in the lab, the infants' abilities manifest in an apparently sudden way and their causes are difficult to be identified. In this article, we propose two hypotheses on such processes, operationalized in a computational model.

The proposed hypotheses, and their implementation, might also be relevant for autonomous robotics. Psychological concepts related to intrinsic motivations (IMs) [13]–[16] initially inspired the development of new algorithms for autonomously learning agents [17]–[23] and are now attracting a growing attention within developmental robotics and machine learning (e.g., [24]–[34]). The models as the one proposed here, developed at the boundary between developmental psychology and autonomous robotics, might inspire new ideas, especially at the level of general principles, on algorithms and architectures useful for robotics.

We now first consider the interpretations of the target experiments on tool use [8], [9] proposed by their authors, who introduced concepts relevant to this article. We then expand those concepts and introduce new ones drawn from the psychological and computational literature: these concepts are used to design the model proposed here. We close this section by overviewing the model and highlighting its novel contributions. The remainder of this article is organized as follows. Section II overviews the architecture and functioning of the model (the details are presented in the supplementary material). Section III shows the results of the model tests. Section IV discusses the results and compares the model with other relevant models. Finally, Section V draws the conclusions.

A. Psychological Processes and Computational Mechanisms That Might Underlie Tool-Use Emergence

In [8], the authors of the target experiment interpret the development of tool use through Piaget's development stages [35]. At sensorimotor stage 4 (8–12 months), infants start to assemble sequences of actions to attain goals but they can retrieve an out-of-reach object with a tool only if the two are linked, as when the toy is on a piece of fabric [36]. At stage 5 (12–18 months), infants become able to combine two objects but they are still unable to use tools to retrieve objects if a spatial gap separates them [37]. Finally, at stage 6 (18–24 months), infants become able to mentally manipulate representations of objects and to plan, and likely due to this they become able to use tools to retrieve objects even if a spatial gap separates them [2], [37].

The understanding of how to use a tool might rely on the capacity to mentally imagine the consequences of own actions ([38]; see [39] for a model). Evidence shows that human infants form expectations on tool-use effects after they have acquired actions related to their use [40]. The ability to predict action consequences might rely on the acquisition of internal representations of the world dynamics ("forward models"): the capacity of imagining the possible consequences of actions might allow infants to plan the action sequences needed to solve tool-use tasks before performing them in the environment. Planning is indeed a powerful means to achieve complex goals, as shown by a vast computational literature ([41], [42]; see [43]–[45] for bioinspired models, where planning is based on "imagination," i.e., the internal reactivation of static/dynamic representations corresponding to percepts in the absence of their world referents).

The concept of *affordance* is also important to explain the development of tool use as it plays a key role for linking perception, action, and cognition (e.g., [46] and [47]). Affordances have been defined as what the environment offers to an animal for its needs [48] or, more specifically, the actions that are readily perceivable by an actor seeing an object [49]. Affordances are hence relational properties that depend on the agent's body and needs and the physical features of objects. Tools, and also the physical and social context, participate in the perception of object affordances [50], [51]. The concept of object affordance has been operationalized in the field of developmental robotics as the effects that a set of (usually hardwired) actions can produce on objects [52]. This concept has been also applied to tools as "intermediate objects" usable to produce effects on "primary objects" [53]. Recently, affordances have been considered as a broader concept linking objects, actions, and action outcomes to serve different functions, such as planning and action recognition [54]. In [45], we propose that affordances can be computationally specified as the expectation that an agent has that its actions will achieve their intended effects (*goals*) if performed on a given object: this definition, more closely related to the initial notion of affordance, allows one to link affordances to planning processes [41], [55].

Infants acquire knowledge related to actions (motor skills), imagination, and affordances through the spontaneous sensorimotor exploration of the environment with *play* [56]. Motivations guide these processes by driving the performance of exploratory behaviors and furnish the rewards needed to consolidate successful motor skills [16]. Within motivations, IMs play a prominent role, especially during the initial years of development [16], [24]. The psychological literature has proposed that IMs drive the performance of actions "for their own sake" [14], [57]. IMs thus differ from extrinsic motivations (EMs) that in animals are directed to obtain valuable resources, e.g., food or pain avoidance, and in robots are directed to solve externally assigned tasks [24]. Within the computational literature, IMs have been operationalized as mechanisms driving behavior to support the acquisition of new knowledge, in particular, knowledge on *novel* or *surprising* stimuli [58] in order to, respectively, learn object representations [59] and world models [18], [60]. Moreover, IMs can drive the discovery and learning of relevant action outcomes that might be later reactivated as goals, i.e., desirable world states that the agent aims to achieve; goals can then drive the acquisition of new motor skills needed to accomplish them [19], [61], [62]. Knowledge and skills acquired through IMs could later support extrinsically motivated behaviors and learning processes directed to satisfy survival and reproduction purposes, in the case of animals, or to solve tasks defined by the user, in the case of robots [24].

Concerning IMs and the empirical literature, the experiment reported in [8] shows different overt attention behaviors of the infants at different ages. These attentional behaviors might reveal the underlying IM mechanisms since attention targets the stimuli that elicit the highest interest [63]–[65]. The authors observe: "at 14 months, even if infants express interest in the toy at some times during the trial, they seem to be mainly interested in exploring the tool": we interpret this as an interest for exploring the novel tool object which overcomes the interest for the toy. "At 16 months, infants are more likely to focus their attention on the goal of retrieving the toy, but ignore or discard the tool": now, the tool might be familiar, and thus uninteresting, and its sight does not yet afford actions relevant for retrieving the toy. "From 22 months onward, infants seem to become able to spread their attention simultaneously on the toy and the tool and to make the link between the two": at this age, infants are interested in retrieving the toy and also perceive the tool affordances.

B. Overview of the Model: Two Hypotheses Incorporated by the Model and Its Novel Contributions

At the moment, there is no model that links in a coherent fashion all the elements considered this far: motor learning, imagination and forward models, goals and planning, affordances, and IMs; nor one that relates them to tool-use development. In this article, we propose such a framework and possible explanations of the tool-use puzzle in the form of a computational model. The model has an architecture formed by different components. During free exploration, some hardwired detectors of "intrinsically interesting events" allow the agent to perceive some action outcomes, e.g., an action moving an object in a certain way, as *salient* [19], [29]. In an initial free-exploration *intrinsic phase*, typical of open-ended learning models [24], [45] and corresponding to the infant's daily play before the lab tests, these detectors drive the learning of the motor skills ("actions") needed to produce those outcomes and the learning of the affordances and forward models of the skills. The possible actions the agent might discover are for example: "move the tool tip to a certain position close to the toy" or "push the toy leftward with the tool"; affordances predict which actions can reliably produce their intended effect if performed on a target object; and forward models predict the next state of the target object if a certain action is performed on it. In the following *exploitation phase*, corresponding to the lab test, if the agent desires to achieve a certain world state (goal), the previously acquired actions, affordances, and forward models allow it to use planning or trial-and-error processes to assemble action sequences to do it.

The model accounts for the target experiment by incorporating two alternative hypotheses formulated on the basis of the literature on IMs and planning discussed in Section I-A. The first hypothesis is that infants of all ages already possess the motor skills, affordances, and forward models needed to solve the tool-use task, but their planning capabilities fully develop only with time. In particular, older infants become able to think and imagine for longer times what to do and this allows them to "understand" the types of interactions between the tool and objects that they might produce with action. When these planning capabilities achieve a certain level of sophistication, the infants become able to solve the tool-use task. The second hypothesis is that during daily play, infants are driven by IMs to learn an increasing number of actions, affordances, and forward models related to objects. At a certain age, these learning processes also involve multiple objects and so allow the infants to solve the tool-use task in the lab. In particular, the needed action sequence is rapidly assembled by imagination and planning, or by trial-and-error (in the case of planning, the planning process is the same as in the first hypothesis but it is fully developed since the beginning of the simulation).

The two hypotheses are tested with the model controlling a simulated iCub humanoid robot engaged in a tool-use task similar to the one considered by Rat-Fisher *et al.* [8], [9]. In the setup, the tool is always attached to the hand of the robot, so the model does not tackle the problem of the decision to grasp the tool. This was done for simplicity as autonomous learning of grasping is still not easy for robots and as we think the core difficulties posed by tool use are still present in our scenario. In particular, the problem posed by tool use is to realize that a *sequence of actions* is needed to retrieve the object with the tool: the fact that here we consider a shorter sequence of actions of the type "hook the object with the tool; retrieve the object" rather than the longer sequence "grasp, hook, and retrieve," only changes the "quantity" of the challenge but not its quality. This is supported by the target experiments where the experimental condition, in which the tool is given to the infants, had the same results as when the tool was set on the table [8], [9]. This is also similar to what done in [66] where an iCub robot decided which tool to choose but then the tool was set in the robot's hand by the experimenter. In [67], a robot with a gripper could grasp the tool but it did this through a hardwired motor routine.

The results of the tests of our model show that both hypotheses are able to reproduce the increasing ability for tool use of the considered *groups* of infants. However, only the second hypothesis is able to reproduce the *sudden* emergence of tool-use abilities exhibited by different *single infants* at different ages around 18 months.

As further discussed in Section IV-A concerning other models, the model presented here has the following novelties with respect to understanding tool-use development: 1) the operationalization of the two hypotheses discussed above and 2) the use of the two hypotheses to account for specific tool-use empirical data. Regarding existing computational systems, the model presents these novelties: 1) the use of IMs to drive the free exploration of tools allowing the acquisition of multiple actions, affordances, and forward models, a feature shared with few other models and 2) the capacity of assembling the acquired actions on the basis of either the trial-and-error or planning processes, with the latter based on imagined action effects.

II. Methods

A. Simulated Robot and Tool-Use Scenario

Fig. 1 shows the robotic setup used to reproduce the behavior of the infants involved in the target experiments of Rat-Fisher *et al.* [8], [9]. The model controls the right arm of the simulated iCub robot [68]. iCub is a humanoid robot with

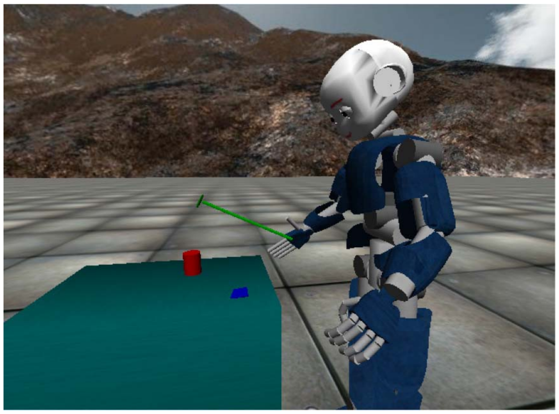


Fig. 1. Simulated tool-use scenario. The green rake-like tool is attached to the right hand of the robot. The small red cylinder represents the toy the robot has to retrieve with the tool. The blue square area on the table is the target position where the robot has to bring the toy through the tool to "successfully retrieve" it.

