What are Intrinsic Motivations?
A Biological Perspective

Gianluca Baldassarre
Laboratory of Computational Embodied Neuroscience, Istituto di Scienze e Tecnologie della Cognizione,
Consiglio Nazionale delle Ricerche (LOCEN-ISTC-CNR), Via San Martino della Battaglia 44, I-00185 Roma, Italy
Email: gianluca.baldassarre@istc.cnr.it

Abstract—The concept of “intrinsic motivation”, initially proposed and developed within psychology, is gaining an increasing attention within cognitive sciences for its potential to produce open-ended learning machines and robots. However, a clear definition of the phenomenon is not yet available. This theoretical paper aims to clarify what intrinsic motivations are from a biological perspective. To this purpose, it first shows how intrinsic motivations can be defined contrasting them to extrinsic motivations from an evolutionary perspective: whereas extrinsic motivations guide learning of behaviours that directly increase fitness, intrinsic motivations drive the acquisition of knowledge and skills that contribute to produce behaviours that increase fitness only in a later stage. Given this difference, extrinsic motivations generate learning signals on the basis of events involving body homeostatic regulations, whereas intrinsic motivations generate learning signals based on events taking place within the brain itself. These ideas are supported by presenting some examples of biological mechanisms underlying the two types of motivations. The paper closes by linking the theory to the current major computational views on intrinsic motivations and by listing the main open issues of the field.

I. INTRODUCTION

The goal of this paper is to define the concept of intrinsic motivations from a computationally-informed biological perspective. In particular, this goal will be achieved by contrasting intrinsic motivations to extrinsic motivations.

The concept of intrinsic motivations has been proposed and developed within the psychological literature to overcome the difficulties of the behaviourist theory on learning and drives (e.g., [1], [2]) to explain why animals spontaneously engage in puzzles [3] or can be instrumentally conditioned to produce particular responses on the basis of apparently neutral stimuli (e.g., a sudden light onset, [4]) as it happens with standard primary rewards (e.g., food). Subsequent proposals have highlighted how the properties of certain stimuli can drive animals’ exploration and guide their learning processes, for example when they are complex, unexpected, or in general “surprising” [5]. Another important thread of psychological research has highlighted the importance that action plays in intrinsic motivations, for example in relation to the fact that a person manages to affect the environment with her behaviour (“effectance”, [6]), or can autonomously set her own goals and master their achievement [7]. Recently, neuroscience has started to investigate brain mechanisms that are relevant to explain intrinsic motivations. Examples of these mechanisms are illustrated in Sec. V.

Lately the issue of intrinsic motivations has drawn the attention of both computational modelling of behaviour and machine learning. The reason is that intrinsic motivations have the potential, if translated into suitable algorithms, to greatly enhance the learning capabilities of autonomous intelligent machines and robots. Some major computational theories on intrinsic motivations are presented in Sec. VI.

Notwithstanding this rising interest in the topic, the very concept of intrinsic motivations is not fully clear, in particular it is not clear if and how they differ from extrinsic motivations. The goal of this theoretical paper is to clarify the concept of intrinsic motivations from a biological perspective. To this purpose, I contribute to the understanding of intrinsic motivations answering the four types of questions that, according to the ethologist Tinbergen [8], can lead to a full understanding of an observed behaviour.

The first two questions proposed by Tinbergen relate to the ultimate mechanisms behind a behaviour. The first question is: “What is the adaptive value (“survival value”, in Tinbergen’s terms) of intrinsic motivations?” To answer this question one has to explain how intrinsic motivations improve the fitness of the organisms that possess them. This question plays a key role in this paper, and Sec. II addresses it. The second question is: “How did intrinsic motivations evolve during phylogeny (“evolution”)?” To answer this question one has to explain how intrinsic motivations emerged and changed during evolution, for example by comparing intrinsic motivations in different species having common ancestors but different degrees of differentiation from them. This question is only briefly considered here (Sec. II) due to the limited information and space available.

