This preprint may not exactly replicate the final version of the article published in: *Adaptive Behavior*, vol. 22, no. 5, pp. 304-329.
It is not the copy of record. The final publication is available at SAGE via http://adb.sagepub.com/content/22/5/304

Modular and hierarchical brain organization to understand assimilation, accommodation and their relation to autism in reaching tasks: a developmental robotics hypothesis

Daniele Caligiore[1], Paolo Tommasino[2], Valerio Sperati[1] and Gianluca Baldassarre[1]
1 Laboratory of Computational Embodied Neuroscience,
Istituto di Scienze e Tecnologie della Cognizione,
Consiglio Nazionale delle Ricerche (LOCEN-ISTC-CNR), Italy
2 School of Mechanical and Aerospace Engineering,
Nanyang Technological University, Singapore
corresponding author: Daniele Caligiore
email: daniele.caligiore@istc.cnr.it
Tel:+39 06 49 936 230

By "assimilation" the child embodies the sensorimotor experience into already built mental structures. Conversely, by "accommodation" these structures are changed according to the child's new experiences. Despite the intuitive power of these concepts to trace the course of sensorimotor development, they have gradually lost ground in psychology. This likely for a lack of brain related views capturing the dynamic mechanisms underlying them. Here we propose that brain modular and hierarchical organization is crucial to understanding assimilation/accommodation. We devised an experiment where a bio-inspired modular and hierarchical mixture-of-experts model guides a simulated robot to learn by trial-and-error different reaching tasks. The model gives a novel interpretation of assimilation/accommodation based on the functional organization of the experts allocated through learning. Assimilation occurs when the model adapts a copy of the expert trained for solving a task to face another task requiring similar sensorimotor mappings. Experts storing similar sensorimotor mappings belong to the same functional module. Accommodation occurs when the model uses non-trained experts to face tasks requiring different sensorimotor mappings (generating a new functional group of experts). The model provides a new theoretical framework to investigate impairments in assimilation/accommodation processes and proposes that such impairments might be related to the autistic syndrome.
Keywords: Brain functional modularity, brain functional graph, Piaget, autism, connectome, hierarchical reinforcement learning, Mixture of experts neural networks.

Introduction

The theoretical concepts of assimilation and accommodation were used by Piaget to indicate the cardinal processes by which the child constructs sensorimotor structures for behaving adaptively in the world. By the process of *assimilation* the *child incorporates* the outer world experience into the internal structures that are already built in her/his mind. Conversely, by *accommodation* these structures are changed according to the new experience the child makes in the world (Piaget, 1953). Assimilation is necessary as it assures the continuity of structures and the integration of new elements into these structures. On the other side, accommodation is important to permit structural changes and the transformation of structures as a function of the new elements encountered. In this perspective, assimilation and accommodation concepts might crucially help to understand, and theoretically frame, important developmental processes related, for example, to the cumulative learning of motor skills and to the transfer of learned skills to new conditions (Caligiore, Mirolli, Parisi, & Baldassarre, 2010; Tommasino, Caligiore, Mirolli, & Baldassarre, 2012; Taylor & Stone, 2009; Tommasino, Caligiore, Mirolli, & Baldassarre, submitted). Assimilation and accommodation are closely related to the con-

This research has received funds from the European Commission under the 7th Framework Programme (FP7/2007-2013), ICT Challenge 2 "Cognitive Systems and Robotics", project "IM-CLeVeR – Intrinsically Motivated Cumulative Learning Versatile Robots", grant agreement no. ICT-IP-231722. We are very grateful to one anonymous reviewer who pointed out the importance of stressing the coexistence of assimilation and accommodation processes in most learning experiences, and its implications.

cept of "schema". Piaget used this term to indicate the basic building block used by the child to organize knowledge. Schemas are "units" of knowledge each relating to objects, actions, interpretation processes, predictions, etc. Assimilation implies the use of existing schemas to deal with new experiences. Accommodation, in contrast, happens when existing schemas do not work as desired with the new situations and thus need to be suitably updated.

Despite their highly intuitive power, assimilation and accommodation notions have gradually lost ground in developmental psychology. The possible reason for this is that the two concepts are formulated in generic/descriptive rather than mechanistic/generative terms (Forssell, 2004). In this respect, some authors have arrived to say that Piaget's "theory is so vague as to be virtually unfalsifiable" (Boden, 1994). We agree that the concepts of assimilation and accommodation would have greatly benefited from more specific and operational definitions. However, we also think that they contain a critical intuition on child development that should be preserved and valued as they hints to *key mechanisms* that might underlie development. In particular, this is the idea that an important aspect of development can be understood in terms of assimilation as a *re-use of existing internal structures* to face novel experiences, when these are sufficiently similar to previous ones, and accommodation as a *progressive modification of such structures* when they are increasingly dissimilar from them. How to value this intuition?

Some authors argue that the weakness of the assimilation and accommodation concepts could be due to the lack of a *brain related* framework to concretely express the dynamic mechanisms underlying them (Mareschal, 2003; Parisi & Schlesinger, 2002). An important attempt to cope with this issue has been made using computational models. In this line, connectionist scientists used simple neural networks models as tools to make explicit and less abstract the computational mechanisms behind assimilation and accommodation (McClelland, 1995; Elman et al., 1996). Basically, according to this approach the changes in the neural network weights represent a form of accommodation whereas the transformation (by the network weights) of input patterns into internal patterns of activation constitutes assimilation (Parisi & Schlesinger, 2002; Rasheed & Ali, 2009). Tani and Nolfi (1999) and Nishimoto and Tani (2009) presented other neural hierarchical and modular architectures linked to the analogous structure of brain and capable of capturing some aspects of assimilation and accommodation on the basis of supervised learning processes. Sugimoto, Haruno, and Kawato (2012) proposed another relevant hierarchical and modular architecture using reinforcement learning to select among different experts together with other prediction-based mechanisms. Sec. "Related works" expands the review of these and others models relevant to study assimilation and accommodation, and presents a comparison with the model presented here.

This article proposes that the modular and hierarchical organization of the brain (Hamilton & Grafton, 2007; Fuster, 2001; Chen, He, Rosa-Neto, Germann, & Evans, 2008; Thill, Caligiore, Borghi, Ziemke, & Baldassarre, 2013) plays a key role in understanding these adaptive processes, and it is the key to specify and value the Piagetian concepts of assimilation and accommodation. The modular and hierarchical organization of the brain is a general anatomical and physiological design principle involving both cortical and subcortical areas (Baldassarre, Caligiore, & Mannella, 2013; Meunier, Lambiotte, & Bullmore, 2010; Schwarz, Gozzi, & Bifone, 2009). At cortical level, for example, the motor cortex is organised by modular neural columns. Different assemblies of columns might participate to encode multiple repertoires of skills, from limited to very different (Aflalo & Graziano, 2011). At sub-cortical level the basal ganglia have a hierarchical structure that is based on partially segregated loops linked to different cortical areas (Alexander, DeLong, & Strick, 1986; Middleton & Strick, 1996, 2000). Different loops encode motor actions (in particular, the loops with motor and premotor cortex), or context and goals (in particular, the loops with prefrontal cortex). These loops seem to be characterized by certain degree of modularity possibly subserving the encoding of different actions and goals (Gurney, Prescott, & Redgrave, 2001a, 2001b). In the same line, another sub- cortical area crucially involved in motor skill acquisition and expression, the cerebellum, receives input from, and sends output to, the cerebral cortex through multisynaptic partially-segregated loops performing distinct functional operations (Middleton & Strick, 2000; Strick, Dum, & Fiez, 2009; Houk et al., 2007; Caligiore, Pezzulo, Miall, & Baldassarre, 2013).

The modular organization of the brain has a number of advantages. It allows brain to break problems into identifiable sub-tasks thus making possible to use neural chunks encoding them across multiple problems (Schwarz et al., 2009). Moreover, it allows a faster adaptation of the system in response to a changing environment since the opportunity to build new skills from a rich repertoire of existing skills is more efficient than generating new abilities from scratch (Meunier et al., 2010). Modularity can also act to improve robustness and evolvability (Calabretta, Di Ferdinando, Wagner, & Parisi, 2003).

To support our proposal we devised an experiment using the TERL model (Tommasino et al., 2012, submitted). This model captures some essential aspects of the modular and hierarchical arrangement of brain. In particular, as we shall see it is based on a hierarchical modular mixture-of-experts neural network architecture (Jacobs, Jordan, Nowlan, & Hinton, 1991) adapted to work with reinforcement learning. The model is used here to autonomously learn the behaviour of a simulated humanoid robot engaged in the acquisition of different reaching skills. While the system solves new tasks, when useful it starts from previously acquired neural structures (assimilation) and then updates them (accommodation) to an extent that depends on the degree of similarity of such tasks with previously solved tasks.

The model's sensorimotor competence, and the capacity to evaluate actions, pivots on neural *experts*. The model modifies these experts to acquire new skills in various ways. It can re-use experts as they are, or slightly modify them, to solve tasks that require the same or very similar sensorimotor mappings. Alternatively, when the new tasks share some feature with the solved tasks, but are also quite different from them, the system can learn to solve them starting from experts that are "copies" of experts already trained to solve previous tasks. This allows the system to transfer useful knowledge from previously solved tasks to the new tasks, and at the same time to avoid disrupting the acquired information. Finally, the new tasks might require the development of sensorimotor mappings so unrelated to all previous experiences that starting from the modification of copies of current experts would be worse than "starting from scratch", i.e. from simple-default/random/innate experts. Here we will consider the set of experts which store similar sensorimotor mappings as forming a *functional module*, and the experts storing different sensorimotor mappings as forming different functional modules. This perspective agrees with recent neuroscientific evidence suggesting that the primate motor cortex does not contain a simple map of the body's muscles as proposed by the traditional somatotopic view (Penfield & Boldrey, 1937) but it rather contains a topographical map of behaviorally relevant actions clustered in the neural space according to their behavioural similarity (Meier, Aflalo, Kastner, & Graziano, 2008; see Ring, Schaul, & Schmidhuber, 2011, for a model capturing this type of organisation but not linked to assimilation and accommodation).

In the interpretation proposed here, the basic *mechanism underlying assimilation* will be considered the *re-use of previously developed connection weights to solve new tasks in more efficient ways with respect to starting from scratch*. Instead, the *mechanisms underlying accommodation* will be considered those that lead to *modifications of previously developed neural structures*, in particular: the modification of experts already used to solve previous tasks, the modification of neural copies of them, or the modification of experts never used in previous experiences. These modifications involve changes of increasing complexity, from changes of the connection weights within one expert, to changes of a functional module, and finally to the creation of a new functional module.

This interpretation of assimilation and accommodation has the strength of being formulated in terms of *neural mechanisms*, so it specifies the original Piagetian terms formulated only at a functional/behavioural level. Moreover, it also departs from, and supposedly improves, a previous interpretation we proposed in Tommasino et al. (2012). In this work, we already used a hierarchical modular model, a precursor of the one presented here, and mentioned a preliminary possible interpretation of assimilation and accommodation concepts. The present work fully expands this issue and presents a related, but also rather different, interpretation of assimilation and accommodation. In particular, the different interpretation is based on the idea that most of times assimilation and accommodation mechanisms are *simultaneously present* rather than working in isolation as assumed by Tommasino et al. (2012), and that the two processes are present in different degrees depending on the similarity/novelty of the new tasks with respect to the solved ones. The view is in line with Piaget's perspective for which most motor and higher-level mental activities involve both assimilation and accommodation. In fact, there are only two special cases in which they operate in a "pure" form: in "imitation", mainly based on accommodation, and in "play", mainly based on assimilation (Piaget, 1951).

The accommodation processes modelled in the paper captures Piaget's general idea on the change of existing schemas within the rather simplified domain of motor behaviour, and in particular with respect to a particular behaviour, namely *reaching*. This might appear a restrictive test with respect to the original concept proposed by Piaget. Indeed, Piaget talked of accommodation, *at a behavioural level*, to refer to substantial changes of existing schemas, e.g. to acquire behaviours which are qualitatively different from those already acquired (e.g., pushing, blocking, throwing, etc.). Moreover, he talked not only of *sensorimotor* schemas but also of interpretative schemas, prediction schemas, and other types of schemas capturing mind contents related to higher-level cognition. For the sake of focussing on *key neural processes*, however, here we will refer only to sensorimotor behaviour and in particular to reaching, with the changes of "schemas" involving only some features of this behaviour, in particular the location of the reached objects. Notwithstanding this simplification, we expect that the principles illustrated here *at the neural level* should scale up to capture the learning of qualitatively different actions (sensorimotor schemas) as those mentioned above. The reason is that the performance of any action ultimately requires to control the trajectories and/or final states of body configurations in space. Thus the principles proposed here should continue to hold although the specific controllers used (in particular the experts implementing the actions) should become more sophisticated to be able to produce richer body trajectories and states that might be behaviourally deemed as "different actions". Also, we expect that to some extent the neural mechanism proposed here to capture the processes of assimilation and accommodation might also be relevant for other domains of cognition, as further discussed in Sec. "Discussion".

The model presented here provides a new theoretical framework to study the possible consequences of damaging assimilation and accommodation processes during sensorimotor learning. This kind of deficit might occur in human subjects showing a lack in the abilities to discriminate and generalize motor behaviours, as one sees in autistic disorder (Burack, Charman, Yurmiya, & Zelazo, 2001; Gowen & Hamilton, 2012). Although sensorimotor impairments are not considered to be core features in the Autism Spectrum Disorder (hereafter we will refer to this as simply "autism"), there is increasing acknowledgment that they are instead present and can have a significant impact on quality of life and social development (see Gowen & Hamilton, 2012 for a recent review). As acquisition of good sensorimotor skills is important for a range of everyday abilities such as communication and language development (Gernsbacher, Sauer, Geye, Schweigert, & Hill, 2008), playing and interacting with others (Clearfield, 2011), mental imagery (Williams, Thomas, Maruff, & Wilson, 2008) and perception (Blaesi & Wilson, 2010; Eskenazi, Grosjean, Humphreys, & Knoblich, 2009), it is likely that abnormal development of sensorimotor behaviour can have far reaching consequences on development of other cognitive functions (Gernsbacher et al., 2008). Moreover, as mentioned above, the same neural mechanism studied here for motor behaviour might underlie also other cognitive functions as the basic modular organisation of basal ganglia and cortex tend to repeat at different levels of complexity of cognition, from the motor to the premotor cortex areas underlying sensorimotor behaviour, to prefrontal cortex areas underlying decision making, planning and reasoning.

Some pivotal researches suggested that assimilation and accommodation concepts may help to trace more precisely the course of aberrations in the autistic sensorimotor development (Burack et al., 2001; Morgan, 1986; Rosenthal, Massie, & Wulff, 1980). These studies, however, were not further developed likely because they did not rely on a proposal about the brain mechanisms potentially responsible for impaired functions during sensorimotor development in autism (Morgan, 1986).