TABLE I
RANGE OF THE ROBOT-CONTROLLED JOINTS IN DEGREES

	Joint number						
Setting	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Min	-60	30	-20	0	-30	-60	-20
Max	0	90	40	60	30	0	20
Home	0	30	0	40	0	0	0

multiple degrees of freedom (DOF), built for studying cognitive development [68]–[70]. Each iCub arm has 16 joints: three for the shoulder (J_{0-2}), one for the elbow (J_3), three for the wrist (J_{4-6}), and nine for the hand (J_{7-15}).

The simulated setup is similar to the scenario of the target experiment. The robot is located in front of a table and a toy is placed at a location on the table outside the reaching space of the robot. The rake-like tool consists of a stick with a flat rectangle on the tip that can be used to hook the toy. For simplicity, the tool is attached to the palm of the robot hand during the whole test so that movements of the right arm directly cause movements of the tool. The task requires the robot to use the tool to bring the toy to a "target area" located close to the robot, where it could possibly grasp it with hands. To solve the task, the robot can use seven DOFs of the right arm, while all other DOFs are kept to a fixed value (e.g., the fingers are kept straight open). Each controlled joint can assume values within the range shown in Table I. This setup has been also reproduced with the real robot in a companion paper [71] that used a different model (see Section IV-A for more details).

B. Overview of the Model

This section overviews the architecture, functioning, and learning mechanisms that allow the model to solve the tool-use task (Fig. 2). This description is qualitative and allows all readers, also those with little mathematical knowledge, to have an intuition of how the model works and reproduces the target results. The detailed equations of the model are presented in the supplementary material.

In the target experiments [8], [9], the behavior of infants is observed in the lab but not in daily life. An important

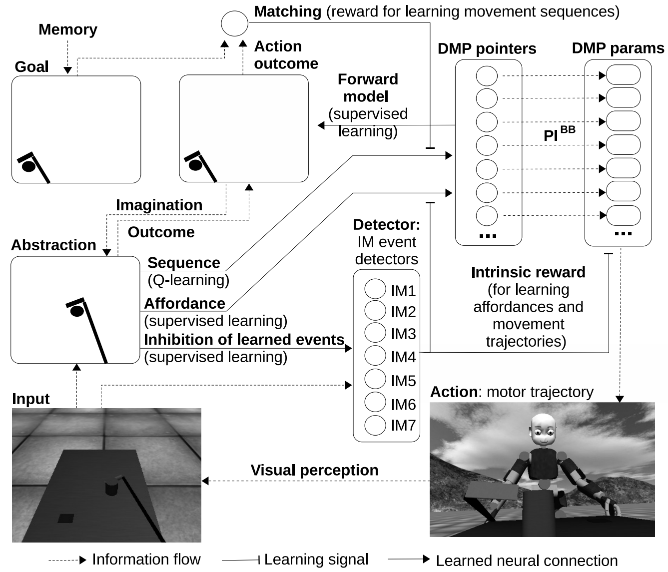


Fig. 2. Architecture of the model. Bold names: components of the model architecture. DMP: dynamic movement primitive. PI^{BB}: policy iteration black box.

idea behind the model is that infants undergo a developmental process during life that allows them to progressively acquire knowledge and skills driven by IMs [24]. When observed in the lab, infants are confronted with a new problem and solve it by relying on the knowledge and skills acquired during daily life. The model operationalizes these ideas by dividing the life of the model in two phases. In the first *IM learning phase*, reflecting daily life exploration and learning, infants learn how to interact with objects directly or through other objects (tools). In the *lab phase*, corresponding to the lab tests, motivations (possibly extrinsic) drive the infant to retrieve the toy. The two phases are interleaved as the lab tests are repeated at different ages.

In both phases, the model behavior is divided in *actions*, each consisting in a movement (arm motor trajectory) lasting 5 s (*trial*). The robot performs an action through a control model called dynamic movement primitive (DMP; [72]). A DMP has a set of parameters that leads it to perform a given motor trajectory. Here, we will refer to different "actions" or "DMPs" to mean different values of the DMP parameter set causing different motor trajectories.

In the simulations, the actions of the intrinsic phase, and the object retrieval attempts of the lab phase each involving a sequence of actions (as illustrated below), started from four possible conditions featuring different tool–object relations. The target experiments [8], [9] considered five conditions: 1) toy attached to the tool; 2) toy hooked by the rake and touching it; 3) toy hooked by the rake but not touching it; 4) toy distant from the rake; and 5) rake put in the infant's hand. As mentioned earlier, in the simulations, the robot always holds the tool to avoid the difficult problem of grasping. We thus considered two conditions.

1) *Tool-Close-To-Toy [Fig. 3(a)]:* The tool tip is hooked and is close to the toy but detached from it; this condition, although similar to condition "2" of the target experiment, actually corresponds to the easiest condition

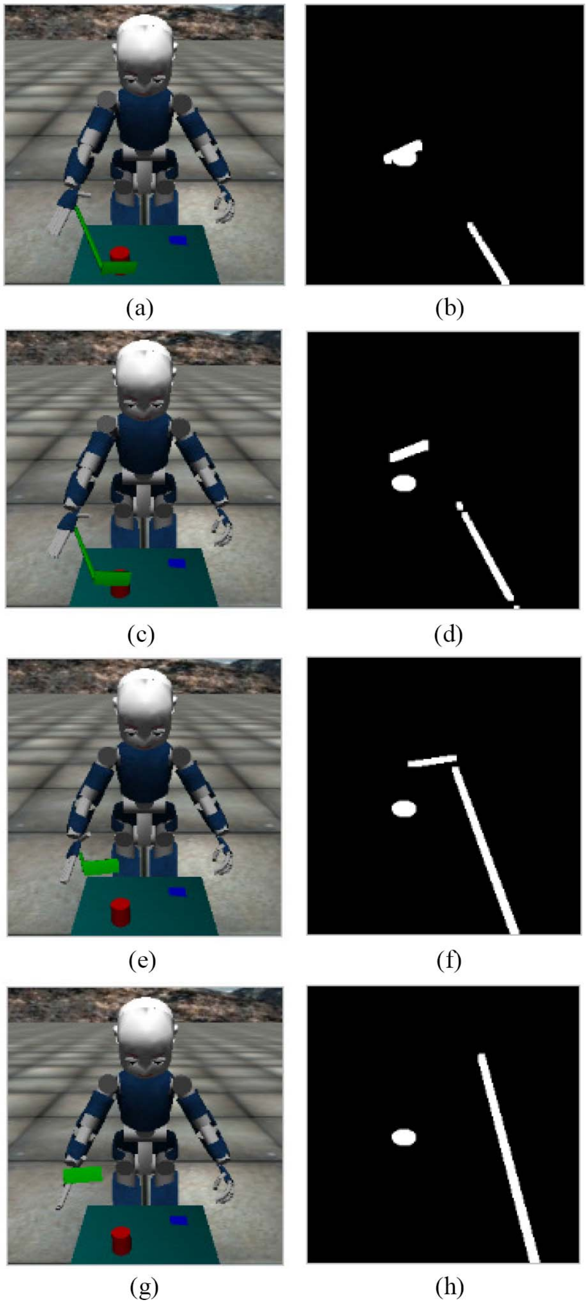


Fig. 3. Four initial conditions involving different tool–toy relations (left) and resulting filtered visual images captured by the robot camera (right). (a) and (b) No spatial gap between toy and tool tip: this condition mimics the "tool-attached-to-the-toy condition" of the target infant experiment; (c) and (d) Tool tip close and above the toy; (e) and (f) Tool far and above the toy; and (g) and (h) Tool far, above, and on the side of the toy: this condition mimics the "tool-far-from-toy condition" of the target experiment. Note that in the (g) and (h) case, the right snapshot shows that the tool tip is not visible to the robot as its color is too different from the green reference color used for the tip visual filtering (see the dark color of the tip in Fig. 4); this is an example of the challenges posed by the use of a realistic simulation.

"1" of such experiment where infants have not yet understood the tool–object relation but they are successful because the toy is physically attached to the rake; indeed, as shown below, the robot is endowed with an action of "pulling the object whatever it is" that is always successful.

2) *Tool-Far-From-Toy:* The tool tip is far from the toy; to require that the model is robust to different initial postures, for this condition, we considered three different tool–toy initial possible relations [Fig. 3(c), (e), and (g)]; this condition corresponds to the more challenging conditions "4" and "5" of the target experiment that infants cannot solve by chance and require that they "understand" that the tool is important for retrieving the toy.

During the intrinsic phase, at the beginning of each trial, the toy and robot are set to one possible initial condition randomly chosen between the four conditions described above. During the lab phase, at the beginning of each trial, the toy and robot are set to either the first condition ("tool-close-to-toy") or the fourth condition ("tool-far-from-toy"; the performance in the second and third conditions, ensuring the system robustness, is similar to the one in the fourth condition): these first and fourth conditions reproduce the first and fourth/fifth conditions of the target experiment. The robot then performs a sequence of actions that terminates with either a successful toy retrieval or a maximum number of performed actions.

Before and after performing each action, the robot perceives the tool–toy relation through its right camera. During the intrinsic phase, the robot uses this information to evaluate the effects of actions and feed the IM mechanism illustrated below and detecting interesting "events" after action execution. During the lab phase, the camera input is used to trigger actions in sequence to retrieve the toy.

The simulations mimic the intrinsic phase (daily infants' life) by allowing the robot to freely explore an object with the tool that is also used in the lab phase. This differs from the target experiment where during daily play, infants are not given the tool of the lab test. This simplifying assumption is an abstraction of the fact that infants have generalization capabilities allowing them to use different objects as tools once they have learned to use some of them. The assumption is justified by the fact that developing artificial systems with generalization capabilities similar to human ones is a difficult challenge [73] going beyond the scope of this article.

During the exploration, the robot performs actions (DMPs) with randomly chosen parameters to find movement trajectories that cause interesting effects (events) on the toy, where the "interestingness" of such effects is based on IMs. We have seen in Section I that IMs can drive the learning of knowledge and skills and that different IM mechanisms exist. Here, we are not interested in investigating the specific *causes* of IMs, but rather their possible *effects* on the emergence of the tool-use capacity. We hypothesize that infants are endowed with a general mechanism (studied in more detail in [29] and [71]) for which the outcome of an action is considered *interesting* if it causes a *change* in the environment ("event"). The rationale of this mechanism is that what infants (and robots) ultimately need to learn are the actions that cause changes in the environment as this gives them the power to change it as desired in the future. We further hypothesize that when an infant discovers the possibility to cause an event with an action, then she *actively exercises* multiple times to cause the event so as to learn to produce it reliably when desired (a process mimicking

Piaget's circular reactions [35] (see [6], [74] for models). We also hypothesize that infants progressively discover an increasing number of different interesting events, each guiding the acquisition of a different motor skill. As a consequence, at different ages, infants have an increasing number of skills usable to obtain different desired effects in the environment, either through the direct manipulation of objects or by means of other objects.

