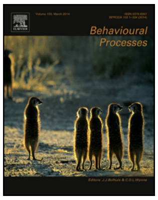
Behavioural Processes 115 (2015) 1–18

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc



Investigating intertemporal choice through experimental evolutionary robotics

Fabio Paglieri[a,*], Domenico Parisi[b], Massimiliano Patacchiola[c], Giancarlo Petrosino[b]

[a] Goal-Oriented Agents Lab (GOAL), ISTC-CNR, Roma, Italy
[b] Laboratory of Artificial Life and Robotics (LARAL), ISTC-CNR, Roma, Italy
[c] Centre for Robotics and Neural Systems, School of Computing and Mathematics, University of Plymouth, UK

ARTICLE INFO

Article history:
Received 2 May 2014
Received in revised form 7 February 2015
Accepted 21 February 2015
Available online 23 February 2015

Keywords:
Intertemporal choice
Evolutionary robotics
Ecological rationality
Delay discounting
Spatial discounting
Preferences

ABSTRACT

In intertemporal choices, subjects face a trade-off between value and delay: achieving the most valuable outcome requires a longer time, whereas the immediately available option is objectively poorer. Intertemporal choices are ubiquitous, and comparative studies reveal commonalities and differences across species: all species devalue future rewards as a function of delay (delay aversion), yet there is a lot of inter-specific variance in how rapidly such devaluation occurs. These differences are often interpreted in terms of ecological rationality, as depending on environmental factors (*e.g.*, feeding ecology) and the physiological and morphological constraints of different species (*e.g.*, metabolic rate). Evolutionary hypotheses, however, are hard to verify *in vivo*, since it is difficult to observe precisely enough real environments, not to mention ancestral ones. In this paper, we discuss the viability of an approach based on evolutionary robotics: in Study 1, we evolve robots without a metabolism in five different ecologies; in Study 2, we evolve metabolic robots (*i.e.*, robots that consume energy over time) in three different ecologies. The intertemporal choices of the robots are analyzed both in their ecology and under laboratory conditions. Results confirm the generality of delay aversion and the usefulness of studying intertemporal choice through experimental evolutionary robotics.

1. Introduction

Intertemporal choices concern options that can be obtained at different points in time: buying a luxury item today or saving the money to ensure a sizable pension in the future, enjoying a yummy dessert or dieting to ensure a better health over time, waiting for a fruit to ripen rather than eating it unripe. These choices often oppose a smaller but sooner prize (*e.g.*, a modest amount of food ready at hand) against a larger but delayed outcome (*e.g.*, a more distant but also richer foraging opportunity). How subjects behave in such circumstances is considered indicative of different degrees of self-control (Ainslie, 2001; Frederick et al., 2002; Berns et al., 2007), and the capacity to delay gratification in order to obtain a better outcome has been found to correlate with a variety of adaptive skills or traits, such as planning for the future, being able to interact socially, and enjoing physical health and well-being (Logue, 1988; Mischel et al., 1989; Tangney et al., 2004; Duckworth and Seligman, 2005; Moffitt et al., 2011).

After an initial emphasis on finding what mathematical model (typically a delay discounting function) would better fit the empirical data (Strotz, 1956; Ainslie, 1975; Laibson, 1997), current research on intertemporal choice has identified new priorities: most notably, the need to investigate the cognitive mechanisms that produce the observed behavioral patterns, and to trace their evolutionary roots (Berns et al., 2007). Comparative studies are essential for that purpose, and indeed intertemporal choices have been studied in insects (Cheng et al., 2002), birds (Logue and Peña-Correal, 1985; Chelonis et al., 1994; Mazur, 2007), rodents (van Haaren et al., 1988; Tobin et al., 1993; Green and Estle, 2003), non-human primates (Tobin et al., 1996; Stevens et al., 2005a; Rosati et al., 2007; Addessi et al., 2011; Stevens and Mühlhoff, 2012; Paglieri et al., 2013a,b), and of course humans – both adults (Green et al., 1994; Logue et al., 1996; Lawyer et al., 2010; Paglieri et al., 2013a,b, 2015) and children (Mischel, 1974; Schwarz et al., 1983; Thompson et al., 1997; Garon et al., 2012; Addessi et al., 2014).