The computational approach proposed here overcomes this limitation as it suggests an explicit link between assimilation and accommodation and the possible brain mechanisms underlying them. In more detail, here we use the model to propose that: (a) in the early stage of sensorimotor learning autistic brain prefers to accommodate rather then assimilate; (b) the abnormal accommodation mechanism in autism is reflected by a weak development of intra-module connections during the formation of functional modules. Importantly, this latter point is in agreement with recent imaging literature showing a deficit in brain functional modularity in autistic subjects (Boersma et al., 2013; Catarino et al., 2013; Meunier et al., 2010).

The rest of the paper is organized as follows. Sec. "The TERL model: biological and computational constraints" presents recent evidence supporting the use of a hierarchical reinforcement learning architecture to study assimilation and accommodation. Sec. "Methods" presents the robot and the task used to test the model and explains the functioning and the learning mechanisms of this. Sec. "Results" shows the results obtained by testing the system, and Sec. "Discussion" discusses such results. Sec. "Related works" presents a focused overview of theories and computational models capturing important mechanisms related to assimilation and accommodation. Sec. "Conclusions and future work" draws the conclusions and suggests future work.

The TERL model: biological and computational constraints

To investigate the mechanisms underlying assimilation and accommodation we used the neural network model TERL (Figure 2) presented in Tommasino et al. (2012) (see also Tommasino et al., submitted and Baldassarre, 2002; Caligiore, Mirolli, et al., 2010, for neural architectures represented precursors of TERL). TERL was developed taking into account several bio-inspired and computational constraints that could be critical for the aim of this paper. First, TERL is a reinforcement learning (RL) actor-critic system (Sutton & Barto, 1998) which can be used to abstract the biological action learning in the basal ganglia: the actor in particular has been suggested to correspond to the *matriosomes* whereas the critic to the *striosomes* compartments of *striatum*, the basal-ganglia input station (Houk, Adams, & Barto, 1995).

Second, the actor-critic components of TERL have a hierarchical and modular architecture formed by a *gating network* and a number of *experts*, as in the mixture of experts models (Jacobs et al., 1991; Jordan & Jacobs, 1994). These structures can be used to model the softly-modular redundant organization of some parts of brain, in particular: (a) the organisation of motor cortex based on neural columns (Aflalo & Graziano, 2011); (b) the organisation of basal ganglia in *channels*, where channels support the trial-and-error learning and the selection of different actions (Gurney et al., 2001b; Graybiel, 1998).

Third, the functioning and learning algorithms of all components of TERL have been modified for working with a continuous RL system (cf. Baldassarre, 2002; Caligiore, Mirolli, et al., 2010). This allows TERL to drive the behaviour of an embodied system (here the simulated humanoid robot iCub) interacting with an environment with continuous states through continuous actions similarly to real organisms.

Finally, the learning algorithm used by TERL relies upon the temporal difference (TD) reinforcement learning algorithm (Sutton & Barto, 1998). This has been shown to be able mimic the processes that guide trial-and-error learning processes through which infants and adults acquire reaching skills (Berthier, 1996; Berthier, Rosenstein, & Barto, 2005; Caligiore, Guglielmelli, Parisi, & Baldassarre, 2010; Herbort, Ognibene, Butz, & Baldassarre, 2007; Caligiore, Parisi, & Baldassarre, 2014). Moreover, the TD-error signal has been shown to have the same dynamics, during learning, of biological learning signals based on phasic dopamine, an important brain neuromodulator supposed to guide trial-and-error learning in organisms (Houk, Davis, & Beiser, 1995; W. Schultz, Dayan, & Montague, 1997).

Importantly, all these features make TERL able to autonomously decide whether to encode skills in the same or in different neural structures (experts) on the basis of the similarities and differences between the required sensorimotor mappings. In this way, the model can face the problem of deciding if and which of previously learned skills can be used as a starting point to solve new tasks faster than starting from scratch. This kind of behaviour, important to explain assimilation and accommodation, has also recently received attention within the RL community under the research agenda called *Transfer Reinforcement Learning* (TRL; see Taylor & Stone, 2009, for a recent survey). Within TRL framework, the challenge of transfer is described in these terms: identifying the possible source tasks, among those previously learned, on the basis of which to learn a new target task so as to maximize the transfer of knowledge and decrease the learning time needed to achieve the steady-state performance.

Among the TRL systems relevant for the problems faced here, one (Fernández & Veloso, 2006) proposes a RL method that, like TERL, explicitly reasons on a library of already acquired policies to solve a new task and selects already acquired policies to be re-used on the basis of the reward they obtain in the new task. However, the method strongly exploits the off-policy learning features of Q-learning, tests the source tasks (policies) serially, and was developed for grid-work tasks. For these reasons it would be difficult to use it in robotic contexts. Recent works on TRL highlight how we still lack systems that can solve this problem in principled ways (Taylor & Stone, 2009). TERL deals with this issue. Indeed, TERL proposes a novel and effective way to solve TRL issues pivoting on a reinforcement learning version of the mixture of experts hierarchical architecture (the acronym TERL stands for Transfer Expert Reinforcement Learning).

Methods

The robot and task used to test the model

In this article the assimilation and accommodation processes are studied by considering the development of a reaching skill. In more detail, we devised an experiment where the TERL model controls the motor behaviour of a 3D four degrees of freedom (4DOFs) simulated humanoid robotic arm involved in learning a reaching task.

Figure 1 shows the simulated humanoid robot iCub (Tikhanoff et al., 2008) and the environment used to test the TERL model. The iCub simulator replicates the same body and control scheme of the real iCub robot (Sandini, Metta, & Vernon, 2007), an open source robotic platform built for studying cognitive development in children. The iCub's body roughly reproduces the body of a five year old child. Each arm of the iCub 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}) ($http://wiki.icub.org/wiki/ICub\_joints$). Here we used TERL to control the movements of four joints of the right arm. In particular, TERL gives motor commands for the "shoulder pitch" joint J_0, responsibles for the front-back movement when the arm is aligned with gravity; for the 'shoulder roll" J_1, affecting the adduction-abduction movement of the arm; for the 'shoulder yaw" J_2, affecting the yaw movement when the arm principal axis is aligned with gravity; and for J_3 affecting the elbow angle. The positions of the remaining joints are set at fixed values ($J_4 = -10°$; $J_5 = -30°$; $J_9 = 80°$; $J_{6-8} = J_{10-15} = 0°$). The torso joint affecting the yaw with respect to the gravity is fixed to $-30°$. During the simulation J_0 can assume values in the range [-80°;-15°], J_1 in the range [10°;110°]; J_2 in the range [-10°;75°]; and J_3 in the range [20°;85°].

The 3D environment used to test the model is formed by three spherical objects (diameter equal to 15cm) displaced around the robot (Figure 1). The objects allow the formation of three reaching tasks, each requiring that the model learns how to control the right arm of the iCub in order to reach one specific object (object A, or object B, or object C). We will call these tasks "Task A", "Task B" and "Task C" depending on the target object.

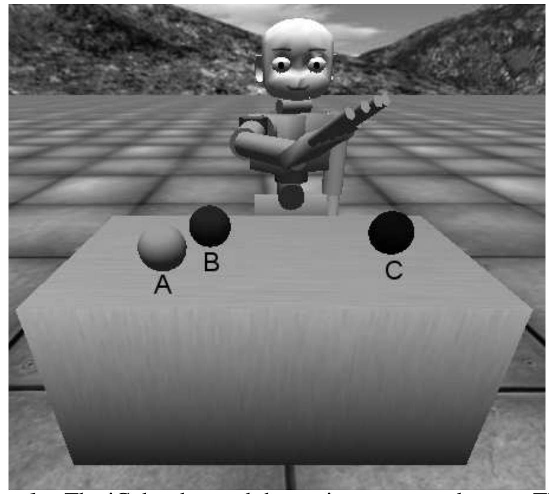

Figure 1. The iCub robot and the environment used to test TERL. The picture refers to the beginning of an trial. The objects A, B and C represent the target objects that the robot has to reach.

TERL has to learn the three tasks in a sequential fashion, each for 5000 trials. All trials involving the solution of the tasks start with the arm set at fixed posture ($J_0 = -90°$, $J_1 = 50°$, $J_2 = 90°$, $J_3 = 40°$) corresponding to the center of the workspace (Figure 1). During learning of a task the arm randomly explores the environment and each trial ends when the hand hits the target object, or when a time out of 8.0s occurs. If simulation cycles when iCub's hand touches the target object the model receives a reward equal to 1.0, otherwise it receives a reward of 0.0. We now explain the functioning and learning of TERL.

Functioning of TERL

For clarity the main symbols used in this section are reported in Figure 2 showing the model architecture. Figure 3 shows the information flow between the components of the model during functioning and learning.

Input The system gets two types of inputs: the current goal and the arm posture. The gating networks get as input the current task, or *goal*, encoded with a different binary vector for different objects: A=[1,0,0], B=[0,1,0] and C=[0,0,1]. As the Task A and the Task B require similar sensorimotor mapping but are coded with two different input goals sent to the gating networks, as they allow the evaluation of TERL capacity of reusing the expert used to solve Task A when learning to solve Task B. By contrast, as Task C requires a rather different sensorimotor mapping with respect to Task A and Task B, TERL has to use a novel expert not used to solve the previous tasks. The complex formed by the gating network and the goal input vector reproduces in an abstract way the role played by the ventro-medial and orbitofrontal portions of the prefrontal cortical areas that reciprocate connections with the basal ganglia and are important for the selection of goals and, via these, the sensorimotor mappings that lead to pursue them (Yin & Knowlton, 2006; Baldassarre et al., 2013). This aspect is further discussed in Sec. "Conclusions and future work".

The experts get as input the arm postures ($J_0(t)$, $J_1(t)$, $J_2(t)$, $J_3(t)$) encoded in a four dimensional neural map (with *population coding*, cf. Pouget & Latham, 2003) formed by 8^4 normalized Gaussian radial basis function units x_i (as in Doya, 2000). The choice of using the arm proprioception (joint angles) as sensory input for the actor, rather than a combination of both proprioceptive and visual signals, is based on empirical evidence suggesting that at its onset reaching is strongly based on proprioception. Vision, instead, plays an important role in indicating the approximate position of the target in space, possibly on the basis of the proprioception of the gaze direction (Berthier & Carrico, 2010; Carrico & Berthier, 2008) (in the model, information on the target position is represented by the information sent to the selectors). These ideas agree with experimental evidence showing that the first reaching attempts in infants do not require visual perception of hands or arms although vision is sufficiently developed to provide a good sense of the target location in the reachable space (Thelen et al., 1993; Clifton, Muir, Ashmead, & Clarkson, 1993). In addition, in adults proprioception plays a key role in guiding reaching in the early phases of the movement while vision is important in the later phases when the hand arrives in proximity of the target (Sarlegna et al., 2003). These assumptions are shared with several influent models on reaching development preceding this work. In particular, the model proposed in Berthier et al. (2005) used as input the arm joint angles and velocities, and the more abstract model presented in Berthier (1996) used as input the sensed position of the end effector.

The difference in the input between the gating and the experts networks reflects what is done in the transfer reinforcement learning literature where the system is typically informed about the task it is facing. The different task in most cases have to be accomplished in the same environment (as here), and the input (here the arm proprioception) sent to the part of the system that have to solve the task (here the experts) does not change (but there are other possibilities, see Taylor & Stone, 2009). We cannot expand this issue here, but this arrangement can be considered an abstraction of the hierarchical organisation of the striato-cortical loops in real brains, the core structures that underpin trial-and-error learning in organisms (Botvinick, Niv, & Barto, 2008). Moreover, the fact that the behaviours acquired by the model are hierarchically driven by the goals of the system (encoded by the gating networks) is in line with the original account of Piaget as well as with recent perspectives on behaviour development (von Hofsten, 2007; Schlesinger & Cangelosi, 2013).

Actor gating network The actor gating network (AG) has ten output units (indexed with e) which receive the task input z_i via connections with weights w_{AGei}. The activation potential, p_{Ae}, of output unit e is filtered with a soft-max function, and the resulting activation, g_{Ae}, represents the *responsibility*

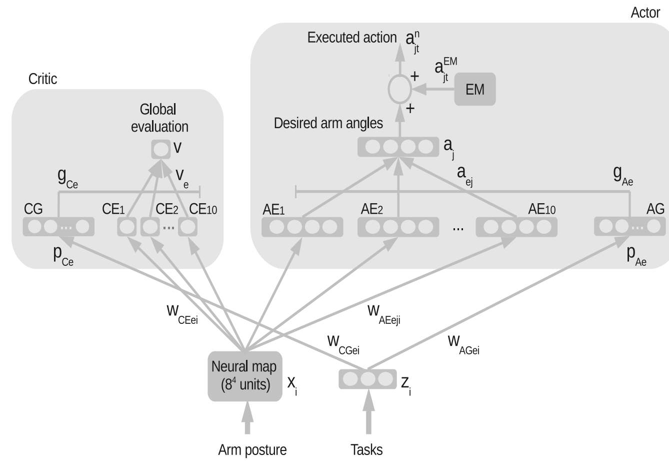

Figure 2. The TERL neural architecture used to study the mechanisms underlying assimilation and accommodation.

of expert e (Bayesian probability *prior*; Jacobs et al., 1991):

$$g_{Ae} = \frac{e^{(p_{Ae}/T)}}{\sum_{e=1}^{10} e^{(p_{Ae}/T)}} \quad (1)$$

where T, set to 0.1, is a "temperature" parameter allowing the enhancement or flattening of the differences between priors and hence of the relative contribution of experts to action.

Actor experts Each actor expert (AE_e) has four output units with sigmoidal activation a_{ej} which encode the control signals sent to the arm (the four desired joint angles). These output units receive input from the arm-posture map units x_i, via connections with weights w_{AEeji}, and a bias unit (constantly set to one). The global action a_j (desired arm angles) of the actor is computed on the basis of the actor gating network *priors*:

$$a_j = \sum_e g_{Ae} \cdot a_{ej} \quad (2)$$

To foster exploration, the executed action, a^n_j, includes noise, as further explained below.

Critic gating network The critic gating network (CG) works analogously to the AG on the basis of the connection weights, w_{CGei}, the unit activation potentials, p_{Ce}, and the *priors* of the critic experts g_{Ce}.

Critic experts Each critic expert (CE) has a linear output unit v_e encoding the evaluation of the current state and receives input from the arm-posture map units x_i via connections with weights w_{CEei}:

$$v_e = \sum_i w_{CEei} \cdot x_i \quad (3)$$

The global evaluation v of the critic is computed on the basis of the *priors*:

$$v = \sum_e g_{Ce} \cdot v_e \tag{4}$$

Learning signals

Global TD-error Couples of successive global evaluations, together with the reward r_t, are used to compute the global TD-error, δ_t, as in standard reinforcement learning (Sutton & Barto, 1998):

$$\delta_t = \begin{cases} r_t - v_{t-1} & if\ end\ of\ trial \\ (r_t + \gamma v_t) - v_{t-1} & if\ during\ trial \\ 0 & if\ start\ of\ trial \end{cases} \tag{5}$$

where γ is a discount factor ($\gamma = 0.99$).