To contain complexity, the model reproduces these developmental processes in an abstract way. In particular, we endow the robot with the capacity of visually detecting gross changes/events involving the objects. The change is detected by the agent by comparing the images (the pixel-based Euclidean distance) of the environment before and after the action to determine if there was a variation (distance above a threshold) in the environment itself. For simplicity, the images are taken with the arm out of the scene so the robot's body does not contribute to the change. This mechanism for goal self-generation is partially related to IMs as: 1) different from the distinctive features of IMs [24], "change" is not related to an increase of knowledge of the agent but to something happening in the outer world and 2) however, to be relevant, the change has to be *novel* with respect to already experienced changes, a feature detected by the model. The detected changes are classified into five possible categories depending on the average position of the image pixels that change: four involving movements of the object toward the four possible cardinal directions (north, east, west, and south), and a "jerk-on-the-spot" change corresponding to the object being hit from the top. This classification is an abstract representation of the fact that infants can distinguish between different effects and learn different actions to cause them (one hypothesis explored here is based on the progressive development of this capacity). Fig. 4 shows such possible directions of movement of the object that trigger the detection of an event (called here the "IM-event") marked as interesting by the robot IM system. The five IM-events have the potential to cause the learning of five different motor skills called for reference "*N*, *E*, *W*, *S*, and *T*" actions.

In addition to these five actions, we also assume that the robot acquires two additional actions related to retrieving objects located in the hand, whatever they are. These actions are acquired under the (possibly extrinsic) reward related to having the target object in proximity to own body so as to use it (e.g., to eat it if it is food) or to further interact with it (e.g., in the case it is a toy). The first action is a "retrieval" action ("*R*") that is trained by giving the robot a reward when it succeeds to bring the object attached to the tool tip within the retrieval area, while the second is a "pull" ("*P*") action that is trained by giving a reward when the robot succeeds to bring the tool tip, without object, within such area (see Fig. 4; the idea of using a retrieval area is taken from [67]). These two basic actions are possibly the ones observed in the experiments of Rat-Fischer *et al.* when infants have a young age and retrieve objects by directly using the hand: at this age, if the tool is detached from the toy, the infants pull it to themselves to further explore it; if the tool is connected to the toy, they pull both to themselves thus seemingly solving the task. In

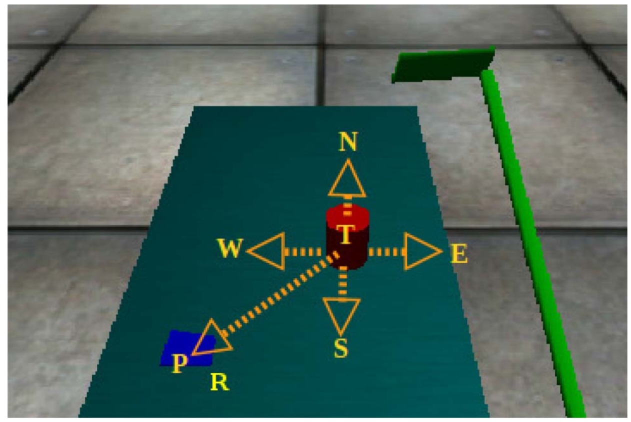


Fig. 4. Tool-use scenario with the tool (green) and object (red). Possible movements of the object ("*N*," "*E*," "*W*," "*S*," and "*T*") that are considered as interesting IM-events by the robot and can drive the acquisition of suitable motor skills to cause them. The blue area represents the "target/retrieval area," used to reward the robot when it manages to bring the object within it: this reward drives the acquisition of the "retrieval" (*R*) motor skill. It is also assumed that the robot has a "pull" (*P*) motor skill that simply brings the tool tip within the retrieval area without the object.

the model, the need to consider the two actions *R* and *P* separately, notwithstanding they require the same motor trajectory, derives from the need of planning to discriminate between different conditions and action outcomes so as to concatenate actions.

When the exploration of the toy produces an IM-event, the robot focuses on such an event and spends some time to learn (Fig. 2): 1) a suitable *DMP params* set to reliably cause the IM-event from different possible initial conditions; 2) the affordances of actions, related to the fact that if executed from a certain initial condition, the action causes the IM-event under focus (*Abstraction*→*DMP-pointers* connections); and 3) a forward model, able to predict the image (tool–object image) caused by the action: the model allows the robot to plan suitable courses of actions before performing them in the environment (*DMP-pointers*→*Action-outcome* connections).

During the lab phase, infants are challenged to retrieve the toy object with the tool. The model mimics this phase by allowing the robot to perform multiple action sequences to retrieve the toy. The model can learn and perform these action sequences either through reactive behavior and reinforcement learning or through planning. Planning is based on reinforcement learning run "in the head" through the forward model, rather than in the environment [44], [75]. The use of reinforcement learning or planning depends on which one of the two hypotheses proposed here to explain tool-use development is tested with the model. In both cases, the account of the target experiment on the tool use is obtained by testing the model multiple times at "different ages" characterized by an increased planning capability (first hypothesis) or by an increasingly complex intrinsic-motivation system (second hypothesis: this uses either planning or reinforcement learning).

We now consider more in detail how the two hypotheses account for the target experiment. According to the first hypothesis, the robot already has the motor skills needed to solve the task at all the tested ages. However, its planning capabilities needed to assemble sequences of actions to retrieve the

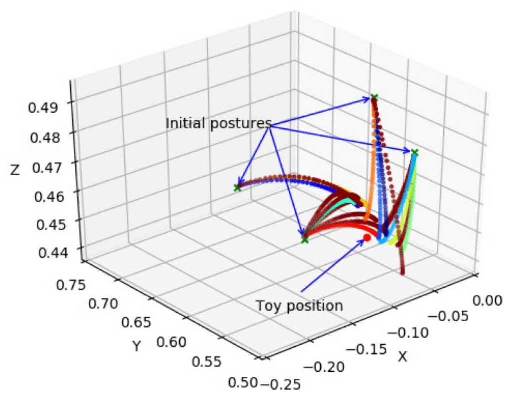


Fig. 5. Intrinsic-motivation phase: examples of movement trajectories during skill acquisition guided by IM-event detectors. The graph shows four sets of trajectories related to different initial postures.

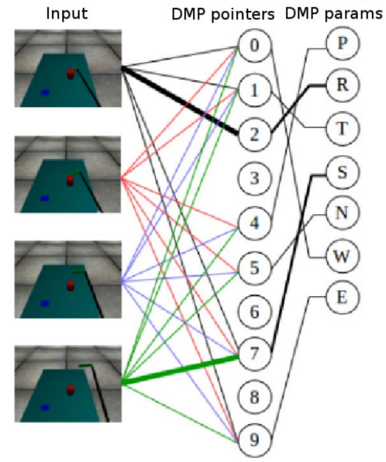


Fig. 6. Example of "links" (summarizing the connection weights different from 0) between the abstraction component (here, the input images corresponding to the four possible initial postures are reported) and the DMP-pointer component units acquired during the intrinsic learning phase (thin links) and the solution of the tool-use lab task (thick links). The example refers to a simulated infant that learned to retrieve the toy from both the "tool-close-to-toy" (through the action R) and "tool-far-from-toy" conditions (through the action sequence $<S, R>$).

object are initially limited and increase only with age. The increasing planning capacity is simulated by endowing the robot with an increasing number of planning cycles. When the available planning cycles are limited, the robot fails to "realize" (i.e., to mentally simulate) that a suitable sequence of actions might lead to retrieve the toy. For example, the robot might not think that the action of "moving the object to the left" would lead to hook the toy and create the preconditions to successfully perform the retrieval action. Only when the planning capabilities reach a certain degree of sophistication, the robot can foresee such possibilities.

According to the second hypothesis, the robot that simulates infants at different ages is endowed with an increasing number of IM-event detectors (at the same age, different simulated infants are endowed with different IM-events). This mimics the fact that infants' IM system might get progressively more sophisticated with age, possibly because of maturation processes or because of the possibility to explore an increasing number of possible effects of actions. When in the lab at different ages, each infant is thus endowed with an increasing number of different actions to assemble in sequence to solve the task. When the available actions are enough to solve the task, infants have a high chance to solve it, otherwise, they fail. In this hypothesis, the planning capacity is assumed to be fully developed from early ages. However, we will see that the same results are obtained if the skills are assembled by trial and error rather than by planning (real infants might actually mix planning and reinforcement learning).

III. Results

A. Skill Learning Based on Intrinsic Motivations and Solution of the Tool-Use Lab Task

In this section, we consider the results related to the autonomous acquisition of actions based on the IM-events during the IM phase. We, in particular, refer to the case where the robot has all the seven IM-event detectors.

Fig. 5 shows some examples of exploratory motor trajectories (of the tool tip) performed by the robot in order to discover an IM-event. The test used four different initial postures and the toy was always located in the same position.

Fig. 6 refers to two developed (late-age) infants having all the seven IM-events and maximum planning capability. In particular, the graph shows examples of the connection weights that the infants might learn, that link the units of the abstraction component to the units of the DMP-pointer component (left part of each graph). For each DMP-pointer, the graph also shows with a link the type of trajectory movement ($<P, R, T, N, E, W, S>$) produced by the corresponding parameter set. In the figure, the thin connections linking the abstraction units to the DMP-pointers units represent the affordances acquired during the intrinsic learning phase, whereas the thick connections represent the weights acquired with Q-learning during the tool-use lab test. Regarding such affordances, the model learned that the first condition (tool-close-to-toy) affords the actions W, T, R, S, and E but not N (which requires to move the tool around the toy and then push it) nor P (which requires to drag the toy to the retrieval area with the tool). Regarding the other three postures involving the tool far from the toy, the model learned that they afford all actions with the exception of R, which would require a complex trajectory (hooking the toy and bringing it to the retrieval area) to be discovered by the exploratory DMPs.