Based on the available evidence, there is growing consensus on the fact that inter-specific differences in intertemporal choice behavior under laboratory conditions are tied to both environmental factors, *e.g.*, feeding ecology (Stevens et al., 2005a,b,b), social dynamics (Amici et al., 2008), and tool use (Rosati et al.,

---

* Corresponding author. Tel.: +39 0644595336.
E-mail address: fabio.paglieri@istc.cnr.it (F. Paglieri).

http://dx.doi.org/10.1016/j.beproc.2015.02.014
0376-6357/© 2015 Elsevier B.V. All rights reserved.

2007), and to the physiological and morphological constraints of different species, such as metabolic rate (Tobin and Logue, 1994). All these hypotheses emphasize the role of ecological features in determining the rationality of intertemporal choices observed under laboratory conditions. It has been argued (Gigerenzer and Goldstein, 1996) that ecological rationality is the most faithful and promising present-day development of Herbert Simon's intuitions on bounded rationality. In particular, ecological rationality, as the name implies, takes as central Simon's idea that cognitive skills should be understood and assessed in relation to the environment in which they have evolved (see for instance Simon, 1956). This is in line with a general effort at understanding decision-making in the context of evolution (Hammerstein and Stevens, 2012; Fawcett et al., 2014).

While evolutionary hypotheses on intertemporal choices are fascinating, they are also extremely hard to verify. In particular, explanations in terms of ecological rationality (Stephens and Anderson, 2001; Stevens and Stephens, 2010; Addessi et al., 2011; Fawcett et al., 2012) rely on key assumptions about the environment characteristic of a given species. Unfortunately, real environments are typically hard to observe with the degree of precision required to pinpoint the effects of a given ecological variable (*e.g.*, risk of predation) on the behavior under study (*e.g.*, intertemporal choice). Moreover, ecological rationality assumes that laboratory responses are indicative of behavioral patterns adapted to cope efficiently with the ecology of a given species (Todd and Gigerenzer, 2012). When applied to evolutionary explanations, this makes reference to the ancestral environment where that species evolved, which may or may not coincide with its current ecology. Even though there are methods to acquire data on ancestral environments (*e.g.*, paleobiology may provide clues on past feeding behavior, and primate archeology can shed light on the kind of tools used ancestrally by non-human primates; Haslam et al., 2009), they are bound to deliver incomplete information at best, in spite of substantial research efforts.

This highlight the need to find also other methods for studying the interaction between evolution, ecology and behavior – more directly and with a greater degree of control on all the relevant variables. In this paper, we discuss the viability of an approach based on evolutionary robotics (Nolfi and Floreano, 2000; Floreano et al., 2008; Doncieux et al., 2011; Bongard, 2013). The basic idea is to evolve populations of simulated robots under specific ecological pressures,[1] and then observe their behavior under laboratory conditions, with the aim of drawing interesting implications for our understanding of natural organisms faced by similar tasks. Let us label this methodology experimental evolutionary robotics, to stress the fact that these robots are studied not only in the ecology where they evolve, but also under artificial laboratory conditions (for early examples of this method, see Da Rold et al., 2011; Saglimbeni and Parisi, 2011; Paglieri et al., 2014; for a general discussion see Parisi, 2014).[2]

In this paper, we describe a series of studies where simple robots evolve in an environment containing two different types of food, and then they are required to make various intertemporal choices between these two types of food under laboratory conditions. Our results, albeit preliminary, demonstrate the viability of this approach and how its results can be meaningfully compared with naturalistic data. Moreover, both strengths and weaknesses of experimental evolutionary robotics are highlighted in the process, and they turn out to be complementary to those of behavioral studies in psychology, economics, and biology. This argues in favor of an integrated approach to bounded rationality in both natural and artificial organisms.