The other two questions proposed by Tinbergen relate to the proximate mechanisms behind a behaviour. In particular, the third question is: “What are the mechanisms (“causations”) behind intrinsic motivations?” To answer this question one has to understand the physiological and neural mechanisms that implement intrinsic motivations in the brain of organisms. This question plays a central role for the distinction between intrinsic and extrinsic motivations put forward here: Sec. III-V address this problem in detail. The fourth question is: “How are intrinsic motivations acquired during the life of organisms (“ontogeny”)?” This question might not be relevant here as Tinbergen’s questions were formulated to study behaviours, but intrinsic motivations are not a behaviour: they are rather
an important set of mechanisms that drive ontogeny itself and as such they seem to have mainly an innate origin. However, although not done here, it would be interesting to investigate if and how they are characterised by any form of plasticity that changes them during life. To this respect, however, it is important to stress that intrinsic motivation mechanisms should not be confused with those related to secondary reinforcers studied within psychology [9]. Secondary reinforcers are initially-neutral stimuli that acquire reinforcing properties based on a repeated association with primary reinforcers, either extrinsic (e.g., food) or intrinsic (e.g., a novel stimulus; Sec. II-V will discuss primary reinforcers at length). As such, the distinction between primary and secondary reinforcers is orthogonal to the distinction between intrinsic and extrinsic motivations. An analogous irrelevant distinction is the one involving “internal and external” elements: both intrinsic and extrinsic motivations involve external stimuli suitably processed by mechanisms internal to the organisms [10].

Before considering the differences between extrinsic and intrinsic motivations, some caveats are in order. Although the main effort here is to find ways of distinguishing between the two types of motivations, I recognise that there might be a sort of continuum between them, as it is generally the case in biology. In particular, there exist some classes of motivations that posses the features of both (a case is presented in Sec. IV). Moreover, extrinsic and intrinsic motivations share a number of underlying brain mechanisms (see Sec. III-V), and interact in several interesting ways (I will say little on this for lack of space). However, I think that this paper shows that it is possible to define the prototypical features of intrinsic and extrinsic motivations. For a cognitive scientist this distinction has a status similar to the one between birds and mammals for a zoologist: it is an important and useful distinction notwithstanding the existence of interesting intermediate species such as the platypus and the echidna.

Aside the aforementioned sections, Sec. VI reviews some major computational perspectives on intrinsic motivations and clarifies their relation with the theory proposed here, and Sec. VII lists some of the major open challenges of the field.

II. THE NATURE AND ADAPTIVE FUNCTIONS OF INTRINSIC AND EXTRINSIC MOTIVATIONS

Evolution is the key principle for understanding any biological phenomenon. In this respect, extrinsic and intrinsic motivations are not an exception.

Organisms evolve body structures that allow the increase of their survival and reproductive chances (fitness) in the particular niche where they live. 1 Bodies posses actuators (muscular and skeletal systems) that allow organisms to move in the environment and to modify it. Bodies also posses sensors (various sensory systems) that allow organisms to perceive both the external environment and the visceral body. With visceral body I refer to the inner organ systems of organisms (the integumentary, ingestive, excretory, circulatory, endocrine, lymphatic, and reproductive systems) and I do not include in it actuators and sensors. As we shall see, visceral body is very important for extrinsic motivations whereas sensors and actuators are important for both extrinsic and intrinsic motivations. 2 Brains (nervous systems) are specialised portions of bodies that can store skills (i.e., sensorimotor mappings) and knowledge (i.e., the capacity to abstract sensations and to predict). Skills and knowledge allow organisms to produce behaviours that eventually increase fitness.

Behaviours can be innate. In this case the structure of the brain (or the plan to build it) is encoded in the DNA and evolves so as to directly produce adaptive behaviours. Since Lorenz [12] and Tinbergen [8], [13], ethology has systematically studied innate behaviours, the stimuli that release them, their relation to needs, their hierarchical organisation, etc. [14]. Innate behaviours are fundamental for adaptation as they represent readily-available fast answers to many (“most”, for many species) needs the animals might have.

However, innate behaviours have also a fundamental drawback: they emerge in the course of several generations based on the accumulation of random mutations, so they fail to adapt when the environment changes rapidly or when the organisms move into new environments [15]. When the flexibility of learning can produce advantages that overcome its additional costs (e.g., higher metabolic costs and the possibility to learn wrong behaviours) then learning mechanisms evolve [16]. Learning generally requires the evolution of two types of mechanisms having two distinct functions: (a) Learning mechanisms: these allow the physical modification of brain, for example the formation of new synapses, so that it can produce new behaviours; (b) Guiding mechanisms: these generate learning signals in charge of guiding learning mechanisms to acquire behaviours that increase fitness (in other words, they establish the direction of behavioural change). I anticipate that extrinsic and intrinsic motivations tend to involve common learning mechanisms, and to change the same brain structures, but to involve different guiding mechanisms. Since Pavlov [17] and Skinner [18], psychology, and later psychobiology, have studied the effects of learning mechanisms on behaviour in the lab [9]. They have also studied the origin of the guiding mechanisms (i.e., of primary reinforcers) but to a much lesser extent, and when they have done so they have mainly focussed on those related to extrinsic motivations [9].