The Q-learning connections shown in the figure were learned by the model during the tool-use lab task solution, in the conditions involving the first posture (tool-close-to-toy) and the fourth posture (tool-far-from-toy). In the first condition, where the robot sees the tool close to the toy, the robot learns to select the DMP-pointer 2 related to R that leads to the retrieval of the object and to get the reward. In the second condition, where the robot sees the tool far from the toy, the robot learns to select the DMP-pointer 7 related to S as this brings the tool behind the toy: this is a condition where the robot can then perform a second R action to get the reward. This example shows how the acquisition of the ability to move objects with the tool (e.g., with action S), acquired with free

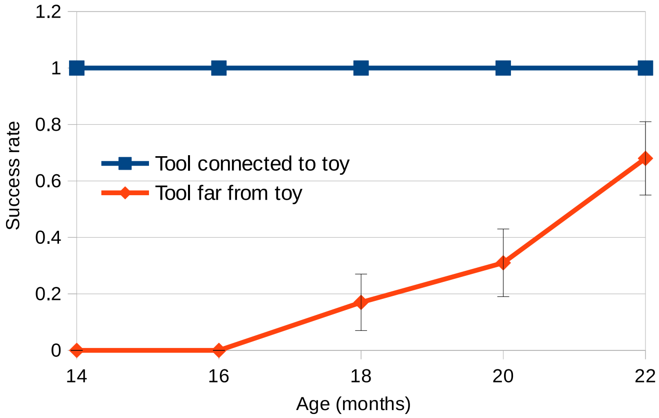


Fig. 7. Real infants: success rate measured in the target tool-use lab test, related to the "tool-connected-to-toy" condition (corresponding to the tool-close-to-toy condition in the simulations) and the "tool-far-from-toy" condition, for groups of infants with different age (cross section experimental design; mean and standard error of 12 real infants per group). Data from [8] (these data are complementary to those reported in [9] using the same experimental conditions but a longitudinal experimental design, Fig. 10).

exploration under the drive of intrinsically motivating events, later reveals useful to solve the lab task.

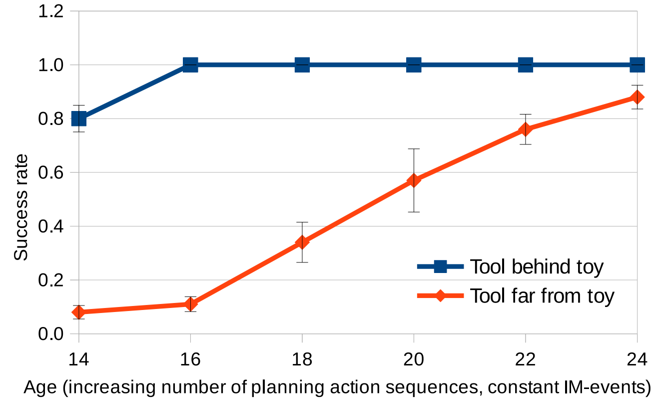
B. Solution of the Tool-Use Lab Task at Different Ages

This section presents the results related to the ability of the model to solve the tool-use lab task at different ages simulated on the basis of the two hypotheses, namely, the increase of the planning abilities or the increase of the available skills (IM-event detectors). The two hypotheses were, in particular, tested with the data from real infants shown in Fig. 7. The figure shows that when the toy is connected to the tool, then the infants at all tested ages readily retrieve it. Instead, when the tool is far from the toy, only infants with 18 months or older succeed to retrieve the toy.

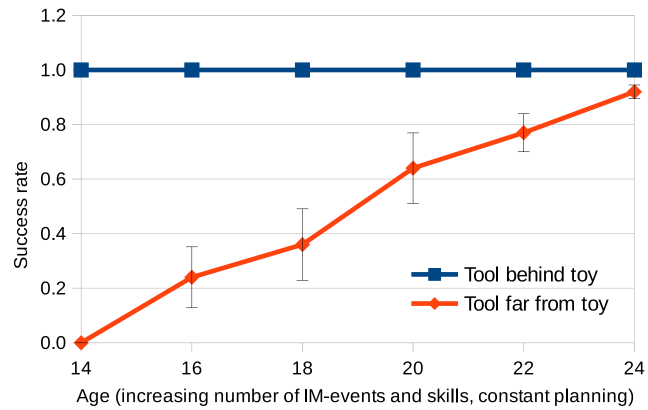
We addressed these results with the model set on the basis of the two hypotheses. For each age (14, 16, 18, 20, 22, and 24 months), we replicated the experiment ten times with different seeds of the random number generator: these replications represented different infants that learned during the IM phase and later were tested with the tool-use lab test, thus forming different connection weights at the level of motor skills (DMP parameters), affordances, Q-learning connections, forward models, and IM-event inhibition. We then tested such artificial infants for the two conditions "tool-close-to-toy" and "tool-far-from-toy." We now consider the results obtained with the two hypotheses.

1) Hypothesis 1 (Development of Tool Use When the Advancement of the Robot Age Increases on the Number of Planning Trials): We first tested the model where age was simulated on the basis of the hypothesis illustrated in Section II, i.e., by giving all IM-event detectors to infants of all ages, but the number of planning sequences increased with age.

Fig. 8(a) shows that in this condition, in the case of the toy connected to the tool, the simulated infants of all ages succeed to retrieve the toy. The reason is that the sight of the toy connected to the tool immediately affords the R action that leads to successfully retrieve the toy.



(a)



(b)

Fig. 8. Simulated infants: tool-use performance with increasing age (14–24 months). (a) Age simulated as an increasing number of planning cycles (action sequences each formed by up to five actions), in particular from 0 to 50 with increases of 10. (b) Age simulated as an increasing number of available IM-event detectors, and hence of skills acquired before the test, in particular from 3 to 7 IM-events. Each graph reports the average performance and standard error of ten simulated infants in the two conditions where the tool tip is either close to the toy or far from it. The two curves are qualitatively similar to the one related to real infants, reported in Fig. 7; in particular for the tool-far-from-toy condition, Pearson's R correlation coefficient was 0.77 and 0.73 for, respectively, the two hypotheses with respect to the real infant data.

Instead, with the tool-far-from-toy, the younger infants are not able to retrieve the toy. The reason is that even if they have the actions to bring the rake to the toy, they fail to mentally "understand" that the outcome of some actions would be useful to produce the condition from which to trigger the retrieval action. Only the older infants that are able to plan for time, rather than just triggering whatever action is afforded by the object and/or tool, manage to retrieve the toy and exhibit a success rate increasing with age (i.e., with the time spent planning).

2) Hypothesis 2 (Development of Tool Use When the Advancement of the Robot Age Increases on the Number of the Acquired Motor Skills): We then tested the model where age was simulated on the basis of the second hypothesis illustrated in Section II, i.e., by giving to all infants the possibility of planning for 50 action sequences before acting, but a limited number of IM-events increasing with age.

Fig. 8(b) shows that in the case of the toy connected to the tool, the simulated infants of all ages succeed to retrieve the toy as all of them are endowed with P and R actions (i.e., "pulling the tool" and "pulling the tool and the toy").

Instead, in the condition of the tool-far-from-toy, younger infants are not able to retrieve the toy. The reason is that even if they would have the cognitive capacity to mentally assemble the needed sequences of actions, they still lack the actions and affordances needed to control the tool and bring it to have a spatial relation with the toy that affords a successful retrieval. Only the older infants that are able to bring the tool tip close to the toy and hook it can succeed. This happens because in daily life, they had IMs that drove them to explore different interactions between objects and so learned the actions needed to prepare a successful retrieval, and affordances to "see" such possibility.

What would happen in the case of the second hypothesis if the infants had limited planning capabilities? Would older infants still be able to solve the tool-use lab task in the hard condition? For this to be possible, the infants would still need to form action sequences. A possibility to do this would be to learn such sequences by trial-and-error learning rather than by planning, either before coming to the lab or during the lab sessions (or both). To test the computational viability of this possibility, we prevented the model from planning but allowed it to learn to solve the lab task by trial-and-error for an increasing number of learning sequences, i.e., in a way similar to what done for planning but with the difference that the model performed the actions in the environment rather than "in the head" (i.e., through the forward models).

The results are shown in Fig. 9. These results are qualitatively similar to those obtained with the planning hypothesis [Fig. 8(b)], showing that the capacity of planning is not strictly needed if one assumes that the infants have enough time to either learn to solve the tool-use task during the lab sessions or before then (in the latter case, one should also assume infants can generalize such knowledge to the lab task).

Incidentally, notice how, in the tool-far-from-toy condition and the second hypothesis, the performance of the trial-and-error learning model (Fig. 9) is slightly worse than the performance of the planning model [Fig. 8(b)]. The reason is that with planning, there is no noise that can cause the mental images corresponding to the action outcomes to differ from the image of the desired goal, thus the matching process is more accurate and the formation of the Q-learning connections is faster.

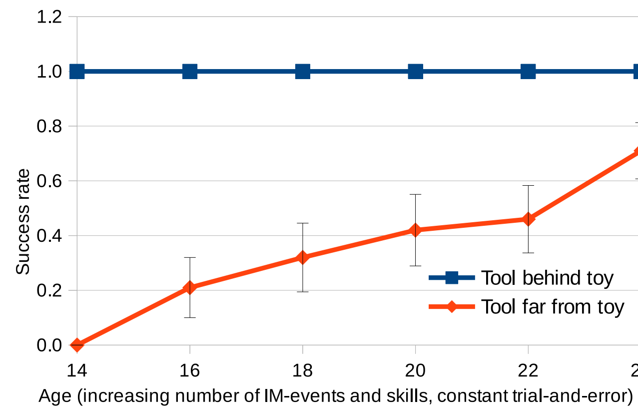


Fig. 9. Simulated infants: success rate of infants with different ages simulated as an increasing number of IM-event detectors (mean and standard error of ten simulated infants). In this simulation, the model learned the action sequence to solve the lab task through a trial-and-error (reinforcement learning) process rather than with planning.

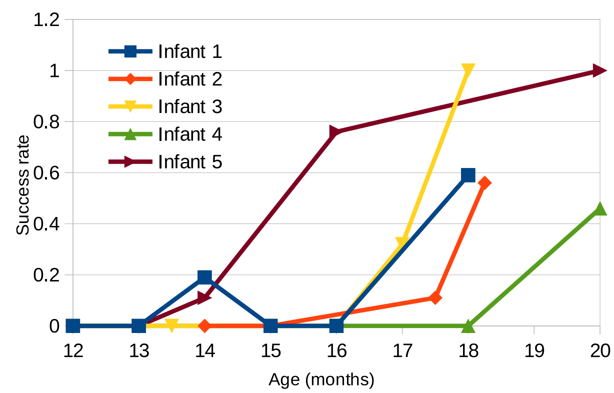
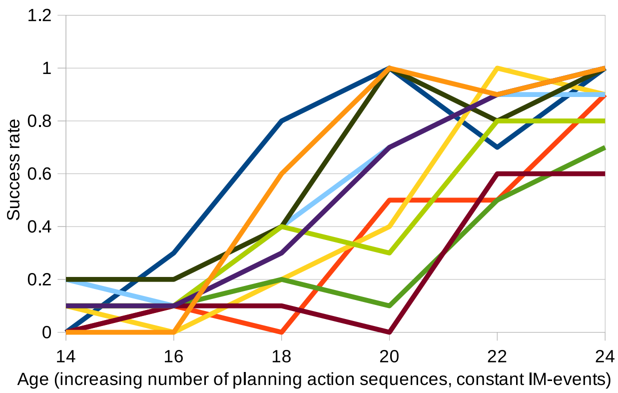


Fig. 10. Real infants, longitudinal data on tool-use performance: individual performance at different ages of five infants. The timing of the tests for the different infants had some variability. Data obtained from [9] (these data are complementary to those reported in [8] using the same experimental conditions but a cross section experimental design, see Fig. 7).