Critic Experts TD-error The expert TD-error signals are calculated as follows:

$$\delta_{et} = \begin{cases} r_t - v_{et-1} & if\ end\ of\ trial \\ (r_t + \gamma v_{et}) - v_{et-1} & if\ during\ trial \\ 0 & if\ start\ of\ trial \end{cases} \tag{6}$$

Actor experts posterior responsibilities To train the actor experts and gating network the algorithm computes the adjusted responsibilities (Bayesian probability *posteriors*, Jacobs et al., 1991) of the experts as follows:

$$h_{Ae} = \frac{c_{Ae} \cdot g_{Ae}}{\sum_e [c_{Ae} \cdot g_{Ae}]} \tag{7}$$

where c_{Ae} is a measure of the *likelihood* that the actor expert e chooses the global action, $\mathbf{a}_t^n$:

$$c_{Ae} = e^{-0.5\frac{(D[\mathbf{a}_t^n, \mathbf{a}_{et}])^2}{\sigma^2}} \tag{8}$$

where $D[\mathbf{a}_t^n, \mathbf{a}_{et}]$ is the Euclidean distance between the two vectors encoding respectively the global action $\mathbf{a}_t^n$ and the action $\mathbf{a}_{et}$, computed by expert e. The width of the Gaussian (σ) is kept constant at 0.3. Notice that this formula implies a higher posterior responsibility for experts whose action was more similar to the noisy performed action.

Critic experts posterior responsibilities The *posteriors* of the critic experts are computed as follows:

$$h_{Ce} = \frac{c_{Ce} \cdot g_{Ce}}{\sum_e [c_{Ce} \cdot g_{Ce}]} \tag{9}$$

where c_{Ce} is a measure of the *likelihood* that the critic expert e gives an accurate evaluation producing a zero TD-error, and is computed as follows:

$$c_{Ce} = e^{-0.5(\delta_{et})^2} \tag{10}$$

Notice that this formula implies a higher posterior responsibility for experts with a lower TD-error.

Learning

At the beginning of learning the connection weights of the actor experts w_{AEeji} are randomly generated in $[-0.1, +0.1]$, whereas the connection weights of the critic experts w_{CEei} as well as the connection weights of actor and critic gating networks, respectively w_{AGei} and w_{CGei}, are randomly generated in $[-0.01, +0.01]$. The learning procedure allows to update the values of these weights in order to make the system able to accomplish the three reaching tasks.

Actor gating network learning The learning of the AG has been developed in analogy with the mixture of experts model (Jacobs et al., 1991; Jordan & Jacobs, 1994). Intuitively, the learning rule tends to increase the responsibility of an expert if its likelihood (i.e., the similarity of its action with the executed action) is higher than average (implying $h_{Ae} > g_{Ae}$) and if it has produced a positive TD-error; otherwise it is decreased. Formally:

$$\Delta w_{AGei} = \eta_{AG} \cdot \delta_t \cdot (h_{Ae} - g_{Ae}) \cdot z_{it-1} \tag{11}$$

where η_{AG} is the learning rate (here set to 3.0).

Actor experts learning Filtering the gating outputs with the soft-max favors the quick specialization of the experts. This means that with learning the prior of the best expert will become close to one and those of other experts to zero. In this case the Bayesian rule returns a posterior close to one for the best expert and posteriors close to zero for the remaining experts. Therefore if posteriors are used to modulate the experts' learning rates, as in the mixture of experts systems, it is not possible to create multiple copies of the behavior of the best experts. To solve this issue TERL uses a different learning rule. The soft-max priors g_{Ae} are ranked and the ranks are used to set a *learning rate modulation parameter*, $l_{Ae} = [0.9, 0.8, 0.1, 0.005, 0, 0, 0, 0, 0, 0]$. Note that, as learning modulation parameters do not determine the priors for actions, they do not need to sum up to one. Importantly, this mechanism used for regulating learning is *decoupled* from the priors used to act: this gives much flexibility to TERL because allows the user to establish the number of copies the algorithm develops and the rate with which those copies are trained (see Tommasino et al., 2012, and Tommasino et al., submitted, for more details).

The decoupling between the priors used to mix the experts' action and learning rate modulation parameters can be grounded on brain organisation. The selectors might indeed correspond to high-level areas of brain (e.g., premotor cortex, or even prefrontal cortex, and the related basal ganglia circuits), whereas the experts might correspond to motor areas (again involving cortex and basal ganglia) (Botvinick et al., 2008; Baldassarre et al., 2013). Within this view, the increasingly focussing of the selectors' priors might reflect a progressive learning of the higher areas to select suitable experts, while the learning rate modulation parameters might reflect a tendency of the motor areas to train not only the main expert but also its neighbours in the neural or functional

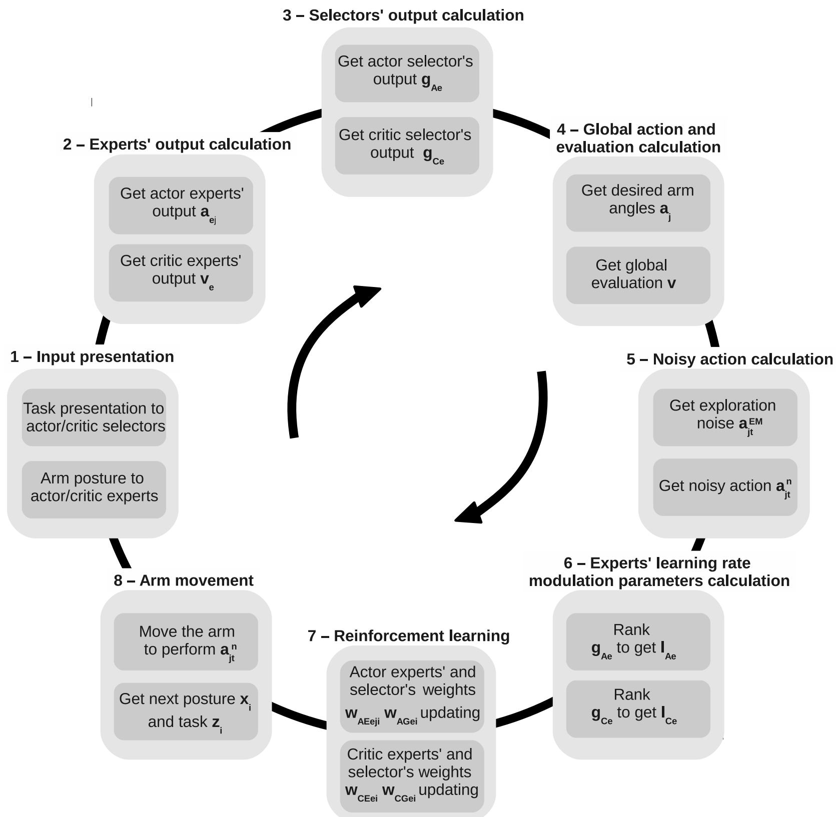

Figure 3. Block diagram showing the information flow between the components of the model during a typical functioning and learning cycle.

space (similarly to what happens in Self Organising Maps, Kohonen, 2001). In this respect, those parameters are now fixed but they might be modulated to best serve the functions of learning, e.g. to balance plasticity/stability based on the learning stage.

The TD(0) learning rule used to train the actor experts and adapted to TERL is:

$$\begin{aligned} e_{AEejit} &= (a^n_{jt} - a_{ejt}) \cdot (a_{ejt} \cdot (1 - a_{ejt})) \cdot x_{it} \\ w_{AEejit} &= w_{AEejit-1} + \eta_{AE} \cdot l_{Ae} \cdot \delta_t \cdot e_{AEejit-1} \end{aligned} \quad (12)$$

where η_{AE} is the learning rate ($\eta_{AE} = 2.0$), and $(a_{ejt} \cdot (1 - a_{ejt}))$ is the derivative of the sigmoid function. This rule implies that the expert action (a_{ejt-1}) gets closer to the noisy performed action (a^n_{jt-1}) if the TD-error (δ_t) is positive, and does so in proportion to the expert rank (l_{Ae}).

Critic gating network learning Similarly to AG, the rule to update the critic gating network was developed on the basis of the mixture-of-experts model (Jacobs et al., 1991; Jordan & Jacobs, 1994): the responsibility of an expert is increased if the expert likelihood is higher (i.e., if its reward prediction error is smaller) than average, that is if $h_{Ce} > g_{Ce}$; otherwise it is decreased. Differently from AG, δ_t is not needed in the formula as the likelihood is already informative of the expert's output quality. Formally:

$$\Delta w_{CGei} = \eta_{CG} \cdot (h_{Ce} - g_{Ce}) \cdot z_{it-1} \quad (13)$$

where η_{CG} is the learning rate ($\eta_{CG} = 1$).

We also introduced a mechanism facilitating the robustness of the system with respect to catastrophic forgetting, particularly severe with sequential tasks involving similar sensorimotor mappings as those considered here. Based on this mechanism, when one responsibility priors of the actor and critic gating networks are greater than a threshold (0.85), then the experts corresponding to them are considered as safely allocated to the task. In particular, when an expert overcomes the threshold the connections weights (of both the actor and critic gating network) linking the output unit with the greatest responsibility prior and the input units related to the other tasks are set to a very low value (-10.0). In this way, the expert corresponding to the high prior is not recruited to solve other tasks and so the skill it stores is protected from interference but continues to be trained for the task to which the expert is dedicated.

Critic experts learning As for the actor we rank the critic priors and obtain the coefficient l_{Ce} to modulate learning rates. The learning rule becomes:

$$w_{CEeit} = w_{CEeit-1} + \eta_{CE} \cdot l_{Ce} \cdot \delta_{et} \cdot x_{it} \quad (14)$$

where η_{CE} is a learning rate ($\eta_{CE} = 0.2$). Note that here the expert TD error δ_{et} is used to update the critics experts instead of the global TD error δ_t.

Exploratory behavior. One important challenge in RL is the regulation of exploratory noise. Different solutions have been proposed for discrete action/state stationary environments (e.g., Gittins & Jones, 1979; Thrun, 1992), but solutions for continuous action/state environments are still preliminary (e.g., see Doya, 2000).

Here we use a noise regulation that exploits the fact that we are interested in episodic RL problems, where learning is based in trials, involving skill transfer as in TRL (Taylor & Stone, 2009). To this purpose, each trial is divided in two phases: a first exploitation phase, with low noise, and a second exploration phase, with high noise. The exploitation phase lasts a time considered sufficient to accomplish the task by a close to optimal system ($1.0s$). After this time, the exploration phase starts. The idea is that if the system has not yet learned the optimal solution then the trial will last beyond the exploitation phase and the system will benefit of the high exploration noise of the exploration phase.

Formally, an *exploratory module* produces stochastic actions obtained by filtering a uniform random noise:

$$a_{jt}^{EM} = \left(1 - \frac{1}{\tau}\right) \cdot a_{jt-1}^{EM} + \frac{1}{\tau} \cdot n_t \quad (15)$$

where $1/\tau = 0.01$ is the filter time constant and n_t is a random variable uniformly distributed in $[-20, +20]$.

The stochastic action is then mixed, through a coefficient c_t, with the global action a_j to obtain the executed action a_{jt}^n:

$$a_{jt}^n = c_t \cdot a_j + (1 - c_t) \cdot a_{jt}^{EM} \quad (16)$$

The parameter c_t is modulated during the exploitation and exploration phase as mentioned above. In particular:

$$c_t = \begin{cases} c_0 & if\ t \leq t_e \\ \beta \cdot c_{t-1} & if\ t_e < t \leq t_T \end{cases} \quad (17)$$

where t_T ($t_T = 8.0s$) is the trial duration, t_e ($t_e = 1.0s$) is the exploitation time during which $c_t = c_0$ ($c_0 = 0.9$), β ($\beta = 0.008$) is a decay coefficient progressively increases noise during the exploration phase. The small noise during the exploitation phase ($c_0 = 0.9$) allows the system to slowly refine the policy even during this phase.

Actions a_{jt}^n are cut within $[0;1]$ and then mapped to the desired angles of the arm.

Results

To understand the model functioning with respect to assimilation and accommodation, we analysed the allocation of the experts during the sequential learning of three reaching tasks. The tasks could require similar sensorimotor mappings (tasks A, B) or different sensorimotor mappings (task C). In particular, we analyzed the dynamics of the output signal supplied by the actor gating network during learning. This signal, indeed, sets the responsibility of experts in action (filtered by a softmax function) and contributes to the entity of their learning (by establishing the learning ranks). In more detail, the priors indicate: (a) which expert has the main responsibility for the selection of the action/evaluation at hand and which are the other experts that contribute to it; (b) which experts are learning a task "in background" (i.e., with a smaller intensity) with respect to the main expert, and so become "copies" of the currently learned skill, available for future exploitation; (c) which expert is used when a new task is introduced, for example an expert "copy" or a completely new expert.

Learning performance

Figure 4 shows the trend of the reward during the learning of the three reaching tasks. From the figure it is evident how the increasing of the reward is faster when the system assimilates (Task B). This because the model is capable of quickly discovering that in order to solve the new task (Task B) it can start from the skill previously acquired to solve Task A. In particular, the model accomplishes Task B starting from an expert copy trained in background during the learning of Task A and this produces a notable advantage on the learning speed. By contrast, when the system accommodates (Task C) the increasing of the reward is less strong as it has to allocate a new expert and train it from scratch.

The curve trajectories performed by TERL to solve the three reaching tasks are shown in Figure 5. The figure illustrates how the model learns to manage the redundant DOFs of the robot arm by developing almost straight trajectories in order to accomplish the three reaching tasks.

The possible neural mechanisms underlying assimilation and accommodation

Figure 6 shows the evolution of the prior responsibilities of actor experts recorded at the end of each trial during the *sequential* learning of the three tasks. At the beginning of the simulation the robot starts to learn Task A for 5000 trials. The actor gating network tests different experts in order to accomplish the task at hand. With the progression of learning the actor gating network gradually selects one expert (the one with the higher responsibility, in this case the expert e3 indicated by the black stripe inFigure 6a to accomplish Task A. At the same time the system gradually trains in background copies of other experts useful to quickly accomplish future same or similar tasks (Figure 6a, gray stripes). After learning Task A the system learns Task B for other 5000 trials. Task B is similar to Task A. As before, at the beginning of the

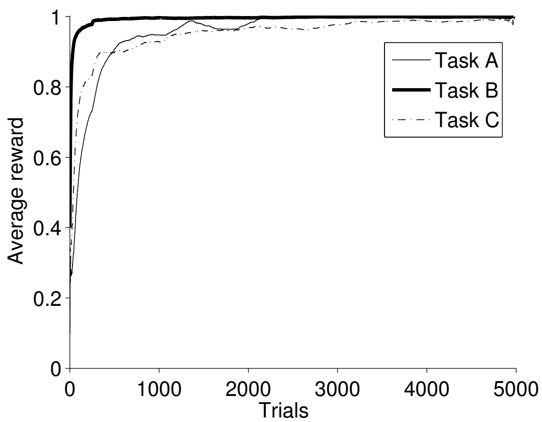

Figure 4. Reward acquired by the model during the sequential learning of the three tasks. The system sequentially learns Task A (thin line), Task B (thick line) and Task C (dot-dash line), each for 5000 trials. Each curve represents the average reward over ten repetition of the experiment calculated at the end of each trial. The curve related to the Task B grows up quickly indicating that the system exploits the knowledge acquired during learning of the previous Task A (experts copies trained in background) to speed up the learning of Task B as the two tasks are similar.