Motivations have at least three important functions for organisms (cf. [19]). One is to guide behaviour to attend, at each moment, the most important needs. A second one is to establish the amount of energy invested in executing the selected actions. A third one is to generate learning signals. In this paper I focus on the latter function, and in particular on the generation of learning signals to guide trial-and-error.
learning. I focus on this type of learning as there is a larger amount of biological literature on the relation of intrinsic motivations with the latter than with other types of learning processes, and as trial-and-error learning is closely related to reinforcement learning methods, the field of machine learning that is more keenly investigating intrinsic motivations.

A main claim of the paper is related to extrinsic motivations. Extrinsic motivations are motivations based on mechanisms that drive learning of skills and knowledge, and the exploitation and energisation of behaviours, on the basis of the levels and variations of homeostatic needs detected within the visceral body. Such skills and knowledge have the adaptive function to produce behaviours that allow the regulation of those homeostatic needs so as to increase fitness. The rationale of this claim is as follows.

The guidance of learning, that should tell how useful a skill or piece of knowledge is, requires measuring how such skill or knowledge might impact on fitness. Now, measuring fitness is not an easy job to be accomplished by brain and the mechanisms guiding learning. The point is that the brain can find good “proxies” of fitness within the visceral body, namely a multitude of variables that correlate with a number of conditions that enhance fitness, for example the level of energy stored in the body, the presence of needed body components, the integrity of body, the contact with a partner that might prelude reproduction, etc. The extrinsic-motivation mechanisms can use these fitness proxies, and in particular their changes towards reference levels, to generate learning signals that lead to evaluate positively, and hence to retain, skills and knowledge that generate adaptive behaviour.

Extrinsic motivations are characterised by two other distinctive features. The first relates to their diffusion among different species. Learning mechanisms emerged to acquire behaviours directed to regulate homeostatic needs. Extrinsic motivations emerged together with them to guide the resulting learning processes. For this reason, all species capable of trial-and-error learning have extrinsic motivations, even if with a degree of sophistication comparable to their overall sophistication. A second feature is that the learning signals produced by extrinsic motivations tend to cease when the homeostatic needs causing them are satisfied, and to come back again and again when the needs come back. This is adaptive as organisms have to attend homeostatic needs for their whole life.

A second main claim of the paper regards intrinsic motivations. Intrinsic motivations are motivations based on mechanisms that drive learning of skills and knowledge, and the exploitation and energisation of behaviours that facilitate this, on the basis of the levels and the variations of such skills and knowledge directly detected within the brain.

Intrinsic motivations have the adaptive function to allow organisms to learn skills and knowledge without the necessity to have a direct impact on homeostatic needs and fitness at the time of the acquisition. These skills and knowledge contribute to increase fitness as they can later be used to learn, relatively quickly, complex behaviours and long chains of actions that regulate homeostatic needs. I now explain this claim.

The mechanisms guiding learning based on extrinsic motivations render a great adaptive advantage to organisms as they allow them to readily change behaviours when the environment changes so as to continue to fulfil their homeostatic needs. However, such guiding mechanisms have a limitation: they cannot drive the acquisition of skills and knowledge that allow to positively affect homeostatic need regulations, and hence generate learning signals, only much later. Indeed, the environment offers a number of opportunities to improve homeostatic need regulation by performing relatively complex behaviours and long action chains. Now, learning mechanisms of animals fail to function when there are long delays between the performed behaviours and the learning signals they cause. Moreover, there are few chances to produce, by trial-and-error, complex behaviours and long action chains that result in a positive impact on homeostatic needs. As a consequence complex behaviours and chains would never be learned based only on extrinsic motivations.