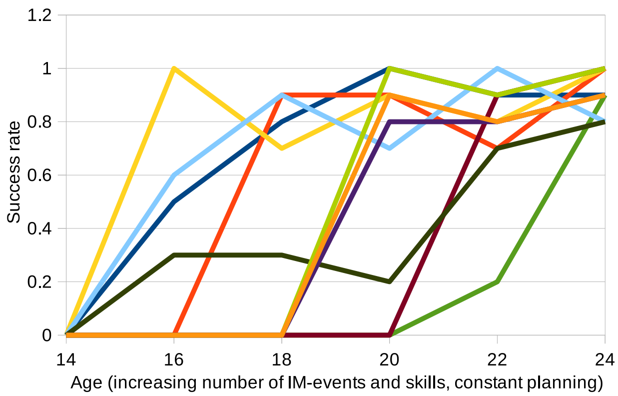
C. Development of Tool Use by Single Simulated Infants Is Better Explained by the Second Hypothesis

The previous section shows that both the planning and the IM-based hypotheses correctly predict the steady increase of infants' performance in the tool-use lab task. However, can the two hypotheses similarly predict the developmental pattern exhibited by *single* infants? In this respect, Fig. 10 shows the capacity exhibited by different infants in the *longitudinal* experiment [9] where the tool-use ability of the same infant was monitored at different ages to track his/her developmental trajectory. Recall that this experiment reports data collected in the same experimental conditions used in the *cross section* experiment considered this far ([8]; Fig. 7). Fig. 10 clearly shows that: 1) each infant seems to "suddenly" acquire the ability at a certain age, meaning that such ability passes from a close-to-zero level to a notable percent of success and 2) different infants acquire the tool-use ability at different ages, thus showing that they have a different developmental history.

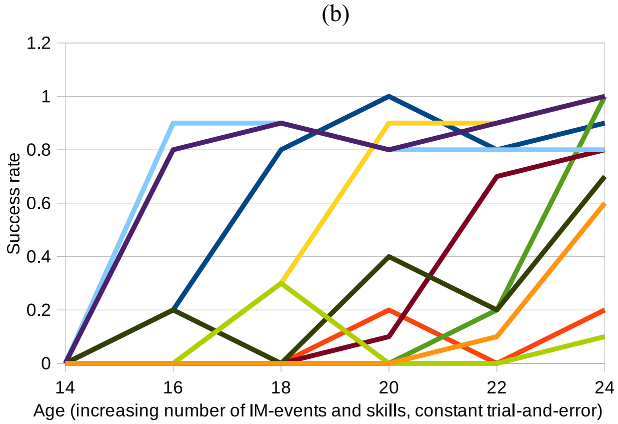
Fig. 11 shows the performance of individual simulated infants at different ages, mimicking the longitudinal experiment. The test refers to three conditions: 1) planning hypothesis; 2) the IM-event hypothesis with planning; and 3) the IM-event hypothesis with trial-and-error learning. The results show that in the case of the planning hypothesis [Fig. 11(a)] each single infant has a tool-use ability that tends to progressively increase with age. This is due to the increasing number of planning cycles that produce an increasing chance to succeed. Instead, in the case of the IM-event hypothesis [Fig. 11(b)], the single infants exhibit a sudden increase of the tool-use ability, and this happens at different ages. This developmental pattern qualitatively mirrors the pattern of real infants (Fig. 10). This also happens in case the actions are assembled by trial-and-error rather than by planning [Fig. 11(c)]. The explanation is that these simulated



(a)



(b)



(c)

Fig. 11. Simulated infants: success rate in the lab longitudinal test of three groups (different graphs) of ten infants each (curves in each graph). (a) Age simulated as an increasing number of planning action sequences. (b) Age simulated as an increasing number of IM-events, and lab test solved with planning. (c) Age simulated as an increasing number of IM-events, and lab test solved with trial-and-error learning.

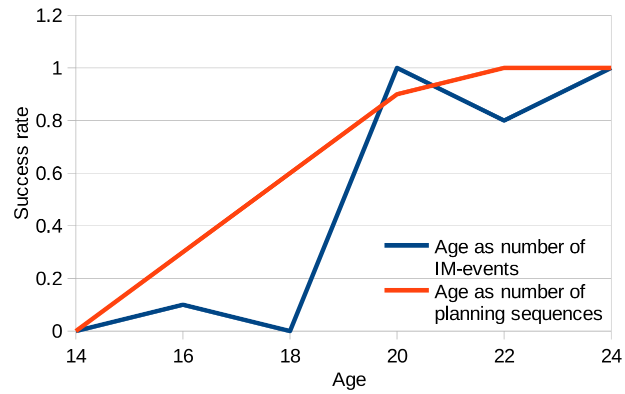


Fig. 12. Simulated infants: success rate in the tool-use lab test of two simulated infants, one with an age simulated as an increasing number of planning cycles before action performance, and the second as an increasing number of IM-events and hence actions acquired in daily life before the lab test.

infants face the lab test endowed with a repertoire of affordances and skills acquired in daily life, and if this repertoire involves the actions needed to solve the task they immediately anticipate this possibility based on object affordances and so readily assemble the needed action sequence by planning or by trial and error.

D. Analysis of the Model Functioning

In this section, we analyze some aspects of the internal functioning of the model. To this purpose, we consider two simulated infants tested longitudinally at different ages, reproduced, respectively, with the planning-based hypothesis (and fully developed IMs) and the IM-event hypothesis (with fully developed planning) in the tool-far-from-toy condition. Fig. 12 shows the performance of the two infants at different ages. We now focus on these two representative simulated infants but the main result, i.e., the progressive development of the tool-use ability with the planning hypothesis and its sudden development with the IM hypothesis, is very robust, as shown in Fig. 11 related to ten simulated different infants for each of the two conditions.

Table II in the supplementary material considers the planning-hypothesis infant and shows data on one action sequence per each age that the infant produced to solve the lab test in the tool-behind-toy condition (for each age, one table row refers to one action cycle of the sequence). Recall that each sequence terminates with either a successful retrieval of the toy (reward = 1) or when a maximum of five actions are produced. For each age/action sequence, Table II reports: 1) the actions that the simulated infant acquired in daily life out of the seven possible ones, and with which it is endowed when it comes to the lab test; 2) the probabilities of selection of the actions produced by the *softmax* function; 3) the action actually selected in the action-selection cycle; and 4) the reward after the action performance. The action selection probabilities of the first sequence (age 14) are caused only by the affordances acquired during the IM phase whereas those of the following ages are caused by both the affordances and the Q-learning connection weights, the latter acquired with planning before the action sequence shown in the table is performed.

Table II shows that the initial image of the tool-behind-toy condition affords only the actions R, W, and S. This implies that during planning, the infant considers only these actions: the restriction of the choice to three actions implies that during the five actions of the sequence, R is selected with a high chance even at 14 months and leads to successfully accomplish the task (reward = 1), here at the third action after the action sequence $<W, S, R>$. At later ages, planning strengthens the useful Q-learning connections of R and so the chance of selecting it in the suitable condition moves from 0.24 to 0.96. This leads to a successful retrieval at the first action-selection most of the times.

Table III in the supplementary material shows similar data for the same infant but for the tool-far-from-toy condition. Now, the initial image affords all actions with the exception of R (see the first line of the table). At 14 months of age, the simulated infant performs two P actions without changing the toy condition and so the affordances remain the same (recall that after P or R, the robot is reset to the initial posture, similar to the target experiment, but the sequence is not interrupted). Then, it performs actions W and S that change the image and hence the affordances: their outcome is that the tool hooks the toy and so R, W, and S actions become afforded while the other ones are no more afforded. The sequence terminates unsuccessfully (reward = 0). In the following ages, with more planning cycles, the infant "understands" that in the initial condition the action S has to be favored as it creates the achievement of a suitable precondition for R (the probability of S passes from 0.19 to 0.98). At the same time, when the toy is hooked (e.g., after the performance of S), the action R progressively acquires a higher chance of being selected (from 0.01 to 1). The result is that at older ages, the robot tends to select an efficient $<S, R>$ action sequence that successfully accomplishes the task (reward = 1).

Table IV in the supplementary material reports analogous data for the simulated infant having an age involving an increasing number of skills, in the tool-behind-toy condition. At the age of 14 months, where only P, R, and T actions were acquired by the infant in the IM phase before the lab test (see table), the scene affords only the action R and so the infant successfully accomplishes the task. With the increasing age, the infant acquires new actions in daily life but both the affordances and the Q-connections resulting from planning allow the infant to focus on the R action and efficiently solve the lab task.

Table V in the supplementary material reports data of the same infant in the tool-far-from-toy condition. In this case, at early ages (14, 16), the simulated infant fails to solve the task as it does not have any action that can allow it to hook the toy with the tool. Only at the age of 18 months, a potentially useful S action becomes available for that purpose and planning favors it (its probability is 0.55). However, the infant still fails to retrieve the toy even when R follows S (as shown in Fig. 12 for ten tests): this happens because the learned motor trajectory of S is not good for hooking the toy, the precondition needed to successfully perform R. In the following months, the infant becomes increasingly able to solve the task by using the now-available action W to hook the toy. Note that in the case reported in Table III in the supplementary material, S was instead preferred to W: that case and the one considered here show an instance of developmental differentiation due to the specific motor trajectories acquired with free exploration.

IV. Discussion

The results show that both hypotheses are able to reproduce the target data. In particular, in both cases, infants of all ages are able to solve the tool-use task when the tool is connected to the toy as this requires only one retrieval action. Instead, when the tool is far from the toy and the solution of the task requires performing an action sequence—bringing the tool behind the toy to hook it and then retrieving it—only older infants can solve the task. Based on the first hypothesis, this happens because younger infants possess all needed motor skills, related affordances, and forward models, but have poor planning capabilities. This leads them to act before they realize by planning that a successful accomplishment of the task requires a two-action sequence. Based on the second hypothesis, younger infants do not possess the skills/affordances/forward models needed to hook the object with the tool. This lack of knowledge is due to the low level of development of their IM system that during daily life activities does not lead them to acquire the necessary abilities before facing the lab task. Only at a later age, the two groups of infants can solve the task, respectively, because they have a sufficient planning capacity or because they have sufficient motivations driving the acquisition of the needed skills/affordances/forward models in daily life.

Notwithstanding the similarity of the predictions of the two hypotheses at the group level, the individual differences in the tool-use behavior exhibited by single infants when analyzed in longitudinal studies (the same infant tested at different ages) can be reproduced only by the second hypothesis. In this respect, a key feature of the target data is that single infants exhibit a lack of ability to solve the tool-use task up to a certain age and then they suddenly exhibit a high capacity to solve it at different ages around 18 months. Only the model where age determines the development of the IM system, but not the one where it determines the planning ability, is able to reproduce this sudden emergence of the tool-use ability. Indeed, the simulated infants whose planning ability strengthens progressively with age also show a progressive increase in the probability of success in the lab task. Instead, the simulated infants that progressively develop the IM system start to suddenly have full success in the lab task as soon as they develop the needed skills in daily life outside the lab due to the development of the related IM interest.