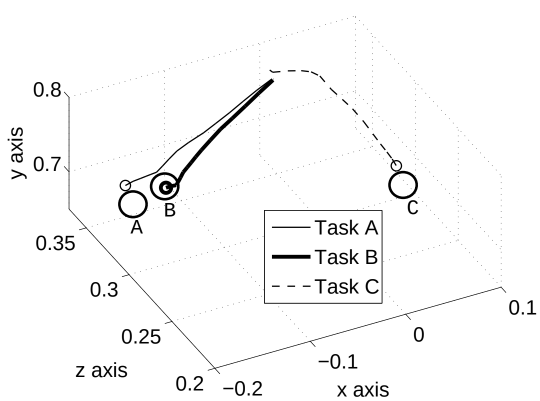

Figure 5. Trajectories showed by the palm of the robotic hand to reach the three objects after learning. The small circle at the end of each curve indicates the end-point of the trajectory; the three big circles represent the three target objects A, B, C. The curves refer to one trained robot.

learning of the new task the actor gating network tests different experts. However, this time the system quickly exploits the knowledge acquired during learning of Task A to speed up the learning of Task B as the two tasks are similar. After few trials the actor gating network recruits a copy expert allocated during the learning of Task A as the expert with the highest prior (e10 inFigure 6b, cf.Figure 6a) for solving Task B. This indicates that the model can learn to allocate experts on the basis of the similarity of the sensorimotor mappings required for solving different tasks. This experiment shows an important case of synergy between assimilation and accommodation processes (see Sec. “Introduction”). A copy of a skill previously used to accomplish Task A is now recruited for the similar Task B and suitably modified to solve a new similar task. Here assimilation manifests in the process of reuse of a copy of a skill previously developed to solve a another task (Task A) to solve a new task (Task B). This allows the system to immediately give a good answer to the new challenge. At the same time, however, the system gradually changes the copy expert to adapt to the requests of the new task, thus manifesting accommodation.

By contrast, when the system learns Task C which is completely different from both the tasks A and B, it recruits a new expert, that is not a copy of the experts recruited for Task A or Task B. This represents a case where assimilation has a very little role (in practice, the system at first uses the best available expert with random weights) while accommodation plays a central role and forms the new skill from scratch (Figure 6c). In this respect, Figure 6 shows that the experts with the three highest priors for tasks A and B, on one side, and those for Task C, on the other side, differ: the system has “understood” that the tasks are radically different, and so it has recruited new experts and created a new functional module.

We chose the $A \Rightarrow B \Rightarrow C$ learning sequence to show the effects of gradually increasing the differences between the sensorimotor mappings the model has in acquiring to solve respectively the tasks A, B, C. If we consider the $A \Rightarrow C \Rightarrow B$ sequence the model solves the Task B (similar to the Task A) again with an expert (e10) belonging to the same functional module of the expert used to solve the Task A. This because the two tasks are similar. In the same line, the expert used to solve the Task C still belongs to a new functional module being very different from the one considered to solve Task A and Task B (Figure 7).

Development of functional modularity

Several recent studies using resting-state functional magnetic resonance fMRI have shown that the human brain functional networks have an intrinsically cohesive modular structure. The modules are mainly composed of functionally and/or anatomically related brain regions in which the connections between regions are denser within each module (He & Evans, 2010; Caeyenberghs et al., 2012). Figure 8 shows an example of this functional modularity of the human brain.

Interestingly, the functional modular organization developed by the model during learning of the three tasks is qualitatively reminiscent of the functional modular organization of the human brain. In this respect, Figure 9 shows the functional graph indicating the emergent functional organization of the experts developed by the model during learning of the three reaching tasks. The nodes of the graph represent the actor experts whereas the links between the nodes are established on the basis of the values of the responsibility priors showed inFigure 6. In particular,

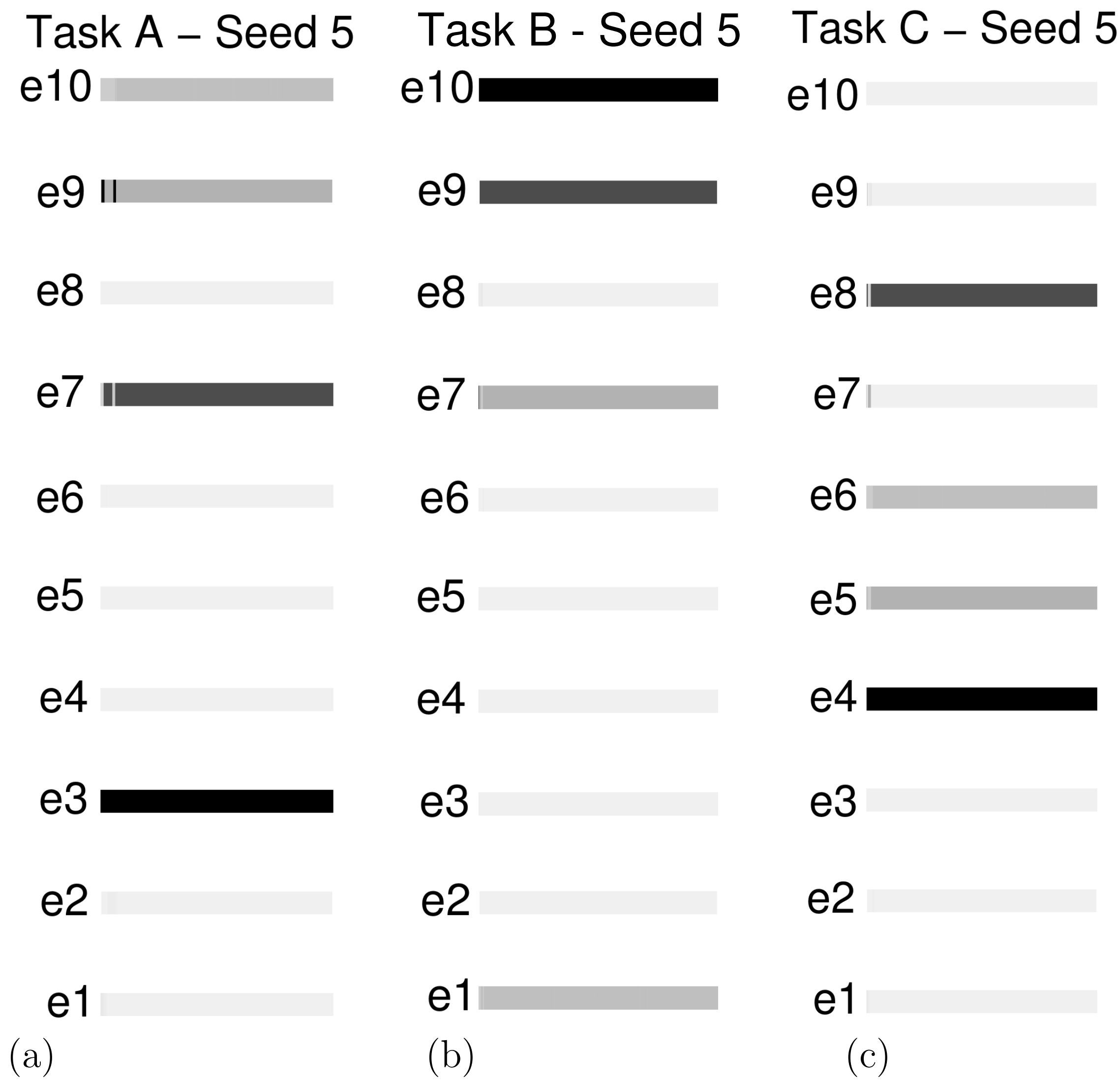

Figure 6. Allocation of actor experts by the model during learning of Task A (a), Task B (b) and Task C (c). Each graph reports the priors of the 10 experts during trials. For each trial of the simulation the highest, second highest, third highest and fourth highest priors are respectively marked with black, dark gray, gray and light gray, while all other priors are not marked (very light gray). The gray tone of the stripes is proportional to the value of the ranked soft-max priors $l_{Ae} = [0.9, 0.8, 0.1, 0.005, 0, 0, 0, 0, 0, 0]$. The black stripes refer to the expert with the higher responsibility chosen to solve the task at hand, while the gray stripes refer to the expert copies trained in background useful to quickly accomplish future same or similar tasks. The data refer to one simulated agent (Seed 5).

for a given task the experts with the four highest priors are linked together. Hence, each module is formed by the expert with the highest responsibility plus the expert copies which learn in background how to solve the task. When the system learns to accomplish Task B which requires a similar sensorimotor mapping with respect to Task A, it chooses as highest prior expert e10, which is a node of the functional module developed during the previous learning of Task A (e3, e7, e9, e10). By contrast, when the system learns to accomplish Task C, requiring a very different sensorimotor mapping with respect to Task A and Task B, it allocates the highest prior expert to a node (node e4) belonging to a completely different functional module (e4, e8, e5, e6).

Note that a comparison between Figure 8 and Figure 9 can be only qualitative as they refer to phenomena taking place at different space scales: the whole brain network in the former, and a possible local network involving motor cortex in the latter. Neuroscience and psychological literature, however, postulates the existence of a "complex posture map" (Graziano, Taylor, & Moore, 2002; Rosenbaum, Loukopoulos, Meulenbroek, Vaughan, & Engelbrecht, 1995), or a "vocabulary of motor acts" (Rizzolatti et al., 1988), within premotor and motor areas, although so far no specific quantitative analysis similar to that of Figure 8 has been given for this finer spatial scale. If this will be done in future work, it

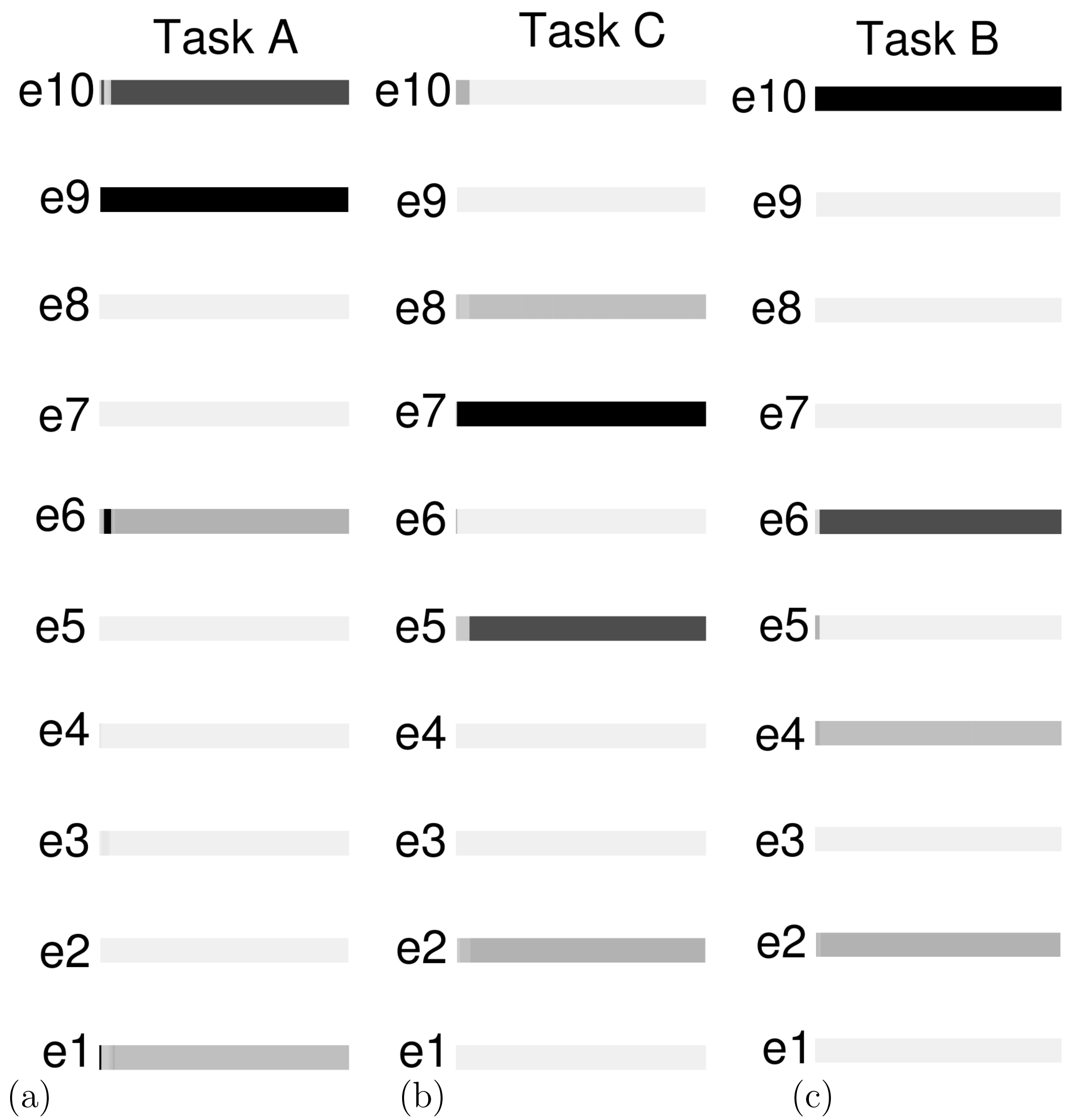

Figure 7. Allocation of actor experts by the model during the sequential learning of Task A (a), Task C (b) and Task B (c). Data plotted as in Figure 6. The figure shows that the system solves Task B by using the expert e10, which is an expert copy trained in background during learning of Task A. The data refer to one trained robot but the results are qualitatively similar if the experiment is replicated with a different random-number generator seed.

will be interesting to evaluate if its organisation qualitatively similar to the one of the model shown in Figure 9.

Impairment of the process leading to the functional modularity

Figure 10 shows the evolution of the prior responsibilities of the actor experts when the mechanism allowing the creation of several expert copies was damaged by setting $l_{Ae} = [1.0, 0, 0, 0, 0, 0, 0, 0, 0, 0]$. In this case the model can allocate only one expert for each task, regardless of the similarity of the sensorimotor mappings required for solving the tasks. As a consequence, the advantages in the learning of Task B (which is similar to Task A) disappears (Figure 11). Using the definitions proposed in Sec. "Introduction" we can say that the damaged system only accommodates.