Intrinsic motivations allow organisms to overcome this problem as they can guide learning of skills and knowledge without having to rely upon learning signals generated by homeostatic need regulation. To this purpose, intrinsic motivations generate learning signals on the basis of the success of the acquisition of skills and knowledge themselves instead of their effects on homeostatic needs. To do this, intrinsic motivations are based on mechanisms that measure the success of the acquisition of skills and knowledge directly within the brain. For example, these mechanisms drive organisms to continue to engage in a certain activity if their competence in achieving some interesting outcomes is improving, or if their capacity to predict, abstract, or recognise percepts is not yet good or is improving: the brain detects all these conditions without involving the visceral body. Once the skills and knowledge are acquired, they can be used as readily available building blocks to rapidly learn and produce complex behaviours and action chains under the guidance of extrinsic motivations, or as a basis for new acquisitions under the guidance of intrinsic motivations (cumulative learning).

Intrinsic motivations are characterised by two additional features that contribute to distinguish them with respect to extrinsic motivations (see above). The first is that they likely emerged, or at least substantially developed, later in evolution with respect to extrinsic motivations. This is shown by the fact that they are almost absent in lower species and are sophisticated only in higher ones, especially those that can rely upon a long period of parental care to fulfil their survival needs. So, for example, birds have a short period of parental care after which they have to strive to attend homeostatic needs. Instead primates, and especially humans [21], have a very long period of parental care and a great part of their

3Note that this does not mean that intrinsic motivations do not support “embodied” and “situated” learning processes in terms of how this concept is intended within the simulation-of-adaptive-behaviour literature (see [20] for a seminal paper on this concept). In fact, as shown in Sec. V, they support learning processes that fully involve the complex circular interactions of organisms with the environment mediated by sensors and actuators.
childhood is dominated by intrinsic motivations that allow them to learn a rich repertoire of skills and knowledge that allow them to rapidly solve useful tasks in adulthood.

A second feature is that the learning signals produced by intrinsic motivations tend to decrease or disappear after the skills or pieces of knowledge that generated them are acquired. For example, when a child has learned to reliably produce a newly discovered effect, or to recognise a novel object, the intrinsic motivation related to them tends to cease and she starts to direct her activities elsewhere. The adaptive function of this mechanism is that once something has been learned cognitive resources should be directed to learn something else.

III. EXTRINSIC MOTIVATIONS

This section reviews empirical evidence that supports the idea that the ultimate origin of the learning signals generated by extrinsic motivations reside in the visceral body. I first present the general idea, and then I corroborate it by reviewing specific examples of mechanisms behind extrinsic motivations related to food and water intake and temperature regulation, some major events affecting organisms’ fitness.

Let us start from the illustration of the general functioning pattern of extrinsic motivations. Organisms perceive the environment through sensors and on this basis produces actions. When these actions are adaptive, they cause a number of effects on the visceral body that enhance the organisms’ fitness. For example, they allow organisms to avoid damage from predators, to increase energy levels, to intake body constituents, and to exchange reproductive materials with mating partners. These effects are detected by a number of sensors scattered within the visceral body and the resulting information is sent to the brain via suitable chemical/neural pathways and systems. On this basis the brain generates learning signals, often based on neuromodulators, hormones, and peptides, that are broadcast to all areas of the brain itself in charge of learning. The learning processes so triggered lead the organisms to acquire knowledge and skills that allow them to produce, with a higher probability and accuracy, the behaviours that caused the original fitness-enhancing effects in the visceral body. A distinctive feature of these mechanisms (cf. Sec. II) is that the learning signals based on them are not produced when the related homeostatic needs are satisfied, but come back again and again for the whole organisms’ life (see [22] on this).

The machinery behind these processes reveals also another general pattern. Evolution developed brain mechanisms that directly modulate the functioning of visceral body, and trigger innate behaviours, in order to keep homeostatic balance. For example hypothalamus (see below) slows down body metabolism if the energy level in the body is low. These same systems contribute to produce extrinsic learning signals that guide learning of behaviour that can lead to similar body homeostatic regulations but this time indirectly via the mediation of the external environment.

I now review some specific examples of mechanisms behind extrinsic motivations, starting with the motivations that generate learning signals based on food intake. Hypothalamus (Hyp) is the major hub of body homeostatic regulations and plays an important role in food intake regulation [23]–[25]. There are other important areas underlying homeostic regulations but we focus on Hyp for simplicity: see for example [22] on the functions played by amygdala. Visceral body informs the Hyp on energy levels and body-constituent needs via a neural pathway involving the nucleus of the solitary tract and the parabrachial nucleus. Based on this information, the arcuate nucleus and other nuclei of the lateral Hyp directly control body metabolism mainly by regulating the autonomic and the endocrine systems (Hyp also controls innate behaviours, as shown by the fact that electrical stimulation of the ventromedial nucleus suppresses feeding and stimulation of lateral nucleus elicits feeding).