1.1. Why should we study choice behavior with evolutionary robots?

Evolutionary robotics is now a well-developed approach to the study of behavior (Nolfi and Floreano, 2000; Floreano et al., 2008; Doncieux et al., 2011; Bongard, 2013), and it has been successfully applied to a variety of biological phenomena, including, among others, spatial navigation, morphological adaptation, cooperation, altruism, predation and evasion strategies (Floreano and Keller, 2010; Waibel et al.„ 2011; Long, 2012; Parisi, 2014). Its potential for social sciences, however, is still largely untapped: evolutionary robotics has been applied to only a handful of relevant psychological phenomena (*e.g.*, risk aversion; Niv et al., 2002; Arbilly et al., 2011; Paglieri et al., 2014), whereas complex, human-specific social dynamics, such as institutions and norms, are typically simulated using other computational techniques (*e.g.*, agent-based modeling in computational sociology; Axelrod, 1997; Macy and Willer, 2002; Squazzoni, 2012). This "division of computational labor" tends to relegate evolutionary robotics to the study of more "basic" biological phenomena: as we intend to show, this limitation is unwarranted, and intertemporal choice behavior constitutes an excellent case-study for testing the fruitfulness of experimental evolutionary robotics to the study of complex psychological dynamics. Nonetheless, the key merits and drawbacks of this methodology are not specific of this domain of inquiry. Thus, before describing our studies on intertemporal choice in robots, it is useful to briefly summarize what makes this method worthy of attention to start with, and what cautions need to be applied in interpreting the resulting data.

- Advantages

- *Full observability*: Experimental evolutionary robotics allows observing behavior both "in the wild", *i.e.*, in the ecological setting where robots evolve, and "in the lab", *i.e.*, under specific test conditions. In both cases, behavior can be studied in extreme detail, since the simulation keeps track of all data – in fact, one of the greatest challenges for the experimental roboticist is to avoid being swamped by data.
- *Full control*: All variables can be easily and precisely manipulated, regarding both ecology and test conditions. This extreme versatility can lead to arbitrariness (see below), but it also ensures that a relatively simple set-up can be adapted to a variety of different real-life environments and/or laboratory protocols. Moreover, it also allows performing "counterfactual experiments", that is, studying how ecological pressures for which no natural correlate is known might affect behavior – an exercise useful to test the predictive value of models based on existing environments.

---

[1] This process of evolution is achieved through evolutionary algorithms: simply put, the idea is to apply to artificial organisms some key mechanisms of natural evolution (*e.g.*, fitness, selection, random mutation) to generate and identify effective behavioral and/or morphological solutions to problems posed by a certain environment. A comprehensive and authoritative introduction to this thriving field is provided by Eiben and Smith (2007); further details on the specific evolutionary algorithms used in our studies are discussed later on.

[2] It is worth noting that all evolutionary robotics is, strictly speaking, experimental, in the sense that it manipulates certain variables (typically, features of the agent and/or the environment where evolution takes place) in order to observe the resulting effects. However, here we use "experimental" to emphasize that robots are not only observed in the environment where they evolved, but also in (an abstraction of) a laboratory setting, where they perform the equivalent of controlled experiments. This particular application of evolutionary robotics is still relatively rare, yet extremely promising, especially for integrating the use of computer simulations into behavioral sciences, where controlled experiments in laboratory settings are the key method.

Ecological shocks (*e.g.*, sudden famine, colonization from other populations, climate change) are also easy to replicate.