These results suggest that a damaged system learning sensorimotor skills based only on accommodation processes cannot develop functional modules (i.e., groups of functionally related experts). In the next Sec. "Implications for

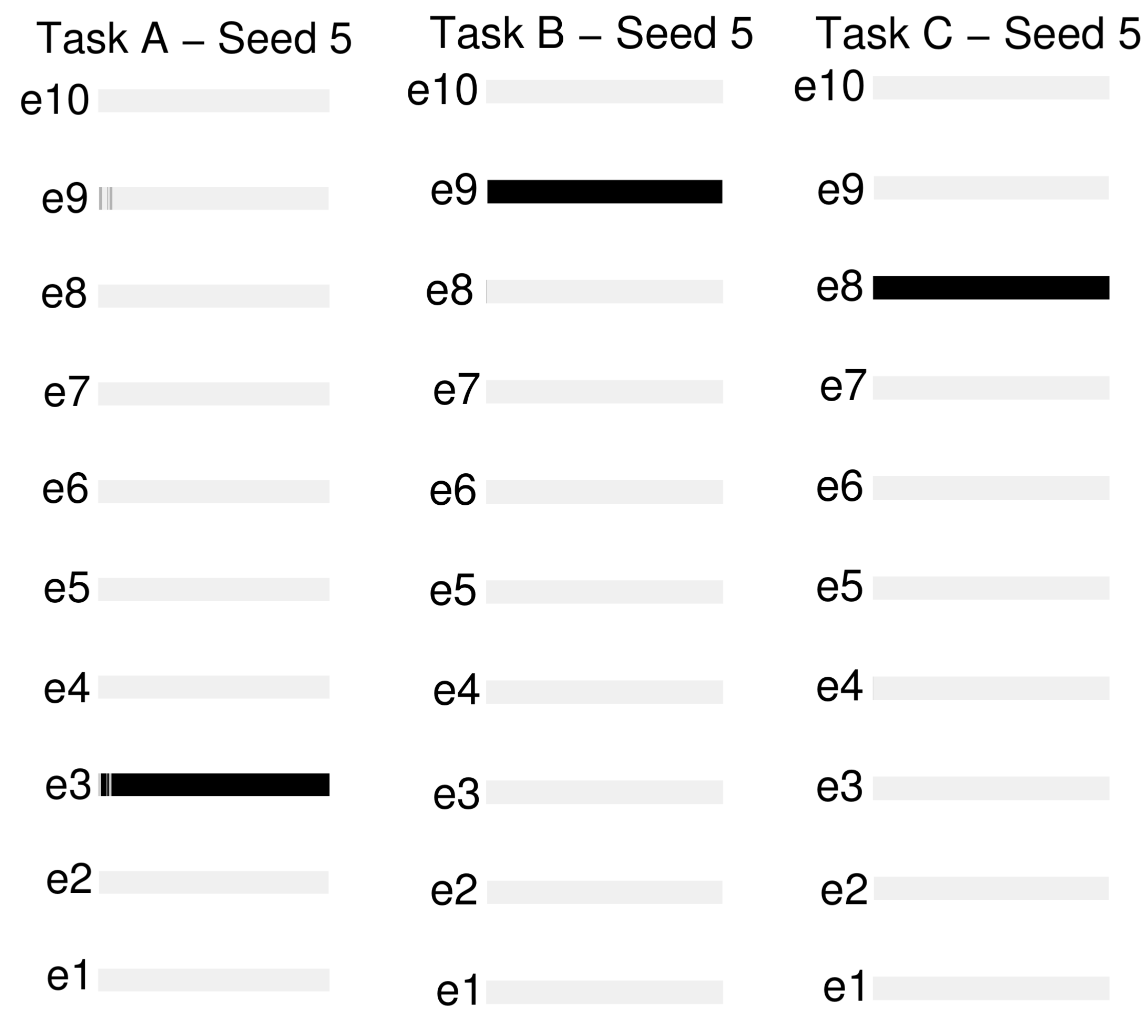

Figure 10. Same data as in Figure 6 but for the damaged model. Notice that in this case there are not gray and light gray stripes as the system cannot train in background expert copies useful to quickly accomplish future same or similar tasks.

autism" we discuss how this kind of deficit might occur in human subjects having impaired abilities to discriminate and generalize motor skills, as in the case of autism.

Discussion

Assimilation and accommodation: neural interpretations

The results shown in Figures 6 and 9 suggest possible computational mechanisms that might underlay assimilation and accommodation. Thus, the experiments showed that when a new task is solved the system is capable of assimilating it to previous similar experiences, so to exploit previously acquired knowledge in the new situation, but also to accommodate the neural structures to a degree that depends on level of novelty of the new challenges. In particular, when the model solves a task similar to some tasks already solved, it recruits copy experts developed in solving those similar tasks (assimilation) and so enlarges the functional module encompassing all those tasks. At the same time, the copy experts are suitably modified (accommodation) to best suit the different features of the new tasks (Figure 6ab and Figure 9). Instead, when the system has to face novel tasks requiring a very different sensorimotor mapping, it recruits non-trained novel experts and so generates a new functional module (Figure 6c and Figure 9). Here assimilation plays a very little role while accommodation plays a very important one. The results on the learning performance (Figure 4) confirms that assimilation occurring in case of similar tasks (Figure 5) allows a fast adaptation thanks to the reuse of a previously trained copy expert then updated to acquire the new behaviour.

The bio-inspired constraints embedded in the TERL model (see Sec. "The TERL model: biological and computational constraints") support the claim that in the brain assim-

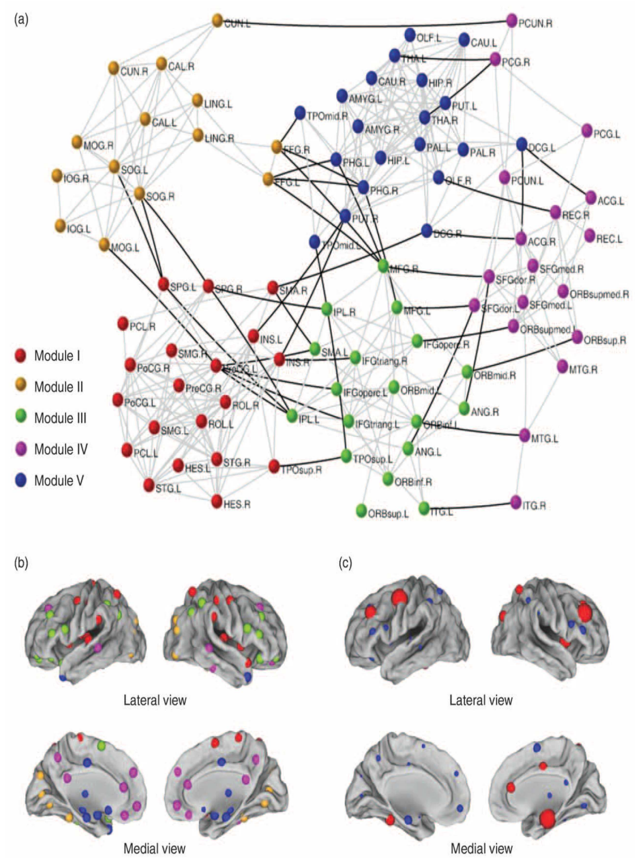

Figure 8. Human brain functional modularity. (a) The functional modular architecture of the human brain. The five functional modules are derived from resting fMRI data and are represented by five different colors. The network is visualized with the Pajek software package (http://vlado.fmf.uni-lj.si/pub/networks/pajek/) using a KamadaKawai layout algorithm. (b) Surface and anatomical representation of the modular architecture of the human brain functional network. The 90 brain regions showed are marked by different color spheres (different colors represent distinct modules) and are further mapped onto the cortical surfaces at the lateral and medial views, respectively, using the Caret software (http://brainvis.wustl.edu). For visualization purposes, the subcortical regions are projected to the medial cortical surface. (c) The global hubs with high topological centralities in the human brain functional networks. The surface visualization of all 90 brain regions is shown, with node sizes indicating their relative node betweenness centrality, Nbc. Regions with high values of Nbc are considered to be hubs (red colors), and otherwise they are considered to be nonhubs (blue colors). Reproduced with permission from He, Y., & Evans, A. (2010). Graph theoretical modeling of brain connectivity. Current Opinion in Neurology, 23, 4, 341350.

ilation and accommodation might pivot on the development of functional modules. These modules could be organized as those shown in Figure 8. In particular, the modular and hierarchical organization of the model could capture general brain design principles representing an important prerequisite to support the development of functional modularity by expert allocation (see Sec. "The TERL model: biolog-

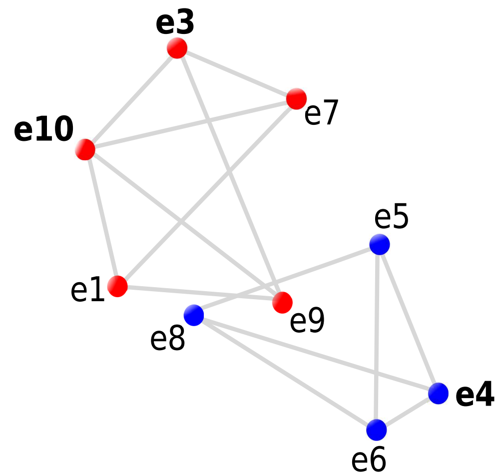

Figure 9. Functional graph illustrating the emerging functional organization of the experts after the model learns the three reaching tasks. The nodes represent the experts while the labels close to each node indicate the expert's name. The name of the highest priors experts performing the three Tasks are indicated in bold. The links between the nodes are established by considering the values of the responsibility priors: the experts with the four highest priors for a given task are linked together. The network is visualized with the Pajek software package (http://vlado.fmf.uni-lj.si/pub/networks/pajek/) using a KamadaKawai layout algorithm.

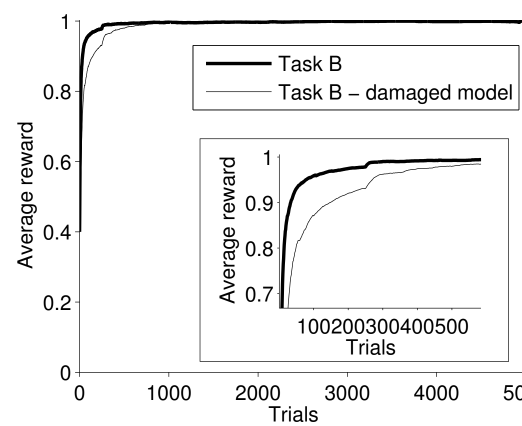

Figure 11. Reward acquired by the damaged model (thin line) and by the not damaged model (thick line) during the learning of the Task B. Both the systems sequentially learn Task A, Task B and Task C, each for 5000 trials. Each curve represents the average reward over ten repetition of the experiment calculated at the end of each trial. The curve related to the damaged model grows up slowly with respect to the curve related to the non damaged one. This trend indicates that the damaged system cannot exploit the knowledge acquired during learning of the previous similar Task A to speed up the learning of Task B because there are no experts copies trained in background. The small box inside the figure shows a zoom of the initial trials where the difference between the two curves is stronger.

ical and computational constraints" and Sec. "Functioning of TERL"). These principles agree with the idea that brain is organized in functional modules each formed by several neural experts (He & Evans, 2010; Caeyenberghs et al., 2012). The explicit link with the brain functional and anatomical modularity is an important step forward with respect to the past computational formulations of Piagetian adaptation processes described in Sec. "Introduction".

In this paper we propose a possible correspondence of assimilation and accommodation to neural processes that differs from the one we proposed in Tommasino et al. (2012). The main idea behind the alternative correspondence proposed in Tommasino et al. (2012) is that assimilation corresponds to the use of the same expert to solve similar tasks, and accommodation corresponds to the re-use and modification of an expert copy to solve rather different tasks. This proposal also considers a third process, called "generation", referred to the use of novel expert to solve a very different task. This proposal was based on the idea that either assimilation or accommodation take over in each situation. Contrary to this, here we proposed the different view for which the two processes are *both present in most cases* that so differ between them for their relative importance.

The view proposed here is summarised in Table 1. The table shows that assimilation is based on mechanisms leading to the re-use of connection weights already used to solve previously solved tasks. Accommodation mechanisms are instead more complex. In general, they involve the update of connection weights to adapt to the new conditions (tasks). However, the updated connection weights can have an increasingly distant relation with the connection weights used to solve previous tasks. Thus, in the basic case a minimal form of accommodation happens when the system updates the connection weights of the best expert used to solve a given task to accommodate some slightly different conditions. This process might also take place in another interesting condition not considered here for the sake of clarity. Both here and in Tommasino et al. (2012) we assumed that the experts are not given any information about the pursued goal. If this assumption is relaxed, and information on the goal (e.g., described in terms of features) is given to the experts, then the *same* expert might acquire different sensorimotor mappings depending on the goal, e.g. it might perform slightly different reaching actions for targets located in close positions. In this case "accommodation" could involve the same experts used to solve previous tasks rather than their copies. Another case of accommodation involving more substantial changes involves the solution of tasks requiring a more different sensorimotor mapping. In this case the system might use and update copy experts to solve the new tasks so to exploit previously acquired knowledge but also avoid to disrupt it. A last case involves the solution of tasks requiring sensorimotor mappings so different from the acquired ones that previously acquired knowledge is not useful, or even deleterious, to solve them. In this case the system accommodates by starting a new expert and a new functional module. Below we expand the strengths of the view proposed here to capture the assimilation and accommodation processes as described by Piaget.

With respect to assimilation, Piaget writes: "[...] repetition of the reflex leads to a general and generalizing assimilation of objects to its activity, but, due to the varieties which gradually enter this activity (sucking for its own sake, to stave off hunger, to eat, etc.), the schema of assimilation becomes differentiated and, in the most important differentiated cases, assimilation becomes recognitory. In conclusion, assimilation belonging to the adaptation reflex appears in three forms: cumulative repetition, generalization of the activity with incorporation of new objects to it, and finally, motor recognition. But, in the last analysis, these three forms are but one: The reflex must be conceived as an organized totality whose nature is to preserve itself by functioning and consequently to function sooner or later for its own sake (repetition) while incorporating into itself objects propitious to this functioning (generalized assimilation) and discerning situations necessary to certain special modes of its activity (motor recognition)." (citeNPPiaget1953OriginsIntelligenceChildren2034688958, p. 37).

Let us consider how the hypothesis proposed here captures the two basic aspects of assimilation, namely "cumulative repetition" and "generalized assimilation". The process of repetition is characterized by Piaget with these words: "[...] reflex is consolidated and strengthened by virtue of its own functioning. Such a fact is the most direct expression of the mechanism of assimilation." (Piaget, 1953, p. 32). Cumulative repetition might be related to motivational mechanisms having the adaptive value of assuring a refinement of the acquired sensorimotor schema, in particular mechanisms such as *intrinsic motivations* that drive the child to interact with the world to acquired knowledge and skills useful in later stages of life (Singh, Lewis, Barto, & Sorg, 2010; Baldassare, 2011; Baldassare & Mirolli, 2013). Although the generation of such motivation is not explicitly captured by TERL (but this might be done in the future, see Sec. "Conclusions and future work"), the model focusses on each task for several "trials". Importantly, this repetition is functional not only to the improvement of the model skill, but also to the gradual formation of background expert copies which are a prerequisite for the following assimilation and accommodation processes. In this respect, the model furnishes a concrete specific interpretation of Piaget's intuition of a strong relation between repetition and assimilation.