Aside these direct regulations, Hyp plays also a fundamental role in generating primary learning signals. In particular, Hyp can activate the ventral segmental area (VTA) both directly and indirectly via the peduncolopontine nucleus [26]. VTA is one of the major sources of dopamine (DA) in brain. DA, especially when phasic (i.e., caused by short intense activation bursts of DA neurons), is one of the most important neuromodulators involved in triggering learning processes in basal ganglia (BG) and frontal cortex (FC; BG are important sub-cortical nuclei in charge of acquiring skills by trial-and-error; FC plays a key role in learning and performing motor actions and in executive control [24]). As shown by the experiments on animal self-stimulation, mammals can quickly learn to perform an action if this causes an electrical stimulation of VTA neurons or their projections [27], [28]. These experiments, and other using natural rewards (e.g., food) and physiological recordings of DA neuron activations, have led many authors to claim that DA is the main learning signal the brain uses to acquire behaviours by trial-and-error [29].

Some authors claim that the main physiological correlate of primary rewards and fitness-increasing events are opioids and not DA, relevant for learning seeking behaviours [19], [30]. This claim is based on various experiments showing a detachment of DA bursts from primary rewarding events (e.g., if rewarding events are repeatedly preceded by a cue they tend to cease to produce DA). Opioids are a family of peptides produced by Hyp arcuate nucleus, solitary tract nucleus, and other brain areas receiving important visceral afferents [31], [32]. Compelling evidence shows that opioids are released during consummatory behaviours (e.g., ingestion of food, sexual activities, maternal cares, etc.). They are also released by the post-ingestion homeostatic consequences of food intake (see [19] for a review). Opioids systems are also related to DA and learning processes, as shown by the fact that their depletion in various areas reduces DA release [30].

The circuits behind temperature regulation and drinking seem to follow the general pattern illustrated above. The Hyp is informed on body temperature by peripheral sensors and by low/high temperature sensors situated within its anterior part. On this basis Hyp regulates temperature [33]. For water, the anterior Hyp has sensors that can detect cells osmolarity (concentration of salt). Moreover, when blood volume decreases
due to insufficient fluids in the body, the kidney produces an enzyme that causes the production of angiotensin in blood. Angiotensin passes into the brain through the subfornical organ, a small neural structure in the third ventricle that projects to the anterior Hyp and can trigger drinking behaviours [24], [34]. As food, temperature and water can activate the dopamine and opioids systems [29], [35] and hence modulate learning. For example, a rat can be trained to press a lever to produce cool-air puffs not only by setting it in a room with high temperature, but also by directly warming its anterior Hyp [36].

All these mechanisms are examples of how changes in the visceral body can ultimately generate learning signals.

IV. EXTRINSIC MOTIVATIONS THAT BYPASS THE VISCERAL BODY

This section introduces, but does not expand for lack of space, an intermediate form of extrinsic motivations that bypass the visceral body and rely only on external sensors and brain (these are called here non-visceral extrinsic motivations for reference). Aside the sensors detecting useful changes happening inside the body, organisms often possess other sensors located on the surface of body that, with their activation, can anticipate fitness-enhancing visceral-body changes caused by behaviour. Smell and taste sensors are examples of these sensors. They are activated not only just before useful resources produce positive effects in visceral body (in particular during consummatory behaviours, e.g. when food is chewed or during a sexual act), but also in anticipation of such events, for example by the odours released by food or a sexual partner located far from the body. These sensors send information to brain centres involved in generating learning signals. Interestingly, these centres tend partially overlap to those underlying the prototypical extrinsic motivations reviewed is Sec. III (e.g., Hyp and amygdala for smell; nucleus of the solitary tract for taste [24]).

The adaptive function of non-visceral extrinsic motivations is the generation of learning signals before those generated by consummatory behaviours and their effects. This facilitates the acquisition of preparatory behaviours as it shortens the delay elapsing between the behaviour and the learning signals and so overcomes the limitation of learning processes that do not tolerate delayed feedbacks [9].