A. Comparision With Other Models

Previous robotic models faced the problem of tool use by relying on the concept of affordance. In [67], during an exploratory phase, an arm-gripper robot learned affordances of different tools by performing a number of hardwired actions (e.g., extend arm and move arm left) and by observing the resulting effects (object displacements). The authors showed how the autonomously learned affordances can be later used to solve a tool-based retrieval task by sequencing the actions (e.g., "grasp tool, move arm, and move object") through a domain-specific procedure using the action expected outcomes. In comparison to our model, the actions were not autonomously learned and the planning procedure was not general. Moreover, despite the fact that the model was built starting from a psychological background on tool use in animals and humans, it was not used to address specific developmental issues.

Another model [66] faced tool-use tasks with the iCub robot. The tool-use scenario involved object retrieval tasks in which a toy was placed far from the robot at different locations on a table and the robot had to select a suitable tool (a rake, a hoe, or a stick) to retrieve it. The model was based on an architecture encompassing different modules each capable of implementing a function needed to solve the task: object recognition, action performance, affordance learning, inverse kinematic models, etc. The robot chose the tool depending on the position of the object that afforded a certain action. The model solved the tool–task based on a specific procedure evaluating the tool–object distance rather than with a general planning procedure as here. Moreover, the model was inspired by the development of tool use in infants but it did not address and explain specific developmental data.

A first model to use "salient states" marked by environment changes to form skills, called "options" as the system was developed within the reinforcement learning framework, was proposed in [19]. The system also used IMs based on the prediction of the termination of options to focus learning on skills needing further training. The system, tested in a grid-world scenario with abstract objects, showed the importance of IMs for guiding the learning of reusable skills. In [29], we proposed a general architecture for goal self-generation and goal-based skill learning called "GRAIL—Goal-Discovering Robotic Architecture for Intrinsically Motivated Learning." The architecture controlled a simulated iCub robot and was able to self-generate goals on the basis of *novel changes* caused in the environment, to focus on them based on competence-based IMs (as here), and so to learn the skills to accomplish them. In the following work [71], we focused on the use of goals and skills previously acquired with goal self-generation as in GRAIL for the solution of extrinsic user-defined goals. The system was, in particular, tested with a real iCub that had to move a ball to multiple desired positions (user-defined goals) by using a tool attached to the robot hand (as here). The robot solved these tasks by recalling, and possibly improving, the skills whose goal was most similar to the user-defined goal. Another work [76] presents a robotic system that learns to control, via a joystick, another robot arm that can act on a ball in turn possibly affecting some lights. The system self-generates goals based on novel action effects and uses competence-based IMs to focus on the different objects/experiences. This allows it to exhibit an autonomous "curriculum learning" starting from the acquisition of the robot control and arriving at the acquisition of the control of the lights. All these systems inspired several aspects of the model presented here but they were not used to study in detail the emergence of tool use as measured in developmental psychology experiments.

The model presented here also builds on previous computational models, tested with the simulated iCub robot, that we proposed to study mental rotation and decision-making abilities in humans [39], [77]. The model presented here has important innovations with respect to such models: 1) the introduction of IM mechanisms that allow the robot to acquire skills by an autonomous exploration of the environment; 2) the use of IM mechanisms to simulate developmental processes leading to the acquisition of multiple motor skills; 3) the introduction of mechanisms able to implement planning based on the actions acquired autonomously by the robot; and 4) the use of these new elements of the model to tackle specific developmental experiments on tool use. The model presented here also draws some features from another model of ours [45]. Designed within an open-ended learning framework, this model used IMs to learn affordances and forward models. The model implemented image-based planning but only for one action ahead. The model was used to study the learning of stochastic affordances based on IMs and their relation with action preconditions. With respect to the model presented here, the model was tested only in a very simplified 2-D environment, involving colored-shaped objects and abstract physical interactions, and used hardwired motor skills. Finally, the model did not tackle tool use and developmental issues related to it. Such a model built on an earlier model that used multiaction planning based on *image trajectories* and *goals* [43], [44]. This model was created after the idea of *Dyna architectures* [78], initially working with grid-world states, to train a reinforcement learning model through internal models rather than in the environment.

V. Conclusion

The model proposed here was built within a developmental robotics framework. Its design and test were based on the experiments on tool use with infants from the developmental psychology experiments of Rat-Fischer *et al.* [8], [9]. These experiments showed that infants developed a capacity to solve tool-use tasks around the age of 18 months. The model proposed two hypotheses to explain these results. The first was related to the development of increasingly powerful planning abilities for composing the action sequences needed to solve the tool-use lab task. The second hypothesis claimed that the lab data were a manifestation of the effects of learning processes happening during the infants' life outside the lab. With age, these processes lead infants to acquire an increasing number of motor skills, and related affordances/forward models, until they also acquire those needed to solve the lab task. Both hypotheses explained the emergence of the tool-use capacity in the whole population of infants, but only the second hypothesis was able to reproduce the *sudden* emergence of the tool-use ability observed in single infants in longitudinal studies. The reason is that the increase of planning abilities leads to a *progressive* increase of the probability to solve the lab task. Instead, the development of IMs leads to a sudden interest in the acquisition, in daily life, of the skills needed to solve the task in the lab.

Future empirical experiments might test the latter hypothesis on IMs. For example, two groups of infants might be allowed to play with two different sets of toys, one affording actions and action sequences in a "tool-like" fashion, and one not doing that. Afterward, the experiment could present different tool-use tasks as those studied here, and monitor the infants' performance. This article would allow the investigation of the link existing between the free explorations during the initial play sessions and the solution of the tool-use tasks [12], [79], [80].

Although the model achieved encouraging results, we are aware it has various limitations that could be tackled in future work. A general limitation, shared with all system-level integrative models, is that it has many mechanisms and parameters that allow it to fit different data. Notwithstanding this problem, we think these types of models have a useful theoretical role in supporting clear formal thinking on psychological problems, and a heuristic value to suggest new experiments. Moreover, as expanded in the supplementary material, various strategies might be followed to face the problem of the many mechanisms/parameters: 1) addressing many experiments rather than one; 2) introducing constraints at multiple brain/behavioral levels; and 3) requiring the models to scale up to more complex scenarios and behaviors.

Also from the autonomous robotics perspective, the model has limitations to be faced in future work. In particular, the model has the components needed to face the two key phases of open-ended learning [23], [45]: 1) the intrinsic phase where the robot autonomously acquires affordances/motor skills/forward models and 2) the extrinsic phase where the robot uses the previously acquired knowledge to solve tasks useful for the user. Such components were kept simple to focus on the target developmental issues, so they should be further developed to scale up to more complex scenarios useful for autonomous robotics, for example, in the ways we now discuss.

The IM-event detectors used here could only detect interesting events corresponding to moving the object in the four cardinal directions. In future work, more general IM-event detectors should be investigated [29], [30], [81]. Moreover, here we considered the IM-event detectors as the markers of interesting action outcomes, but such outcomes were not stored by the system, in particular as possible future goals: the only goal considered here was the whole tool-use task solution stored in a hardwired fashion. Future work should hence introduce general mechanisms to store goals so as to allow the robot to solve multiple extrinsic tasks [65], [71]. In addition, for each IM-event, we considered here a limited number of initial postures: future work should investigate if and how DMPs could cover larger portions of the continuous posture space (see [82], [83]).

Planning was here implemented for combining actions drawn from a small action repertoire to build action sequences formed by at most two actions. Future work should scale up the proposed planning mechanisms to more complex scenarios [84], [85]. Moreover, the used forward models work through detailed primitive perceptual inputs (camera images). Future work should develop prediction and planning processes working at different levels of abstraction (e.g., [81], [85], and [86]).

Overall, the model proposes new possible explanations of the puzzle of the sudden emergence of tool-use abilities in infants, worth further investigation. Moreover, the robotic implementation of the model proposes various ideas and challenges relevant to open-ended learning autonomous robots.

References

[1] M. E. McCarty, R. K. Clifton, and R. R. Collard, "The beginnings of tool use by infants and toddlers," *Infancy*, vol. 2, no. 2, pp. 233–256, 2001.

[2] R. Keen, "The development of problem solving in young children: A critical cognitive skill," *Annu. Rev. Psychol.*, vol. 62, no. 1, pp. 1–21, 2011.

[3] B. A. Kahrs, W. P. Jung, and J. J. Lockman, "Motor origins of tool use," *Child Develop.*, vol. 84, no. 3, pp. 810–816, 2013.

[4] B. A. Kahrs and J. J. Lockman, "Building tool use from object manipulation: A perception-action perspective," *Ecol. Psychol.*, vol. 26, nos. 1–2, pp. 88–97, 2014.

[5] J. J. Lockman, "A perception-action perspective on tool use development," *Child Develop.*, vol. 71, no. 1, pp. 137–144, 2000.

[6] V. C. Meola *et al.*, "Interplay of rhythmic and discrete manipulation movements during development: A policy-search reinforcement-learning robot model," *IEEE Trans. Cogn. Develop. Syst.*, vol. 8, no. 3, pp. 152–170, Sep. 2016.

[7] F. Guerin, N. Kruger, and D. Kraft, "A survey of the ontogeny of tool use: From sensorimotor experience to planning," *IEEE Trans. Auton. Mental Develop.*, vol. 5, no. 1, pp. 18–45, Mar. 2012.

[8] L. Rat-Fischer, J. K. O'Regan, and J. Fagard, "The emergence of tool use during the second year of life," *J. Exp. Child Psychol.*, vol. 113, no. 3, pp. 440–446, 2012.

[9] J. Fagard, L. Rat-Fischer, and J. K. O'Regan, "The emergence of use of a rake-like tool: A longitudinal study in human infants," *Front Psychol*, vol. 5, p. 491, May 2014, doi: 10.3389/fpsyg.2014.00491.

[10] R. Esseily, L. Rat-Fischer, E. Somogyi, K. J. O'Regan, and J. Fagard, "Humour production may enhance observational learning of a new tool-use action in 18-month-old infants," *Cogn. Emotion*, vol. 30, no. 4, pp. 817–825, 2016.

[11] R. Esseily, L. Rat-Fischer, K. O'Regan, and J. Fagard, "Understanding the experimenter's intention improves 16-month-olds? Observational learning of the use of a novel tool," *Cogn. Develop.*, vol. 28, no. 1, pp. 1–9, 2013.

[12] E. Somogyi *et al.*, "The roles of observation and manipulation in learning to use a tool," *Cogn. Develop.*, vol. 35, pp. 186–200, Jul.–Sep. 2015.