Piaget clarifies "generalised assimilation" by referring to the example of sucking: "We simply maintain that, without any awareness of individual objects or of general laws, the newborn child at once incorporates into the global schema of sucking a number of increasingly varied objects, whence the generalizing aspect of this process of assimilation." (Piaget, 1953, p. 34). According to the interpretation proposed here, assimilation occurs either when the model recruits the same expert to solve two tasks requiring an identical sensorimotor mapping and when the model recruits a copy of an already trained expert to solve a similar task. In both cases the system uses the same functional module. In this case the sys-

Table 1
Assimilation and accommodation mechanisms operating with new tasks involving sensorimotor mappings with a different degree of similarity with respect to already solved tasks.

Degree of similarity with previous tasks	Expert used to solve new task	Assimilation mechanisms	Accommodation mechanisms
Different sensorimotor mapping	Expert from new functional module: novel expert	—	Start and update a new functional module
Related sensorimotor mapping	Expert from previous functional module: copy expert	Use of a copy of connection weights	Update of a copy expert of same functional module
Similar sensorimotor mapping	Expert from previous functional module: same changed expert	Use of slightly updated connection weights	Update of a previously developed expert
Same sensorimotor mapping	Expert from previous functional module: same expert	Use of same connection weights	—

tem assimilates in the sense that it incorporates into the same functional module different objects by re-using connection weights developed in the past to face the new situation.

Piaget's characterizes accommodation as follows: "Concerning its adaptation, it is interesting to note that the reflex, no matter how well endowed with hereditary physiological mechanism, and no matter how stable its automatization, nevertheless needs to be used in order to truly to adapt itself, and that it is capable of gradual accommodation to external reality. [...] it sometimes happens that the child does not adapt at the first attempt. Only practice will lead to normal functioning. That is the first aspect of accommodation: contact with the object modifies, in a way, the activity of the reflex, [...]" (Piaget, 1953, p. 30). This definition of accommodation captures its important aspect involving the modification of existing structures (the "reflex"). The model presented here allowed the articulation of the mechanisms implementing such modification by specifying possible changes involving the neural structures of the system having an increasing impact on the system. The appropriateness of the mechanisms proposed here for accommodation could only be established on the basis of detailed evidence from real brain. In particular, this evidence might show how learning a new different task might cause the modification of existing experts, or their copies (and if copies are actually formed), or the recruitment of unrelated neural structures to start a new functional module.

A observation is relevant regarding the latter mechanism involving the generation of a new functional module. This observation is prompted by the finding on brain functioning showing that primate motor cortex contains a topographical map of densely packed behaviourally relevant actions clustered in the neural space according to their behavioural similarity (Meier et al., 2008). This suggests a full employment of neural resources since the beginning of the learning process. The observation concerns the model rank-based learning rates of experts, learning rates that here are fixed. These learning rates are very important as they establish the number of experts involved in the learning processes, and the size of such processes. Thanks to the decoupling between functioning and learning of TERL, in the future it might be possible to modulate the learning rates depending on the advancement of learning, analogously to what happens in Self Organising Map where the number and entity of learning of neighbouring units decreases with the progression of learning (SOM, Kohonen, 2001). If one assumes initial learning rates involving most experts (as in SOMs), after few learning cycles there would be no more "free experts" to recruit for the creation of new functional modules. In this condition, new functional modules could be created only by subtracting "peripheral" neural resources/experts (i.e. resources with low use) from other functional modules.

Note that here we focused only on reaching behaviours. Thus, the accommodation process considered involved sensorimotor mappings that differed only in terms of the location of the target of reaching. Instead, Piagetian accommodation is often referred to more important changes leading to the construction of qualitatively different schemas, and possibly involving domains different from motor behaviour. Notwithstanding this are important differences, we expect the neural principles explained here would scale up to account for more complex forms of accommodation processes. In particular, extending the model to produce actions different from reaching would mainly involve empowering the experts used in the model to implement more complex motor behaviours. Indeed, actions with any type of complexity and variability can ultimately be performed by generating suitable joint trajectories or desired joint postures. Accounting for different domains beyond the motor one would require representation devices different from those used here. However, also in this case the mechanisms used here are expected to some extent to continue to function, in particular the assimilation of new mental contents (e.g., representations of categories or predictions) into the same neural structures, the formation of "background copies", and the re-use of such copies to more quickly learn new mental contents.

Implications for autism

The results shown in Figure 10 suggest that a system with an impairment in the mechanism allowing the creation of several expert copies cannot develop functional modules. As a consequence, the impaired system can learn sensorimotor skills only based on accommodation processes and cannot generalize to tasks requiring similar sensorimotor mappings (Figure 11). We propose the hypothesis that autistic subjects might show similar deficits. This claim agrees with some studies suggesting that autistic brain may show an early divergence between assimilation and accommodation functions, with the latter progressing much further than the former (Morgan, 1986; Burack et al., 2001). Remarkably, the deficit in the development of functional modules suggested by the model is also in line with recent brain imaging data claiming that processes of modularization might be disrupted in autism (Boersma et al., 2013; Catarino et al., 2013; Meunier et al., 2010). The overall idea we propose here with the model is that autistic subjects tend to acquire multiple pieces of knowledge (here sensorimotor skills) in a segregated, non-integrated fashion, as if they had little or no relation between them. Knowledge on similarity and structure linking difference experiences is hence lost.

Figure 10 refers to a model where the mechanism allowing the creation of several expert copies is completely damaged. This is an extreme case used here to emphasize the difference between damaged and not-damaged behaviours. This does not take into account the heterogeneous profiles of real autistic subjects who could show a certain degree of assimilation processes alongside the strong accommodation ones. We notice, however, that the model could account for the partial functioning of assimilation processes in real subjects by partially (rather than fully) damaging the mechanisms supporting the creation of expert copies (indeed, recall that the number and learning rates of background copies can be regulated in the model). This regulation might allow accounting for inter-subject differences.

Finally, the computational approach proposed in this paper to study assimilation/accommodation and functional modularity in autism support the claim that artificial neural networks can be powerful theoretical tools to address some issues related to autism (Grossberg & Seidman, 2006). In this respect, these models have a great potential for studying the effects of specific abnormalities in sensory stages (Thomas, Knowland, & Karmiloff-Smith, 2011) and in motor computations (Gowen & Hamilton, 2012), or to explain how under/over innervated networks in autism subjects may affect their abilities to discriminate and generalize (Grossberg & Seidman, 2006; Thomas et al., 2011; Gowen & Hamilton, 2012).

Related works

In the literature various computational models have been proposed to capture important mechanisms related to assimilation and accommodation. Sec. "Introduction" already mentioned models where the changes in the neural network weights represent a form of accommodation whereas the transformation (by the network weights) of input patterns into internal patterns of activation corresponds to assimilation (Mareschal, 2003; Parisi & Schlesinger, 2002; Rasheed & Ali, 2009). Other computational accounts suggested that Piagetian adaptation processes are the expression of the intrinsic dynamics of an adaptive system (Van Geert, 1998). This "dynamic system approach" (Thelen & Smith, 1994) conceives assimilation and accommodation as a result of a self-organizational process based on a dualism between conservative (assimilation) and progressive (accommodation) forces (Van Geert, 1998).

Some key ideas of the "neuronal groups selection theory" proposed by Edelman (1987) are relevant for the issues addressed here. According to Edelman's perspective at birth the brain is formed by a redundant multitude of competing neural groups "selected" (preserved) on the basis of their overall activity. Different neural groups reach high levels of specialization in processing different classes of stimuli and the selection process leads to the progressive death of groups that fail to specialise. The concept of "neural group" has similarities with the one of "functional module of experts" of TERL but there are also important differences between the two. In particular, neural groups are mainly defined at the anatomical level as neurons of the same group are anatomically linked, whereas experts in TERL are tied by functional links as experts of the same functional module are those that tend to implement similar sensorimotor mappings. In addition, the selection of neural groups is mainly based on self-organizing processes based on their specialisation to process different types of inputs stimuli, whereas the formation and selection of the functional modules in TERL is based on the operation of the selector networks and is mainly based on the capacity of the experts to generate appropriate actions, or action evaluations, to best respond to different environmental challenges. Thus, group selection is mainly stimulus oriented whereas TERL expert selection is mainly action oriented.

Drescher (1991) proposed a sophisticated architecture directly inspired by Piaget's theory. The architecture was used to produce behaviour of a simulated agent endowed with a body through which it interacted with a simulated environment. The notion of *schema* employed by Piaget is pivotal for Drescher's architecture. A schema is formalized as a triplet formed by *context*, *action*, and *result*. The context consists of a Boolean combination of propositions on the state of the world and that can be either true or false. Initially, the simulated agent has a few schemas whose context and result fields are empty. The schema mechanism is equipped with a learning process referred to as *marginal attribution*. Using this process, the agent builds a model of the environment by learning the effects of different actions (the "results") in different contexts. An agent enters the environment with elementary initial schemas. Once learned the consequences of own actions, the agent can use its schemas to accomplish desired goals. Through the process of adaptation the agent builds a meta-level of schemas based on the existing schemas. Similarly to TERL, Drescher's formal proposal for novel concept formation highlights the importance of the agent's interactions with the environment for the system development. However, Drescher's architecture has aspects that would qualify it as a "symbolic" system. Instead, the computational approach proposed by TERL emphasises the continuous nature of perception and action in organisms, an element also at the basis of its incremental "sub-symbolic" learning processes.

Tani and Nolfi (1999) proposed a model based on a hierarchy of recurrent neural-network experts. The system experienced a sensorimotor flow of information collected by navigating in different rooms. Modules of lower levels learned and specialised to anticipate information at a fine time and spatial scale. Modules of higher levels learned and specialised to anticipate the sequence of activations of the lower level modules, thus encoding information at a higher level of abstraction. The authors gave an example of this abstraction process with a simulated robot that had to distinguish between two rooms. After spending some time in room A the lower level of the hierarchical architecture learned primitive concepts like corridors, corners, and crossings. The higher level then learned to distinguish between room A and room B by relying on regularities in the low-level activation sequence. As in the approach proposed here, Tani and Nolfi (1999) used mixtures of neural modules. However, their model focused on prediction learning of the sensorimotor flow rather than on autonomous action learning processes as in TERL. Thus, experts were formed by recurrent neural networks learning by the error back-propagation through time algorithm (Rumelhart, Hinton, & Williams, 1986), whereas TERL experts are formed by feedforward neural networks learning through reinforcement learning.

Nishimoto and Tani (2009) used a multiple time-scales recurrent neural network to control a humanoid robot involved in the learning of multiple goal-directed tasks consisting in manipulating objects under the experimenter supervision. The architecture of the model consisted of input-output and context-output units. The input units received the current proprioceptive and visual signals whereas the output units supplied the proprioceptive and visual signals for the next step. The context units were divided into fast units whose activity changed quickly, and slow units whose activity, in contrast, changed much more slowly. Depending on the top-down signal conveying the goal information flowing from the upstream slow context units, different sensorimotor mappings (called by the authors "behaviour primitives") were adopted which could explain dynamic mechanism of assimilation. The behaviour primitives were the products of the neuronal self-organization with having rich sensorimotor interactions through interactive tutoring. This may account for accommodation, that was understood as recruitment of new fast context units. This perspective is similar to what happens in the TERL model where accommodation implies the recruitment of a different functional module. An important difference between TERL and the model proposed by Nishimoto and Tani (2009) is that the latter is trained with a supervised learning procedure (the experimenter guides the robot's hand along the trajectory of the goals action) whereas our approach is based on a reinforcement learning.

Sugimoto et al. (2012) proposed a modular system based on the idea of the MOSAIC (Model for Sensorimotor Learning and Control) models (e.g., see Haruno, Wolpert, & Kawato, 2001). Here experts are formed by pairs of forward (predictor) and inverse (controller) models, and the selection of experts that are mixed and perform the action is the result of a competition between them based on the prediction error of the forward models: as a low error of a predictor implies a high competence of the related controller for the condition at hand, the experts with low errors receive a higher responsibility in issuing commands to the controlled plant. Sugimoto et al. (2012) enhance this mechanism integrating in the experts' responsibility signals the task-related rewards, thus making the system closer to the approach followed with TERL. The main difference between the two models is however that TERL has been directly designed to decide which skills to use to face novel tasks fully based on the rewards, i.e. on information on how well the already acquired skills perform in the new conditions. This makes TERL more suitable to study assimilation and accommodation processes.

A modular RL system capable of solving different tasks and focused on transferring knowledge between goals was also proposed by Castro da Silva, Baldassarre, Konidaris, and Barto (2014) (see also Castro da Silva, Konidaris, & Barto, 2012). This system solves multiple tasks and directly learns to map information on the task goals to the parameters of parameterised policies (*Dynamic Movement Primitives*, Schaal, Peters, Nakanishi, & Ijspeert, 2005) learned to solve them with a policy search method (Kober & Peters, 2009). When the system encounters a new task defined in terms of a new goal, e.g. a new target position to hit by throwing a ball to it with a robotic arm, it can immediately formulate an initial policy to solve them by mapping the goal to the parameters of the policy. A further RL process can then refine such initial policy. This approach is similar to the one proposed here in that it explicitly investigates how to transfer knowledge between different tasks based on a modular system. However, while this system is fully based on the information on the task goal (it migth be said that assimilation is here based on the similarity between goals), TERL is capable of working without such information and to identify the experts relevant to solve the new tasks on the basis of only the information of how they work in such new task.

Another system proposed by Baldassarre et al. (2012) (see also Taffoni et al., 2013) was based on a trial-and-error learning hierarchical architecture strongly constrained on the basis of the macro anatomy of basal ganglia and cortex in brain. The system was capable of controlling a simulated iCub engaged in learning, through intrinsic motivations, the actions that could be performed on a responsive board ("mechatronic board", Taffoni et al., 2012a) in order to produce "interesting" events (e.g., pressing a button to cause a light switching on, Taffoni et al., 2013). The system is also capable of recalling skills on the basis of "goals" relate to them. However, it is not capable of skill transfer as TERL, and hence cannot capture assimilation/accommodation processes.