This section is closed with an important caveat. Non-visceral extrinsic motivations are a case of extrinsic motivations only if the learning signals they generate are not actually generated by secondary reinforcers (see Sec. I): for now this possibility cannot be ruled out.

V. INTRINSIC MOTIVATIONS

This section reviews empirical evidence and theories that support the claim that the ultimate origin of the learning signals generated by intrinsic motivations resides in the brain. Unfortunately, the literature available on this issue is much less ample and deep than that on extrinsic motivations, although I think it is still enough to support the claim. I now illustrate the general pattern of intrinsic motivation mechanisms, and then I review specific examples of them. Organisms perceive the world and on this basis they produce behaviour. Behaviour produces other perceptions and on this basis intrinsic motivation mechanisms in the brain generate learning signals if the organisms (i.e., the brain itself) are acquiring new skills and/or knowledge. Importantly, the generation of these learning signals does not involve the visceral body. Another distinctive feature of these mechanisms (cf. Sec. II) is that when a certain skill or piece of knowledge is acquired, it ceases to produce learning signals.

Let us now see a first specific example of mechanisms underlying intrinsic motivations. One of the most articulated and empirically well-supported theories relevant for intrinsic motivations is related to the function of phasic DA [37]. The theory proposes that unexpected events (e.g., a light suddenly switched on) activate the superior colliculus (SC; SC, located in the midbrain, receives input from the retina and plays a key role in saccade generation). In turn, SC activates dopaminergic areas and causes phasic DA bursts that reach BG after few milliseconds after the events activating the SC. The DA acts as a “stamp” that causes the association of all information present in BG at that moment, in particular the action that caused the event and the context where it was executed. This phenomenon leads to repeat the skill leading to the event so that the organism can refine the skill and learn the relation between the skill and its outcome. With repeated experiences, the system also learns to inhibit SC and the learning signals it generates, so the skill ceases to be performed. Later on, the skill can be recalled for execution if its outcome becomes desirable (e.g., on the basis of extrinsic motivations). Notice how these mechanisms possess the typical features of intrinsic motivations: (a) the adaptive function of the performed behaviour is to acquire skills and knowledge (skill-outcome associations); (b) the learning signals are generated within the brain by SC; (c) the learning signals cease after learning.

A second important example of intrinsic motivation mechanisms comes from the studies on hippocampus (Hip), DA, and memory formation (Hip, a phylogenetically ancient cortex, plays a crucial role in the acquisition and consolidation of memory thanks to its widespread connections to most neocortex and its capacity of learning arbitrary associations based on few/episodic exposures to them). Hip is highly activated by the perception of unfamiliar objects, the perception of novel spatial arrangements of familiar objects, and the perception of new sequences of familiar objects [38]. In turn, Hip activates the VTA DA neurons, and this leads to the formation of new memories within Hip itself and the cortical areas targeted by VTA, in particular the FC [39]. This allows Hip and the cortex to learn new stimuli/stimuli associations and to progressively respond less intensely to their presentation, so causing learning signals with lower intensity. Notice how also in this case: (a) the adaptive function is to acquire knowledge, namely to learn.

The timing and stereotyped dynamics of DA is a key element used in [37] to support the theory, and also to claim that brain uses phasic DA mainly to signal unexpected events more than extrinsic rewards.
new objects and associations; (b) the learning signals originate
within the brain, namely within the same Hip; (c) the learning
signals cease after the percepts are memorised.

A last example of intrinsic motivation mechanisms is drawn
from [40]. This work presents an interesting theoretical-
neuroscience proposal related to possible functions of the
neuromodulators acetylcholine (ACh) and noradrenaline (NE),
which play important but subtle roles in learning [24]. NE is
produced by the locus coeruleus (LC) in the pons and targets
most cortex. ACh is produced by the nucleus basalis magnocellularis,
mainly targeting the neocortex, and by medial septum, targeting Hip. The starting point of the proposal is
the assumption that the brain is “generative”. Briefly, this
means that brain has top-down expectations on the world
functioning, and these expectations are used to integrate noisy
inputs to produce more accurate actions. The expectations
are acquired on the basis of their mismatches with actual
experiences, for example by cortex. A key idea of the proposal
is that ACh levels are high with high uncertainty, that is
when expectations are not considered very reliable. On this
basis, ACh regulates both how much the system relies on
expectations/actual perception, and how intensely it updates
the expectations. Classical conditioning experiments support
this idea as show that animals learn faster when stimuli have
consequences with higher variability [41]. A second key idea
is that NE signals unexpected uncertainty, happening when
there is a mismatch between top down expectations considered
reliable and direct perception. In this respect, it has been
shown that LC activates in correspondence to novel stimuli
and objects encountered during free exploration [42], [43],
and then rapidly habituates after few experiences. Notice again
the typical features of intrinsic motivations: (a) ACh and
NE signals have the function to support the acquisition of
knowledge on the world; (b) the learning signals originate
within the brain (mismatches in cortex); (c) the learning
signals fade away when the world behaves as expected.