- *Neurocomputational transparency*: When applied to robots whose behavior is controlled by an artificial brain (a neural network of units linked together by unidirectional connections), this methodology also allows to measure the dynamics of the control system, to see what patterns of activation of the neural network's units correspond to different overt behaviors – similarly to what is done by looking at the brain in natural organisms. This provides a further level of comparison between natural and artificial cognition.
- *Individual differences*: Laboratory tests are typically performed on robots that, through evolution, have adapted well to the needs of their environment (although in principle nothing prevents from testing in the lab also "dysfunctional" robots from earlier, less well-adapted generations). While most individuals of the last generation can efficiently cope with their ecology, they do so in different ways and with different levels of proficiency: in other words, substantial individual differences are observed in their behavior, and the effect of such differences on their test performance can be precisely assessed. This is potentially very fruitful, in light of the analogous individual differences observed in virtually all species of animals.
- *Non-deterministic responses*: Behavior varies not only across robots, but also when the same robot is faced more than once by the same input, provided that a minimum amount of noise is introduced in its control system. The fact that evolutionary robots are typically non-deterministic, with respect to external stimuli, is important for comparing them with natural organisms, who also do not react always in the same way to identical inputs from the environment. Typically, an animal repeating the same choice across *N* sessions will not always opt for the same alternative, and this is true also for evolutionary robots.

- Current limitations

- *Abstraction*: Both the ecology and the artificial laboratory are much simpler than most natural counterparts, and the same is true for the structure of the robot's body and its control system. While this high level of abstraction invites extreme caution in making a straightforward comparison between the behavior of evolutionary robots and that of natural organisms, their very simplicity highlights how remarkable is their capacity to perform very well (that is, at a level comparable with that of much more complex natural organisms) in many difficult tasks. Nevertheless, great care is needed to ensure that relevant features of the behavior under study are not lost or distorted through this necessary process of abstraction. As a case in point, in the studies presented here we opted to use spatial distance as a proxy for temporal delay: this was done because the kind of simple robots we used had no awareness of time, while they were fully capable of perceiving space. To justify this choice, we made sure that movement would have no additional cost for the agent, aside from the time spent moving – either because the robot did not consume energy at all (Study 1), or because energy consumption depended solely on the passing of time, regardless of whether the robot was moving or not (Study 2). Besides, this choice is not completely arbitrary: many natural instances of intertemporal choice in animals involve a spatial component (*e.g.*, patch exploitation in blue jays and foraging decisions in baboons; Stephens and Anderson, 2001; Noser and Byrne, 2007), and spatial decision making has been experimentally studied in relation to delay discounting in various species (*e.g.*, guppies, marmosets, and cotton-top tamarins; Stevens et al., 2005b; Mühlhoff et al., 2011).
- *Arbitrariness*: In experimental evolutionary robotics, a variety of parameters need to be set by the researcher concerning type and number of food tokens in the robots' environment, presence/absence/number of competitors, the robots' body and control system, and the experimental test conditions. These parameters are likely to have an impact on the resulting behavior, and yet only a small sample of them can be directly manipulated and tested within a study: thus, there is a significant danger of arbitrariness in the results. The ideal solution to this problem is to use data or hypotheses from behavioral studies to set as many parameters as possible in ways that approximate some existing ecology and actual test conditions, and to design artificial organisms that simulate, in some relevant respect, the morphology and physiology of natural organisms (for an example of this method, see Caligiore et al., 2010). However, before undertaking the substantial efforts required by that strategy, it is advisable to check what kind of results can be obtained using a less constrained evolutionary robotic model: this is not only prudent, but also needed to pinpoint the effect of each variable on the behavioral results (see below).
- *Starting small*: Given the number of variables directly controlled by the experimenter and the huge amount of data obtained, experimental evolutionary robotics requires a scalar approach, especially when this method is applied to a given phenomenon for the first time – as it is the case here with intertemporal choice. The only reliable way of disentangling the effects of different variables on the results is by starting from the simplest possible scenario (even if this means making the evolutionary robots less similar to natural organisms), and then add one element of complexity at a time. As a case in point, in the studies presented here we used robots that have no learning ability, no competitors in their natural environment, and no need of a partner for reproduction, since the genetic heritage of the best individuals in a given generation is simply transferred (with mutations) to individuals of the next generation. More generally, these robots maximize their fitness according to a single parameter (amount of food consumed in Study 1, life length in Study 2) and live in an ecology where two different types of food are the only salient objects. This minimal environment can then be enriched in various ways, but it is very interesting to first see what kinds of choice behavior emerge even under these simple circumstances.
- *Simulated embodiment*: The robots discussed in this paper are simulated robots evolved in a simulated environment. Even though the software used, Evorobot*, is based on the e-puck robotic platform (Mondada et al., 2009), in this series of studies we did not explore the role of real physical embodiment in robots' behavior. Clearly, having a physical body is likely to change many things, as studies on embodiment have repeatedly demonstrated (Varela et al., 1992; Clark, 1997; Brooks, 1999; Pfeifer and Bongard, 2006). However, running evolutionary studies on physical robotic platforms raises a host of technical and methodological problems, many of which are bound to distract from the research issues under consideration. Moreover, the physical structure of the robots' body and of their environment would affect behavior in a variety of ways: the only hope to precisely trace all these interrelated influences is by "adding embodiment" one step at the time – which requires first gathering baseline results in the absence of actual physical constraints. For all these reasons, using a robotic simulator designed to replicate several key aspects of an actual robot is currently the best option to start exploring choice behavior in artificial organisms. In all cases, such method is based on simulated embodiment, which is very different from lack of embodiment – that is, using software agents that do not emulate any actual robotic platforms and live in virtual environments with no resemblance to the laws of physics.