[13] R. W. White, "Motivation reconsidered: The concept of competence," *Psychol Rev*, vol. 66, pp. 297–333, Sep. 1959.

[14] D. E. Berlyne, *Conflict, Arousal, and Curiosity*. New York, NY, USA: McGraw-Hill, 1960.

[15] E. L. Deci and R. M. Ryan, *Intrinsic Motivation*. New York, NY, USA: Plenum Press, 1975.

[16] C. von Hofsten, "An action perspective on motor development," *Trends Cogn. Sci.*, vol. 8, no. 6, pp. 266–272, 2004.

[17] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: A survey," *Connection Sci.*, vol. 15, no. 4, pp. 151–190, Dec. 2003. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/09540090310001655110

[18] J. Schmidhuber, "Curious model-building control systems," in *Proc. IEEE Int. Joint Conf. Artif. Neural Netw.*, vol. 2, 1991, pp. 1458–1463.

[19] A. G. Barto, S. Singh, and N. Chentanez, "Intrinsically motivated learning of hierarchical collections of skills," in *Proc. Int. Conf. Develop. Learn. (ICDL)*, 2004, pp. 112–119.

[20] A. G. Barto, "Intrinsic motivation and reinforcement learning," in *Intrinsically Motivated Learning in Natural and Artificial Systems*. Berlin, Germany: Springer, 2013, pp. 17–47.

[21] M. Schembri, M. Mirolli, and G. Baldassarre, "Evolving childhood's length and learning parameters in an intrinsically motivated reinforcement learning robot," in *Proc. 7th Int. Conf. Epigen. Robot. (EpiRob) Model. Cogn. Develop. Robot. Syst.*, vol. 134, Nov. 2007, pp. 141–148.

[22] P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, "Intrinsic motivation systems for autonomous mental development," *IEEE Trans. Evol. Comput.*, vol. 11, no. 2, pp. 265–286, Apr. 2007.

[23] G. Baldassarre and M. Mirolli, Eds., *Intrinsically Motivated Learning in Natural and Artificial Systems*. Berlin, Germany: Springer, 2013.

[24] G. Baldassarre, "What are intrinsic motivations? A biological perspective," in *Proc. IEEE Int. Conf. Develop. Learn. Epigen. Robot. (ICDL-EpiRob)*, Aug. 2011, pp. 1–8.

[25] T. Jung, D. Polani, and P. Stone, "Empowerment for continuous agent environment systems," *Adapt. Behav.*, vol. 19, no. 1, pp. 16–39, 2011.

[26] M. Mirolli and G. Baldassarre, "Functions and mechanisms of intrinsic motivations: The knowledge versus competence distinction," in *Intrinsically Motivated Learning in Natural and Artificial Systems*, G. Baldassarre and M. Mirolli, Eds. Berlin, Germany: Springer-Verlag, 2013, pp. 49–72.

[27] A. Cangelosi and M. Schlesinger, *Developmental Robotics: From Babies to Robots*. Boston, MA, USA: MIT Press, 2015.

[28] R. Der and G. Martius, "Novel plasticity rule can explain the development of sensorimotor intelligence," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 45, pp. 6224–6232, 2015.

[29] V. G. Santucci, G. Baldassarre, and M. Mirolli, "GRAIL: A goal-discovering robotic architecture for intrinsically-motivated learning," *IEEE Trans. Cogn. Develop. Syst.*, vol. 8, no. 3, pp. 214–231, Sep. 2016.

[30] T. D. Kulkarni, K. R. Narasimhan, A. Saeedi, and J. B. Tenenbaum, "Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation," 2016. [Online]. Available: arxiv.1604.06057.

[31] M. G. Bellemare, S. Srinivasan, G. Ostrovski, T. Schaul, D. Saxton, and R. Munos, "Unifying count-based exploration and intrinsic motivation," 2016. [Online]. Available: arXiv.1606.01868.

[32] D. Pathak, P. Agrawal, A. A. Efros, and T. Darrell, "Curiosity-driven exploration by self-supervised prediction," 2017. [Online]. Available: arXiv:1705.05363.

[33] P. Dhakan, K. Merrick, I. Rañó, and N. Siddique, "Intrinsic rewards for maintenance, approach, avoidance, and achievement goal types," *Front. Neurorobot.*, vol. 12, p. 63, Oct. 2018.

[34] S. Doncieux *et al.*, "Open-ended learning: A conceptual framework based on representational redescription," vol. 12, p. 59, Sep. 2018.

[35] J. Piaget, *The Origins of Intelligence in Children*. London, U.K.: Routledge, 1952.

[36] E. Bates, V. Carlson-Luden, and I. Bretherton, "Perceptual aspects of tool using in infancy," *Infant Behav. Develop.*, vol. 3, pp. 127–140, Jan. 1980.

[37] A. L. Brown, "Domain-specific principles affect learning and transfer in children," *Cogn. Sci.*, vol. 14, no. 1, pp. 107–133, 1990.

[38] W. Möhring and A. Frick, "Touching up mental rotation: Effects of manual experience on 6 month old infants' mental object rotation," *Child Develop.*, vol. 84, no. 4, pp. 1554–1565, 2013.

[39] K. Seepanomwan, D. Caligiore, G. Baldassarre, and A. Cangelosi, "Modelling mental rotation in cognitive robots," *Adapt. Behav.*, vol. 21, no. 4, pp. 299–312, 2013.

[40] M. Schlesinger and J. Langer, "Infants' developing expectations of possible and impossible tool—Use events between ages 8 and 12 months," *Develop. Sci.*, vol. 2, no. 2, pp. 195–205, 1999.

[41] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Harlow, U.K.: Pearson, 2016.

[42] M. M. Botvinick, Y. Niv, and A. Barto, "Hierarchically organized behavior and its neural foundations: A reinforcement-learning perspective," *Cognition*, vol. 113, no. 3, pp. 262–280, 2008.

[43] G. Baldassarre, "Planning with neural networks and reinforcement learning," Ph.D. dissertation, Comput. Sci. Dept., Univ. Essex, Colchester, U.K., 2002.

[44] G. Baldassarre, "Forward and bidirectional planning based on reinforcement learning and neural networks in a simulated robot," in *Anticipatory Behaviour in Adaptive Learning Systems*, M. Butz, O. Sigaud, and P. Gärard, Eds. Berlin, Germany: Springer Verlag, 2003, pp. 179–200.

[45] G. Baldassarre, W. Lord, G. Granato, and V. G. Santucci, "Affordances and planning in open-ended learning systems endowed with attention and manipulation capabilities," *Front. Neurorobot.*, to be published.

[46] M. Tucker and R. Ellis, "The potentiation of grasp types during visual object categorization," *Vis. Cogn.*, vol. 8, no. 6, pp. 769–800, 2001.

[47] D. Caligiore, A. M. Borghi, D. Parisi, and G. Baldassarre, "TRoPICALS: A computational embodied neuroscience model of compatibility effects," *Psychol. Rev.*, vol. 117, pp. 1188–1228, Oct. 2010.

[48] J. J. Gibson, *The Ecological Approach to Visual Perception*. Boston, MA, USA: Houghton Mifflin, 1979.

[49] D. Norman, *The Psychology of Everyday Things*. New York, NY, USA: Basic Books, 1988.

[50] A. M. Borghi, C. Scorolli, D. Caligiore, G. Baldassarre, and L. Tummolini, "The embodied mind extended: Using words as social tools," *Front. Psychol.*, vol. 4, p. 214, May 2013.

[51] D. Caligiore, A. M. Borghi, D. Parisi, R. Ellis, A. Cangelosi, and G. Baldassarre, "How affordances associated with a distractor object affect compatibility effects: A study with the computational model TRoPICALS," *Psychol. Res.*, vol. 77, pp. 7–19, Feb. 2013.

[52] P. Fitzpatrick, G. Metta, L. Natale, S. Rao, and G. Sandini, "Learning about objects through action-initial steps towards artificial cognition," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 3, 2003, pp. 3140–3145.

[53] A. Gonçalves, J. Abrantes, G. Saponaro, L. Jamone, and A. Bernardino, "Learning intermediate object affordances: Towards the development of a tool concept," in *Proc. Joint IEEE Int. Conf. Develop. Learn. Epigen. Robot. (ICDL-Epirob)*, Oct. 2014, pp. 482–488.

[54] L. Montesano, M. Lopes, A. Bernardino, and J. Santos-Victor, "Learning object affordances: From sensory–motor coordination to imitation," *IEEE Trans. Robot.*, vol. 24, no. 1, pp. 15–26, Feb. 2008.

[55] R. E. Fikes and N. J. Nilsson, "Strips: A new approach to the application of theorem proving to problem solving," *Artif. Intell.*, vol. 2, no. 3, pp. 189–208, 1972.

[56] T. Power, *Play and Exploration in Children and Animals*. Hillsdale, NJ, USA: Lawrence Erlbaum Assoc., 2000.

[57] D. E. Berlyne, "Curiosity and exploration," *Science*, vol. 153, no. 3731, pp. 25–33, 1966.

[58] A. Barto, M. Mirolli, and G. Baldassarre, "Novelty or surprise?" *Front. Psychol. Cogn. Sci.*, vol. 4, no. 907, pp. 1–15, 2013.

[59] U. Nehmzow, E. Gatsoulis, Y. A. Kerr, J. Condell, N. Siddique, and T. M. McGinnity, "Novelty detection as an intrinsic motivation for cumulative learning robots," in *Intrinsically Motivated Learning in Natural and Artificial Systems*, G. Baldassarre and M. Mirolli, Eds. Berlin, Germany: Springer-Verlag, 2013, pp. 185–207.

[60] P. Y. Oudeyer and F. Kaplan, "What is intrinsic motivation? A typology of computational approaches," *Front. Neurorobot.*, vol. 1, p. 6, Nov. 2009. [Online]. Available: https://doi.org/10.3389/neuro.12.006.2007

[61] G. Baldassarre and M. Mirolli, "Deciding which skill to learn when: Temporal-difference competence-based intrinsic motivation (TD-CB-IM)," in *Intrinsically Motivated Learning in Natural and Artificial Systems*, G. Baldassarre and M. Mirolli, Eds. Berlin, Germany: Springer-Verlag, 2013, pp. 257–278.

[62] F. Mannella *et al.*, "Know your body through intrinsic goals," *Front. Neurorobot.*, vol. 12, no. 30, pp. 1–17, 2018.

[63] M. Schlesinger, "Investigating the origins of intrinsic motivation in human infants," in *Intrinsically Motivated Learning in Natural and Artificial Systems*, G. Baldassarre and M. Mirolli, Eds. Berlin, Germany: Springer-Verlag, 2013, pp. 367–392.

[64] D. Caligiore *et al.*, "Intrinsic motivations driving learning of eye movements: An experiment with human adults," *PLOS ONE*, vol. 10, no. 3, 2015, Art. no. e0118705.