A hierarchical reinforcement learning system was also proposed by Ciancio, Zollo, Guglielmelli, Caligiore, and Baldassarre (2011) (see also Ciancio, Zollo, Baldassare, Caligiore, & Guglielmelli, 2013). As in TERL, this system is based on a hierarchical actor-critic architecture formed by two levels. A first lower level is formed by experts that learn to generate suitable parameters of *central pattern generators* which in turn generate the commands sent to the hand of the iCub robot engaged in a fine manipulation behaviours (turning a cylinder as fast as possible based on rhythmic finger movements). A second higher level learns to decide the mixture of experts to be used to solve the task. This system is modular and hierarchical and learns by RL as TERL. However, differently from TERL, it cannot re-use acquired skills to improve the learning of new similar skills and hence cannot be used to model assimilation and accommodation.

A two-layer hierarchical reinforcement learning system based on the actor-critic model was also proposed by Schembri, Mirolli, and Baldassarre (2007a, 2007b, 2007c). This system controlled a simulated robot navigating on a ground coloured with different patterns seen trough a simplified camera. In a first phase of life of the system ("childhood") a lower level of experts learned by RL to navigate in the environment to maximise the achievement of rewards produced by "reinforcers" evolved with a genetic algorithm. In this phase, the higher level ("selector") learned to select the experts to miximise such "intrinsic" rewards. In a second phase of life of the system ("adulthood") the selector learned to select the experts in order to accomplish some "extrinsically rewarded" targets (''foods"). The amount of food collected during the adulthood was used as a fitness function to guide the genetic algorithm searching the reinforcers and guiding learning in childhood. This system shares with TERL the use of a hierarchical RL architecture and also the learning of multiple tasks. The system can also re-use and compose acquired skills to solve different tasks. However, the system cannot exploit already acquired skills to boost the learning of similar sensorimotor mappings as TERL, the key feature of accommodation.

Modularity and the capacity to balance plasticity/stability has also been investigated in unsupervised and category learning systems. In this respect, the ART (Adaptive Resonance Theory) neural networks represent a family of models developed especially for unsupervised learning (Carpenter & Grossberg, 2010). Abstracting from details, in ART networks each input pattern, formed by feature vector, activates some recognition units in proportion to the similarity of the features with the recognition unit connection weights (as in self-organising maps). Additionally, recognition units inhibit each other giving rise to a lateral competition. If the most active recognition unit overcomes a threshold ("vigilance parameter"), its connection weights are updated to become more similar to the input features. If this does not happen, the unit is switched off (hence the other units will activate more as released by its inhibition) and the next maximally active unit is checked to see if the threshold is overcome. If this search fails, a new "non-committed" unit is recruited to represent the new input pattern. The two learning processes, involving either integration of the new input into the connection weights of already committed category units or the recruitment of new category units, can thus be considered to correspond to respectively assimilation and accommodation processes. The cascade-correlation networks (Fahlman & Lebiere, 1990; Shultz, Schmidt, Buckingham, & Mareschal, 1995; Shultz, 2003) have some similarities with the ART models but are focussed on supervised learning. Similarly to ART systems, the nets can "grow", thus capturing the processes of synaptogenesis and neurogenesis, in particular can recruit new hidden units. The nets also use a novel learning algorithm called "quick-prop" (Fahlman, 1988), which learns faster than traditional error back-propagation algorithms (Rumelhart et al., 1986). The learning of new patterns based on existing units can be considered a process of assimilation. The recruitment of new units can be instead seen as a process of accommodation. The unsupervised and supervised learning techniques used and the growing process of the network structure adopted in the ART models and cascade-correlation nets are the main differences with respect to TERL. In particular, being focussed on reinforcement learning, TERL captures assimilation and accommodation as strongly linked to the sensorimotor interactions with the environment, an important aspect of Piaget's theory.

Conclusions and future work

This paper proposes that Piagetian assimilation and accommodation processes are pivotal for brain functional modularity, the expression of which is facilitated by the brain's modular and hierarchical organization. This claim is supported by running computer simulations using the bio-inspired computational model TERL whose architectural and functioning organization captures some essential aspects of the modular and hierarchical arrangement of the brain. The model drives a simulated humanoid robot to autonomously learn by trial-and-error how to accomplish different reaching tasks. The model decides to assimilate or accommodate according to the degree of similarity between the tasks.

Remarkably, the model helps to build a new theoretical framework to study the possible consequences of damaging the computational mechanisms underlying assimilation and accommodation. In particular, the capacity of the system to accommodate can be "impaired" and as a consequence the system loses its capacity to generalize to tasks requiring similar sensorimotor mappings. We suggest that a similar damage might occur in autistic subjects as these exhibit a difficulty to discriminate and generalize motor behaviours.

The model used here aims at capturing general features of the modular and hierarchical organization of the brain architecture and functioning reproducing typical processes observed during the development of behaviour in children. This approach in part fulfils the *Computational Embodied Neuroscience* (CEN) method (Caligiore, Borghi, Parisi, & Baldassarre, 2010; Mannella, Mirolli, & Baldassarre, 2010; Caligiore & Fischer, 2013) suggesting the importance of developing general system-level models that incorporate constraints from different sources. The CEN approach suggests several types of constraints to make models cumulative: the constraint of reproducing behaviours as measured in several different psychological experiments; the constraint of reproducing the learning of behaviour alongside the final behaviour; the constraint of using architectures and algorithms informed by neuroscientific evidence; and the constraint for which the model should be able to exhibit its behaviour within an embodied agent reproducing the actual circular interactions with the environment of the participants of the target experiments. In the long run the fulfillment of these constraints has the advantage of leading to the progressive isolation of general principles underlying the class of studied phenomena, thereby fostering theoretical cumulativity (see Caligiore, Borghi, et al., 2010, and Mannella et al., 2010, for the application of this method to the study of phenomena different from reaching). Although the model used here does not fully follow the CEN methodology (e.g., it incorporates few neuroscientific and embodiment constraints and reproduces only the learning of the reaching behaviour at a qualitative level), the constraints it incorporates on modular and hierarchical organization are very important to achieve the results presented here.

In future work, further improvements might be introduced in the model architecture, in particular a more realistic visual system and the control of the hand joints. This could allow the study of assimilation and accommodation processes in the acquisition of other motor behaviours such as grasping and the acquisition of other more abstract forms of cognitive processes. This would also open up the interesting possibility of studying how accommodation and assimilation, possibly supported by a suitable motivational guidance of learning based on intrinsic motivations (Barto, Singh, & Chentanez, 2004; Oudeyer, Kaplan, & Hafner, 2007; Baldassare, 2011; Baldassare & Mirolli, 2013), could support a genuine autonomous open-ended development.

Another feature of the model architecture that could be improved involves the goal representation. For the sake of clarity of the results, the goal (or task) pursued at the moment by the system, and sent to the selectors as input, was here abstracted with a simple vector. A more sophisticated representation of goals would instead be very important to support open-ended development as goals are fundamental pivots of behaviour and learning since the very early stages of development (von Hofsten, 2004). Thus, in future work the input to the selectors might become more sophisticated, e.g. it could be based on a visual representation (at a suitable level of abstraction) of the goal that the system is pursuing. This would allow the selectors to have a bias to select specific experts to solve the new tasks, a bias based on the *similarity of the goals* of the new tasks with those of previously solved tasks. This initial bias would then be strengthened or overridden on the basis of the actual capacity of the experts to solve the new tasks. This mechanism would represent an improvement with respect to the current system that, when facing new tasks, assigns "flat priors" of selection, i.e. equal responsibilities, to all experts.

In the future, the biological plausibility of the model could also be enhanced by capturing the macro-structure of hierarchical brain, in particular the organisation of basal ganglia-cortical sub-systems underlying goal-directed and habitual behaviour, and the control of manipulation and overt attention. This might be done on the basis of the mechanisms and principles captured by the models proposed by Baldassarre et al. (2012) and Chersi, Mirolli, Pezzulo, and Baldassare (2013). This could allow the account of other phenomena related to autism, for example the dysfunctions in some visual (R. T. Schultz et al., 2000; Klin, Jones, Schultz, Volkmar, & Cohen, 2002) and attention (Ames & Fletcher-Watson, 2010) behaviours, or the impaired chaining of different motor acts (Cattaneo et al., 2007).

References

Aflalo, T. N., & Graziano, M. S. A. (2011). Organization of the macaque extrastriate visual cortex re-examined using the principle of spatial continuity of function. *Journal of Neurophysiology*, *105*, 305-320.

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, *9*, 357-381.

Ames, C., & Fletcher-Watson, S. (2010). A review of methods in the study of attention in autism. *Developmental Review*, *30*, 52-73.

Baldassare, G. (2011). What are intrinsic motivations? a biological perspective. In A. Cangelosi et al. (Eds.), *Proceedings of the International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob-2011)* (p. E1-8). Piscataway, NJ: IEEE.

Baldassare, G., & Mirolli, M. (Eds.). (2013). *Intrinsically motivated learning in natural and artificial systems*. Berlin: Springer. doi: 10.1007/978-3-642-32375-1

Baldassarre, G. (2002). A modular neural-network model of the basal ganglia's role in learning and selecting motor behaviours. *Journal of Cognitive Systems Research*, *3*, 5-13.

Baldassarre, G., Caligiore, D., & Mannella, F. (2013). The hierarchical organisation of cortical and basal-ganglia systems: a computationally-informed review and integrated hypothesis. In G. Baldassarre & M. Mirolli (Eds.), *Computational and robotic models of the hierarchical organisation of behaviour* (p. 237-270). Berlin: Springer-Verlag.

Baldassarre, G., Mannella, F., Fiore, V. G., Redgrave, P., Gurney, K., & Mirolli, M. (2012). Intrinsically motivated action-outcome learning and goal-based action recall: A system-level bio-constrained computational model. *Neural Networks*, *41*, 168-187.

Barto, A. G., Singh, S., & Chentanez, N. (2004). Intrinsically motivated learning of hierarchical collections of skills. In J. Triesch & T. Jebara (Eds.), *Proceeding of the International Conference on Developmental Learning (ICDL2004)* (p. 112-119). Piscataway, NJ: IEEE.

Berthier, N. E. (1996). Learning to reach: A mathematical model. *Developmental Psychology*, *32*, 811-823.

Berthier, N. E., & Carrico, R. L. (2010). Visual information and object size in infant reaching. *Infant Behavior and Development*, *33*, 555-566.

Berthier, N. E., Rosenstein, M. T., & Barto, A. G. (2005). Approximate optimal control as a model for motor learning. *Psychological Review*, *112*, 329-346.

Blaesi, S., & Wilson, M. (2010). The mirror reflects both ways: Action influences perception of others. *Brain and Cognition*, *72*, 306-309.

Boden, M. A. (1994). *Piaget*. London: Fontana Press.

Boersma, M., Kemner, C., Reus de, M. A., Collin, G., Snijders, T. M., Hofman, D., … Heuvel van den, M. P. (2013). Disrupted functional brain networks in autistic toddlers. *Brain Connectivity*, *3*, 41-49.

Botvinick, M. M., Niv, Y., & Barto, A. (2008). Hierarchically organized behavior and its neural foundations: A reinforcement-learning perspective. *Cognition*, *113*, 262-280.

Burack, J. A., Charman, T., Yurmiya, N., & Zelazo, P. R. (2001). Development and autism: Messages from developmental psychopathology. In N. Y. . P. R. Z. J. A. Burack T. Charman (Ed.), (p. 3-16). Mahwah, NJ: Lawrence Erlbaum.

Caeyenberghs, K., Leemans, A., Heitger, M. H., Leunissen, I., Dhollander, T., Sunaert, S., … Swinnen, S. P. (2012). Graph analysis of functional brain networks for cognitive control of action in traumatic brain injury. *Brain: A journal of neurology*, *135*, 1293-307.

Calabretta, R., Di Ferdinando, A., Wagner, G. P., & Parisi, D. (2003). What does it take to evolve behaviorally complex organisms? *Biosystems*, *69*(2-3), 245-262.

Caligiore, D., Borghi, A. M., Parisi, D., & Baldassarre, G. (2010). TRoPICALS: A computational embodied neuroscience model of compatibility effects. *Psychological Review*, *117*, 1188-1228.

Caligiore, D., & Fischer, M. H. (2013). Vision, action and language unified through embodiment. *Psychological Research*, *77*, 1-6.

Caligiore, D., Guglielmelli, E., Parisi, D., & Baldassarre, G. (2010). A reinforcement learning model of reaching integrating kinematic and dynamic control in a simulated arm robot. In *IEEE International Conference on Development and Learning (ICDL2010)* (p. 211-218). Piscataway, NJ: IEEE.

Caligiore, D., Mirolli, M., Parisi, D., & Baldassarre, G. (2010). A bioinspired hierarchical reinforcement learning architecture for modeling learning of multiple skills with continuous state and actions. In S. E. In Johansson Birger & B. Christian (Eds.), *Proceedings of the Tenth International Conference on Epigenetic Robotics (EpiRob2010)* (p. 27-34). Lund University Cognitive Studies, 149. Sweden.

Caligiore, D., Parisi, D., & Baldassarre, G. (2014). Integrating reinforcement learning, equilibrium points and minimum variance to understand the development of reaching: a computational model. *Psychological Review*, *121*(3), 389-421.

Caligiore, D., Pezzulo, G., Miall, R. C., & Baldassarre, G. (2013). The contribution of brain sub-cortical loops in the expression and acquisition of action understanding abilities. *Neuroscience and Biobehavioral Reviews*, *37*, 2504-2515.

Carpenter, G. A., & Grossberg, S. (2010). *Adaptive resonance theory*. Springer.

Carrico, R. L., & Berthier, N. E. (2008). Vision and precision reaching in 15-month-old infants. *Infant Behavior and Development*, *31*, 62-70.

Castro da Silva, B., Baldassarre, G., Konidaris, G., & Barto, A. (2014). Learning parameterized motor skills on a humanoid robot. In *International conference on robotics and automation (icra2014)* (p. 1-6).

Castro da Silva, B., Konidaris, G., & Barto, A. (2012). Learning parameterized skills. In *Proceedings of the 29th International Conference on Machine Learning (ICML 2012).* New York, NY, USA: Omnipress. Jul. (26 June – 1 July 2012, Edinburgh, Scotland, GB).

Catarino, A., Andrade, A., Churches, O., Wagner, A. P., Baron-Cohen, S., & Ring, H. (2013). Task-related functional connectivity in autism spectrum conditions: an EEG study using wavelet transform coherence. *Molecular Autism*, *4*, 1-14.

Cattaneo, L., Fabbri-Destro, M., S., B., Pieraccini, C., Monti, A., Cossu, G., & Rizzolatti, G. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proceedings of the National Academy of Sciences*, *104*, 17825-17830.

Chen, Z. J., He, Y., Rosa-Neto, P., Germann, J., & Evans, A. C. (2008). Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. *Cerebral Cortex*, *18*, 2374-2381.

Chersi, F., Mirolli, M., Pezzulo, G., & Baldassare, G. (2013). A spiking neuron model of the cortico-basal ganglia circuits for goal-directed and habitual action learning. *Neural Networks*, *41*, 212-224.

Ciancio, A. L., Zollo, L., Baldassare, G., Caligiore, D., & Guglielmelli, E. (2013). The role of learning and kinematic features in dexterous manipulation: a comparative study with two robotic hands. *International Journal of Advanced Robotic Systems*, *10*, E1-21.