VI. RELATION OF THE VIEW PRESENTED HERE WITH
OTHER MODELS AND PROPOSALS

This section presents a selection of computational theories
and models that have contributed to clarify the concept of
intrinsic motivations in ways relevant for a biological or
theoretical point of view. We will see that, in agreement with
Sec. II, most proposals view intrinsic motivations as a source
of learning signals that allow the acquisition of skills and
knowledge when the learning signals from the task to be
solved are still not available, or are difficult to achieve, or are
far in time/space (cf. [44]). I divide the works to review in two
groups: those of the first group are related, or might be used
to model, the biological mechanisms of intrinsic motivations
reviewed in Sec. V, whereas those of the second group are
relevant for the evolutionary framework presented in Sec. II.

The first theory of the first group [45], [46] proposes that
the fundamental aspects of most intrinsic motivations are
captured by systems endowed with a learning predictor (i.e.,
a model of the world). These systems can produce an intrinsic
learning signal on the basis of the prediction error or the
prediction error decrease: this signal can be used to train a
reinforcement learning component to explore portions of the
environment where the predictor learns the most. Successively,
these models have been developed into systems applicable to
real robots, for example on the basis of algorithms that can
partition the space on which prediction learning is applied
[47]. These models can be related to the mechanisms of
acetylcholine and noradrenaline guiding learning on the basis
of top-down expectation errors (Sec. V). As an alternative,
they might be related to tonic dopamine [48]. Directly inspired
by basal-ganglia acquisition of actions under the guidance of
dopamine phasic signals produced by superior colliculus (Sec.
V), [49] has proposed a model that uses the prediction error
in predicting salient events as a learning signal to learn, by
reinforcement learning and in a cumulative fashion, the actions
that produce such events.

Another approach, inspired by habituation mechanisms in
real organisms, has built habituable neural networks capable
of evaluating the level of familiarity of experienced scenes,
for example to perform patrolling tasks with robots [50].
Here a first component of the system (e.g., a self-organising
map) performs a vector quantisation of the input stimuli;
a second one (e.g., a simple linear neural network trained
with supervised learning) gets as input the output of the
first component and returns, with a single output unit, the
familiarity of the input pattern. Initially, the weights of the
second component are set to high values so that the output
of the system is high, signalling a high perceived novelty
of the input patterns. Then, while the self-organising map
learns the patterns, the output unit is trained towards zero so
the perceived novelty of the patterns decreases. The novelty
signals can eventually be used to lead the system to further
engage with those patterns. Note how these models represent
an operational instance of intrinsic motivations having the
function to guide the acquisition of knowledge on stimuli in
the form of representations, abstractions, etc. (Sec. II). In this
respect, models similar to these might be used to capture the
novelty detection mechanisms of hippocampus (Sec. V).