[65] V. Sperati and G. Baldassarre, "A bio-inspired model learning visual goals and attention skills through contingencies and intrinsic motivations," *IEEE Trans. Cogn. Develop. Syst.*, vol. 10, no. 2, pp. 326–344, Jun. 2018.

[66] V. Tikhanoff, U. Pattacini, and G. Natale, and L. Metta, "Exploring affordances and tool use on the iCub," in *Proc. 13th IEEE-RAS Int. Conf. Humanoid Robots (Humanoids)*, 2013, pp. 130–137.

[67] A. Stoytchev, "Behavior-grounded representation of tool affordances," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2005, pp. 3060–3065.

[68] V. Tikhanoff, A. Cangelosi, P. Fitzpatrick, G. Metta, L. Natale, and F. Nori, "An open-source simulator for cognitive robotics research: The prototype of the iCub humanoid robot simulator," in *Proc. IEEE Workshop Perform. Metrics Intell. Syst. Workshop*, 2008, pp. 57–61.

[69] G. Sandini, G. Metta, and D. Vernon, "The iCub cognitive humanoid robot: An open-system research platform for enactive cognition," in *50 Years of Artificial Intelligence. Essays Dedicated to the 50th Anniversary of Artificial Intelligence*, vol. 4850, M. Lungarella, F. Iida, J. Bongard, and R. Pfeifer, Eds. Berlin, Germany: Springer-Verlag, 2007.

[70] A. Cangelosi and M. Schlesinger, *Developmental Robotics: From Babies to Robots*. Boston, MA, USA: MIT Press, 2015.

[71] K. Seepanomwan, V. G. Santucci, and G. Baldassarre, "Intrinsically motivated discovered outcomes boost user's goals achievement in a humanoid robot," in *Proc. 7th Joint IEEE Int. Conf. Develop. Learn. Epigen. Robot. (ICDL-EpiRob)*, Sep. 2017, pp. 178–183.

[72] S. Schaal, *Dynamic Movement Primitives—A Framework for Motor Control in Humans and Humanoid Robotics*. vol. 22, Berlin, Germany: Springer, 2006, pp. 261–280.

[73] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, "Building machines that learn and think like people," *Brain Behav. Sci.*, vol. 40, pp. 1–72, Nov. 2017.

[74] D. Caligiore, P. Tommasino, V. Sperati, and G. Baldassarre, "Modular and hierarchical brain organization to understand assimilation, accommodation and their relation to autism in reaching tasks: A developmental robotics hypothesis," *Adapt. Behav.*, vol. 22, no. 5, pp. 304–329, 2014.

[75] R. S. Sutton, D. Precup, and S. P. Singh, "Between MDPS and semi-MDPS: A framework for temporal abstraction in reinforcement learning," *Artif. Intell.*, vol. 112, nos. 1–2, pp. 181–211, 1999.

[76] S. Forestier, Y. Mollard, and P.-Y. Oudeyer, "Intrinsically motivated goal exploration processes with automatic curriculum learning," 2017. [Online]. Available: arXiv.1708.02190v1.

[77] K. Seepanomwan, D. Caligiore, A. Cangelosi, and G. Baldassarre, "Generalisation, decision making, and embodiment effects in mental rotation: A neurorobotic architecture tested with a humanoid robot," *Neural Netw.*, vol. 72, pp. 31–47, Dec. 2015.

[78] R. Sutton, "Integrated architectures for learning, planning, and reacting based on approximating dynamic programming," in *Proc. 7th Int. Conf. Mach. Learn.*, vol. 216, 1990, pp. 216–224.
[79] F. Taffoni *et al.*, "A mechatronic platform for behavioral analysis on nonhuman primates," *J. Integr. Neurosci.*, vol. 11, no. 1, pp. 87–101, 2012.
[80] F. Taffoni *et al.*, "Development of goal-directed action selection guided by intrinsic motivations: An experiment with children," *Exp. Brain Res.*, vol. 232, no. 7, pp. 2167–2177, Apr 2014.
[81] A. Nair, V. Pong, M. Dalal, S. Bahl, S. Lin, and S. Levine, "Visual reinforcement learning with imagined goals," in *Proc. 2nd Workshop Lifelong Learn. Reinforcement Learn. Approach (LLRLA FAIM)*, Stockholm, Sweden, 2018, pp. 9209–9220.
[82] F. Stulp, G. Raiola, A. Hoarau, S. Ivaldi, and O. Sigaud, "Learning compact parameterized skills with a single regression," in *Proc. 13th IEEE-RAS Int. Conf. Humanoid Robots (Humanoids)*, 2013, pp. 417–422.
[83] E. Cartoni and G. Baldassarre, "Autonomous discovery of the goal space to learn a parameterized skill," May 2018. [Online]. Available: arXiv.1805.07547v1.
[84] S. Hart and R. Grupen, "Learning generalizable control programs," *IEEE Trans. Auton. Mental Develop.*, vol. 3, no. 1, pp. 216–231, Sep. 2011.
[85] G. Konidaris, L. P. Kaelbling, and T. Lozano-Perez, "Constructing symbolic representations for high-level planning." in *Proc. Assoc. Adv. Artif. Intell. (AAAI)*, 2014, pp. 1932–1938.
[86] K. Liu, T. Kurutach, C. Tung, P. Abbeel, and A. Tamar, "Hallucinative topological memory for zero-shot visual planning," 2020. [Online]. Available: arXiv 2002.12336v1.
[87] C. J. C. H. Watkins and P. Dayan, "*Q*-learning," in *Proc. Mach. Learn.*, 1992, pp. 279–292.
[88] M. I. Jordan and D. E. Rumelhart, "Forward models: Supervised learning with a distal teacher," *Cogn. Sci. J.*, vol. 16, no. 3, pp. 307–354, 1992.
[89] T. Kohonen, "Self-organized maps of sensory events," *Philos. Trans. Roy. Soc. A Math. Phys. Eng. Sci.*, vol. 361, pp. 1177–1186, May 2003.
[90] S. Schaal, P. Mohajerian, and A. Ijspeert, "Dynamics systems vs. optimal control—A unifying view," *Progr. Brain Res.*, vol. 165, pp. 425–445, Oct. 2007.
[91] S. J. Paterson, S. Heim, J. T. Friedman, N. Choudhury, and A. A. Benasich, "Development of structure and function in the infant brain: Implications for cognition, language and social behaviour," *Neurosci. Biobehav. Rev.*, vol. 30, no. 8, pp. 1087–1105, 2006.
[92] D. Caligiore, D. Parisi, and G. Baldassarre, "Integrating reinforcement learning, equilibrium points and minimum variance to understand the development of reaching: A computational model," *Psychol. Rev.*, vol. 121, no. 3, pp. 389–421, 2014.
[93] F. Mannella, M. Mirolli, and G. Baldassarre, "The interplay of pavlovian and instrumental processes in devaluation experiments: A computational embodied neuroscience model tested with a simulated rat," in *Modelling Perception With Artificial Neural Networks*, C. Tosh and G. Ruxton, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2010, pp. 93–113.

Kristsana Seepanomwan received the M.Sc. degree in computer science from the King Mongkuts' University of Technology, North Bangkok, Thailand, in 2007, and the Ph.D. degree in computing from the Centre for Robotics and Neural System (CRNS), School of Computing, University of Plymouth, Plymouth, U.K., in 2016, with a thesis titled "Mental Imagery in Humanoid Robots."

He is a Lecturer with the Department of Computing, Faculty of Science, Silpakorn University, Nakhon Pathom, Thailand. His research interests concern cognitive robotics, robot learning, and brain-like processing mechanisms, with a focus on decision making, mental imagery, and planning in humanoid robots.

Dr. Seepanomwan got a scholarship from the Thai Government for doctoral program in the field of robotics and enrolled as a Ph.D. candidate at CRNS, School of Computing, University of Plymouth in 2011.

Daniele Caligiore received the master's degree in electronics engineering from the Universita' degli Studi di Catania, Catania, Italy, in 2003, and the Ph.D. degree in biomedical engineering from the Universita' Campus Bio-Medico di Roma, Rome, Italy, in 2011.

He is a Researcher with the Institute of Cognitive Sciences and Technologies, Italian National Research Council, Rome. He has participated several European projects in the field of embodied cognition and developmental robotics (e.g., ROSSI—Emergence of Communication in Robots through Sensorimotor and Social Interaction; and IM-CLeVeR—Intrinsically Motivated Cumulative Learning Versatile Robots). His research interests include motor development, system-level computational modeling, computational models of brain diseases, and developmental robotics.

Dr. Caligiore was a Guest Editor for a consensus paper of the journal "The Cerebellum" titled "Toward a Systems-Level View of Cerebellar Function: The Interplay Between Cerebellum, Basal Ganglia and Cortex."

Kevin J. O'Regan received the B.Sc. degree in mathematical physics from University of Sussex, Brighton, U.K., in 1968, and the Ph.D. degree on "Constraints on Eye Movements in Reading" from the University of Cambridge, Cambridge, U.K., in 1975.

He is an emeritus Director of the Laboratoire Psychologie de la Perception, CNRS, Université Paris Descartes, Paris, France. After early work on eye movements in reading, he was led to question established notions of the nature of visual perception, and to discover, with collaborators, the phenomenon of "change blindness." In 2011, he published a book *Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness* (Oxford University Press). In 2013, he obtained a five year Advanced ERC grant for the project FEEL directed to explore his "sensorimotor" approach to consciousness in relation to sensory substitution, pain, color, space perception, developmental psychology, and robotics. Currently, within the European project GOAL-Robots, he is exploring how young infants learn the structure of their bodies.

Gianluca Baldassarre received the B.A. and M.A. degrees in economics from the Sapienza University of Rome, Rome, Italy, in 1998, where he attended a Specialization Course in "Cognitive Psychology and Neural Networks," in 1999, and the Ph.D. degree in computer science from the University of Essex, Colchester, U.K., on "Planning With Neural Networks," in 2003.

He did a postdoctoral research with the Italian Institute of Cognitive Sciences and Technologies, National Research Council (ISTC-CNR), Rome. In 2006, he was a Researcher with ISTC-CNR, where he is the Founder and the Coordinator of the Research Group, Laboratory of Computational Embodied Neuroscience. From 2006 to 2009, he was the Principal Investigator for the EU project "ICEA—Integrating Cognition Emotion and Autonomy"; the Coordinator of the European Integrated Project "IM-CLeVeR—Intrinsically Motivated Cumulative Learning Versatile Robots" from 2009 to 2013; and since 2016, he has been the Coordinator of the EU FET Project "GOAL-Robots—Goal-Based Open-Ended Autonomous Learning Robots" ending in 2020. His research interests are on open-ended learning of sensorimotor skills driven by extrinsic and intrinsic motivations in animals, humans, and robots. He has more than 100 international peer-reviewed publications in the above areas.