Ciancio, A. L., Zollo, L., Guglielmelli, E., Caligiore, D., & Baldassarre, G. (2011). Hierarchical reinforcement learning and central pattern generators for modeling the development of rhythmic manipulation skills. In *Proceedings of the first joint IEEE International Conference on Development and Learning and on Epigenetic Robotics (ICDL-EpiRob 2011).*

Clearfield, M. W. (2011). Learning to walk changes infants? social interactions. *Infant Behavior and Development*, *34*, 15-25.

Clifton, R. K., Muir, D. W., Ashmead, D. H., & Clarkson, M. G. (1993). Is visually guided reaching in early infancy a myth? *Child Development*, *64*, 1099-1110.

Doya, K. (2000). Reinforcement learning in continuous time and space. *Neural Computation*, *12*, 219-245.

Drescher, G. L. (1991). *Made-up minds a constructivist approach to artificial intelligence*. MIT Press.

Edelman, G. M. (1987). *Neural darwinism: the theory of neuronal group selection*. New York,NY: Basic Books.

Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Eskenazi, T., Grosjean, M., Humphreys, G. W., & Knoblich, G. (2009). The role of motor simulation in action perception: A neuropsychological case study. *Psychological Research*, *73*, 477-485.

Fahlman, S. E. (1988). *An empirical study of learning speed in back-propagation networks* (Tech. Rep. No. CMU-CS-88-162). Pittsburgh, Pennsylvania, PA 15213: School of Computer Science, Carnegie Mellon University.

Fahlman, S. E., & Lebiere, C. (1990). The cascade-correlation learning architecture. In *Advances in neural information processing systems 2* (p. 524-532). Morgan Kaufmann.

Fernández, F., & Veloso, M. (2006). Probabilistic policy reuse in a reinforcement learning agent. In *Proceedings of the fifth international joint conference on autonomous agents and multiagent systems* (p. 720-727).

Forssell, D. (2004). Descriptive versus generative scientific theories. In D. Forssell (Ed.), *Perceptual control theory: A book* *of readings* (p. 39-44). Hayward, CA: Living Control Systems Publishing.

Fuster, J. M. (2001). The prefrontal cortex–an update: time is of the essence. *Neuron*, *30*, 319-333.

Gernsbacher, M. A., Sauer, E. A., Geye, H. M., Schweigert, E. K., & Hill, G. H. (2008). Infant and toddler oral- and manual-motor skills predict later speech fluency in autism. *Journal of Child Psychology and Psychiatry*, *49*, 43-50.

Gittins, J. C., & Jones, D. M. (1979). A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrika*, *66*, 561-565.

Gowen, E., & Hamilton, A. (2012). Motor abilities in autism: A review using a computational context. *Journal of Autism and Developmental Disorders*, *43*, 323-344.

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory*, *70*, 119-136.

Graziano, M. S., Taylor, C. S., & Moore, T. (2002). Complex movements evoked by microstimulation of precentral cortex. *Neuron*, *34*, 841-851.

Grossberg, S., & Seidman, D. (2006). Neural dynamics of autistic behaviors: cognitive, emotional, and timing substrates. *Psychological Review*, *113*, 483-525.

Gurney, K., Prescott, T., & Redgrave, P. (2001a). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological Cybernetics*, *84*, 401410.

Gurney, K., Prescott, T. J., & Redgrave, P. (2001b). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biological Cybernetics*, *84*, 411-423.

Hamilton, A. F., & Grafton, S. (2007). The motor hierarchy: from kinematics to goals and intentions. In Y. Rosetti, M. Kawato, & P. Haggard (Eds.), *Attention performance* (Vol. 22, p. 381-408). Oxford University Press.

Haruno, M., Wolpert, D. M., & Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural Computation*, *13*, 2201-2220.

He, Y., & Evans, A. (2010). Graph theoretical modeling of brain connectivity. *Current Opinion in Neurology*, *23*, 341-350.

Herbort, O., Ognibene, D., Butz, M. V., & Baldassarre, G. (2007). Learning to select targets within targets in reaching tasks. In Y. Demiris, B. Scassellati, & D. Mareschal (Eds.), *The 6th IEEE International Conference on Development and Learning (ICDL2007)* (p. 7-12). London: Imperial College.

Houk, J. C., Adams, J. L., & Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. C. Houk, J. L. Davids, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (p. 249-270). Cambridge, MA: The MIT Press.

Houk, J. C., Bastianen, C., Fansler, D., Fishbach, A., Fraser, D., Reber, P. J., … Simo, L. S. (2007). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philosophical Transactions of the Royal Society B Biological Sciences*, *362*, 1573-1583.

Houk, J. C., Davis, J., & Beiser, D. (Eds.). (1995). *Models of information processing in the basal ganglia*. Cambridge, MA: The MIT Press.

Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural Computation*, *3*, 79-87.

Jordan, M., & Jacobs, R. (1994). Hierarchical mixtures of experts and the em algorithm. *Neural Computation*, *6*, 181-214.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*, 809-816.

Kober, J., & Peters, J. (2009). Policy search for motor primitives in robotics. *Advances in neural information processing systems*, *21*, 849-856.

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

Mannella, F., Mirolli, M., & Baldassarre, G. (2010). The interplay of pavlovian and instrumental processes in devaluation experiments: a computational embodied neuroscience model tested with a simulated rat. In C. Tosh & G. Ruxton (Eds.), *Modelling perception with artificial neural networks* (p. 93-113). Cambridge: Cambridge University Press.

Mareschal, D. (2003). Encyclopaedia of cognitive science. In L. Nadel (Ed.), (p. 533-539). London, UK: Macmillan..

McClelland, J. L. (1995). Developing cognitive competence: New approaches to process modeling. In T. J. Simon & G. S. Halford (Eds.), (p. 278-304). Hillsdale, NJ: Erlbaum.

Meier, J. D., Aflalo, T. N., Kastner, S., & Graziano, M. S. (2008). Complex organization of human primary motor cortex: a high-resolution fmri study. *Journal of neurophysiology*, *100*, 1800-1812.

Meunier, D., Lambiotte, R., & Bullmore, E. T. (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience*, *4*, 11.

Middleton, F. A., & Strick, P. L. (1996). The temporal lobe is a target of output from the basal ganglia. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 8683-8687.

Middleton, F. A., & Strick, P. L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Research Reviews*, *31*, 236-250.

Morgan, S. (1986). Autism and Piaget's theory: are the two compatible? *Journal of Autism and Developmental Disorders*, *16*, 441-457.

Nishimoto, R., & Tani, J. (2009). Development of hierarchical structures for actions and motor imagery: a constructivist view from synthetic neuro-robotics study. *Psychological Research*, *73*, 545-558.

Oudeyer, P.-Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, *11*, 265-286.

Parisi, D., & Schlesinger, M. (2002). Artificial life and Piaget. *Cognitive Development*, *17*, 1301-1321.

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain: A Journal of Neurology*, *60*, 389-443.

Piaget, J. (1951). *Play, dreams and imitation in childhood*. London: Routledge.

Piaget, J. (1953). *The origins of intelligence in children*. London: Routledge and Kegan Paul.

Pouget, A., & Latham, P. E. (2003). Population codes. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (Second ed.). Cambridge, MA, USA: The MIT Press.

Rasheed, W. A. J., & Ali, H. A. (2009). Generalization aspect of neural networks on upgrading assimilation structure into accommodating scheme. *Journal of Computer Science*, *5*, 177-183.

Ring, M., Schaul, T., & Schmidhuber, J. (2011). The two-dimensional organization of behavior. In A. Cangelosi et al. (Eds.), *Proceedings of the International Conference on Devel-*

opment and Learning and Epigenetic Robotics (ICDL-EpiRob-2011) (p. E1-6). Piscataway, NJ: IEEE.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. area F5 and the control of distal movements. *Experimental Brain Research*, *71*, 491-507.

Rosenbaum, D. A., Loukopoulos, L. D., Meulenbroek, R. G. M., Vaughan, J., & Engelbrecht, S. E. (1995). Planning reaches by evaluating stored postures. *Psychological Review*, *102*, 28-67.

Rosenthal, J., Massie, H. N., & Wulff, K. (1980). A comparison of cognitive development in normal and psychotic children in the first two years of life from home movies. *Journal of Autism and Developmental Disorders*, *10*, 433-443.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533-536.

Sandini, G., Metta, G., & Vernon, D. (2007). The iCub cognitive humanoid robot: an open-system research platform for enactive cognition. In M. Lungarella, F. Iida, J. Bongard, & R. Pfeifer (Eds.), *50 years of artificial intelligence. essays dedicated to the 50th anniversary of artificial intelligence* (Vol. 4850). Berlin: Springer-Verlag.

Sarlegna, F., Blouin, J., Bresciani, J. P., Bourdin, C., Vercher, J. L., & Gauthier, G. M. (2003). Target and hand position information in the online control of goal-directed arm movements. *Exp Brain Res*, *151*, 524-535.

Schaal, S., Peters, J., Nakanishi, J., & Ijspeert, A. (2005). Learning movement primitives. *Robotics Research*, *15*, 561-572.

Schembri, M., Mirolli, M., & Baldassarre, G. (2007a). Evolution and learning in an intrinsically motivated reinforcement learning robot. In F. Almeida e Costa, L. M. Rocha, E. Costa, I. Harvey, & A. Coutinho (Eds.), *Advances in artificial life. proceedings of the 9th european conference on artificial life (ecal2007)* (Vol. 4648, p. 294-333). Berlin: Springer Verlag.

Schembri, M., Mirolli, M., & Baldassarre, G. (2007b). Evolving childhoods length and learning parameters in an intrinsically motivated reinforcement learning robot. In L. Berthouze, G. P. Dhristiopher, M. Littman, H. Kozima, & C. Balkenius (Eds.), *Proceedings of the Seventh International Conference on Epigenetic robotics (EpiRob2007) modeling cognitive development in robotic systems* (Vol. 134, p. 141-148). Lund: Lund University.

Schembri, M., Mirolli, M., & Baldassarre, G. (2007c). Evolving internal reinforcers for an intrinsically motivated reinforcement-learning robot. In Y. Demiris, B. Scassellati, & D. Mareschal (Eds.), *Proceedings of the 6th ieee international conference on development and learning (icdl2007)* (p. 282-287). London: Imperial College.

Schlesinger, M., & Cangelosi, A. (2013). *An introduction to developmental robotics*. Cambridge: MA: MIT Press.

Schultz, R. T., Gauthier, I., Klin, A., Fulbright, R. K., Anderson, A. W., Volkmar, F., … et., a. (2000). Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and asperger syndrome. *Archives of General Psychiatry*, *57*, 331-340.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593-1599.

Schwarz, A. J., Gozzi, A., & Bifone, A. (2009). Community structure in networks of functional connectivity: resolving functional organization in the rat brain with pharmacological MRI. *NeuroImage*, *47*, 302-311.

Shultz, T. R. (2003). *Computational developmental psychology* (M. Press, Ed.).

Shultz, T. R., Schmidt, W. C., Buckingham, D., & Mareschal, D. (1995). Developing cognitive competence: New approaches to process modeling. In T. Simon & G. Halford (Eds.), (p. 347-362). Hillsdale, NJ: LEA..

Singh, S., Lewis, R., Barto, A., & Sorg, J. (2010). Intrinsically motivated reinforcement learning: An evolutionary perspective. *IEEE Transactions on Autonomous Mental Development*, *2*, 70-82.

Strick, P. L., Dum, R. P., & Fiez, J. A. (2009). Cerebellum and nonmotor function. *Annual Review of Neuroscience*, *32*, 413-434.

Sugimoto, N., Haruno, K., M Doya, & Kawato, M. (2012). Mosaic for multiple-reward environments. *Neural Computation*, *24*(3), 577-606.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge MA, USA: The MIT Press.

Taffoni, F., Formica, D., Schiavone, G., Scorcia, M., Tomassetti, A., Polizzi di Sorrentino, E., … Guglielmelli, E. (2013). The mechatronic board: A tool to study intrinsic motivations in humans, monkeys, and humanoid robots. In G. Baldassarre & M. Mirolli (Eds.), *Intrinsically motivated learning in natural and artificial systems* (p. 411-432). Berlin: Springer-Verlag.

Taffoni, F., Vespignani, M., Formica, D., Cavallo, G., Polizzi di Sorrentino, E., Sabbatini, G., … Guglielmelli, E. (2012a). A mechatronic platform for behavioral analysis of non human primates. *Journal of Integrative Neuroscience*, *11*, 87-101.

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

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

Thelen, E., Corbetta, D., Kamm, K., Spencer, J. P., Schneider, K., & Zernicke, R. F. (1993). The transition to reaching: mapping intention and intrinsic dynamics. *Child Dev*, *64*, 1058-1098.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thill, S., Caligiore, D., Borghi, A. M., Ziemke, T., & Baldassarre, G. (2013). Theories and computational models of affordance and mirror systems: An integrative review. *Neuroscience and Biobehavioral Reviews*, *37*, 491-521.

Thomas, M. S. C., Knowland, V. C. P., & Karmiloff-Smith, A. (2011). Mechanisms of developmental regression in autism and the broader phenotype: a neural network modeling approach. *Psychological Review*, *118*, 637-654.

Thrun, S. (1992). *Efficient exploration in reinforcement learning* (Technical report No. CMU-CS-92-102). Pittsburgh, Pennsylvania: School of Coputer Science, Carnegie-Mellon University.

Tikhanoff, V., Cangelosi, A., Fitzpatrick, P., Metta, G., Natale, L., & Nori, F. (2008). Proceedings of IEEE workshop on performance metrics for intelligent systems workshop. In R. Madhavan & E. R. Messina (Eds.), (p. 57-61). Washington, DC: IEEE.

Tommasino, P., Caligiore, D., Mirolli, M., & Baldassarre, G. (2012). Reinforcement learning algorithms that assimilate and accommodate skills with multiple tasks. In *IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob2012)* (p. e1-8). Piscataway, NJ: IEEE.

Tommasino, P., Caligiore, D., Mirolli, M., & Baldassarre, G. (submitted). Transfer expert reinforcement learning (TERL): A reinforcement learning architecture that transfers knowledge between skills when solving multiple tasks. *IEEE Transactions on Autonomous Mental Development (TAMD)*.

Van Geert, P. (1998). A dynamic systems model of basic developmental mechanisms: Piaget, Vygotsky, and beyond. *Psychological Review*, *105*, 634-677.

von Hofsten, C. (2004). An action perspective on motor development. *Trends in Cognitive Science*, *8*, 266-272.

von Hofsten, C. (2007). Action in development. *Developmental Science*, *10*, 54-60.

Williams, J., Thomas, P. R., Maruff, P., & Wilson, P. H. (2008). The link between motor impairment level and motor imagery ability in children with developmental coordination disorder. *Human Movement Science*, *27*, 270-285.

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*, 464-476.