Another relevant thread of research (“IMRL – Intrinsically
Motivated Reinforcement Learning”) has proposed systems
that rely upon intrinsic learning signals generated by salient
events (e.g., a light onset in a room where the system has to
accomplish a certain task) to accelerate the acquisition of new
actions [10] (notably, this seminal work introduced the concept
of intrinsic motivations into the machine learning field). With
a similar focus on action learning, [51] has proposed to
use the TD-error of reinforcement learning algorithms, in
particular the error of prediction of the reward that can be
obtained executing a certain skill, as an index of how much
certain skill is improving: this index allows a higher-level
reinforcement learning component to (learn to) give control to
skills that have the highest learning rate. A similar mechanism
has been investigated in [52] based on a principled framework.
For now, no brain mechanisms have been found that implement
these principle: will these be discovered in the future after the
investigate the involvement of visceral body in motivations, it (involvement or not of the visceral body) and that this distinction is very important for cognitive science. However, it also shows that it is possible to distinguish them whole classes of environments good for to other relevant events (e.g. “box opening”; as in intrinsic fitness-enhancing events (as in extrinsic motivations) but also events; (b) the best reward functions are related not only to of reward functions variously related to fitness-enhancing regularities. The work shows that: (a) there can be a continuum to find a food within the fitness of the robot used by the genetic algorithm to evolve the reinforcers (intrinsic motivations) guiding learning in childhood. Although in a simplified manner, the model captures some key aspects of the theory presented here: (a) intrinsic motivations are mechanisms internal to the brain that guide the acquisition of knowledge and skills; (b) extrinsic motivations are mechanisms based on visceral body changes directly related to fitness; (c) evolution generates intrinsic motivations as these lead to acquire knowledge and skills (e.g. in childhood) that later (e.g. in adulthood) contribute to learn behaviours that enhance fitness.

The second work [53] adopts, as here, an evolutionary perspective to investigate intrinsic motivations. In particular, it uses a systematic exploration of the space of possible reward functions that can be used to learn, by reinforcement learning, to find a food within a class of grid words having certain regularities. The work shows that: (a) there can be a continuum of reward functions variously related to fitness-enhancing events; (b) the best reward functions are related not only to fitness-enhancing events (as in extrinsic motivations) but also to other relevant events (e.g. “box opening”; as in intrinsic motivations); (c) the idea of rewarding the acquisition of skills good for whole classes of environments is a key concept for intrinsic motivations. As mentioned in the introduction, the theory proposed here agrees that, to some extent, there might be continuum between extrinsic and intrinsic motivations. However, it also shows that it is possible to distinguish them and that this distinction is very important for cognitive science.

VII. DISCUSSION AND FUTURE WORK

This section closes the paper by illustrating some of the future challenges opened by the theory proposed here. A first main challenge is to show that the distinction between intrinsic and extrinsic motivations proposed here can be used operationally. Indeed, the distinction is based on both mechanisms (involvement or not of the visceral body) and functions (direct or indirect effects on fitness). While it seems viable to investigate the involvement of visceral body in motivations, it might be difficult to actually measure if a behaviour impacts fitness directly or indirectly via the acquisition of skills and knowledge. This problem might exacerbate for intermediate types of extrinsic motivations that bypass the visceral body (Sec. IV) as here body is no more an element of discrimination (but as seen in Sec. IV, they likely pivot on brain structures underlying the prototypical extrinsic motivations: this reintroduces the possibility to study body aspects). However, although evaluating the impact of behaviour on fitness is difficult, there are techniques for doing it [8], [14]. A further challenge is to corroborate the empirical evidence presented in Sec. III and V. For example, the whole biological literature is debating if dopamine, opioids, or other neural events represent the ultimate primary extrinsic reward signals [19]; or if dopamine is mainly the mediator of extrinsic or intrinsic motivations [37]. In this respect, another main challenge is to highlight how in detail some biological phenomena, e.g. those related to hippocampus processes [39] or acetylcholine dynamics [24], are related to intrinsic motivations. A further challenge is to clarify the relations existing between extrinsic and intrinsic motivations, for example to understand how the brain arbitrates between them, or how extrinsic motivations can impair the operation of intrinsic motivations [54].

On the computational side, much work still needs to be done due to the novelty of the field. First, we have various specific mechanisms to implement intrinsic motivations that lead to acquire knowledge, but few that lead to acquire competence [51]: what indexes can we used to this purpose [44]? Second, what are the best architectures to support true open-ended cumulative learning based on intrinsic motivations [55]? Finally, we have now various types of intrinsic motivation mechanisms, but none of them has proven capable of supporting a long-lasting sequence of ever-changing learning signals as needed by cumulative learning: should we improve the current mechanisms or should we find new ones? The theory proposed here should offer a useful framework to face all these biological and computational open questions.

ACKNOWLEDGMENT

This work is dedicated to Domenico Parisi for his keen effort to show how important the evolutionary perspective is for understanding all aspects of behaviour and brain. This research has been funded by the European Project “IM-CLeVeR - Intrinsically Motivated Cumulative Learning Versatile Robots”, grant agreement no. FP7-ICT-IP-231722.

REFERENCES