- *Residual key differences*: As a result of the limitations discussed above, the choice situation faced by, respectively, an evolutionary robot and a natural organism is bound to differ in various respects. One of the most important differences is that robots do not distinguish between the ecology and the lab. While humans, as well as most species of animals, easily learn to identify a laboratory test as a "special situation" with its own set of rules, the simple evolutionary robots discussed here treat the stimuli presented them in the artificial lab as if it was observed in their everyday ecology. With respect to choice tasks, this implies that decisions will be reversible until the very last moment (while moving toward one food token, the robot remains free to change direction at any time), and robots will not treat the options as being either unique (there could be other food outside of the robot's current field of vision, and indeed a well-adapted robot will often turn about to look for it, when presented with visible options that are very far), or mutually exclusive (the fact that they approach one token first does not imply that they have forsaken the other ones, insofar as in their ecology food tokens do not disappear). In contrast, many animals (and certainly humans) faced by an intertemporal choice task typically understand that the observed options are unique and mutually exclusive, and that their decision cannot be reversed once it has been made. These differences invite caution in comparing our results with patterns observed in real labs. However, the fact that evolutionary robots do not separate the lab from the ecology is not necessarily an epistemological flaw: in fact, one could argue that it is the real lab to be flawed, precisely because it puts experimental subjects in a situation that they perceive as significantly different from "normal conditions" (for a discussion of artificiality and so called external validity in economical experiments, see Loewenstein, 1999; Starmer, 1999).

2. Study 1: robots without a metabolism

2.1. Basic experimental design

In this first study, we used robots without energetic constraints: that is, robots that do not need to eat in order to survive, since they do not consume energy over time. Nevertheless, their ability to eat is what determines their fitness and thereby their likelihood of leaving their genes to the next generation (see details below). Only two types of food were present in their environment, distinguished by color: green food (G) and blue food (B). The total number of food tokens in the environment at any given time was constant because, when a food token was eaten, another identical one appeared in a random position within the environment, and both types of food tokens were equally frequent and randomly distributed.

We evolved the robots under a variety of ecological conditions,[3] manipulating the following variables: magnitude of the rewards (*i.e.*, energetic values assigned to G and B), and total number of food tokens present in the environment. We simulated five different ecologies, as follows:

Ecology 0: five G tokens of value 1 and five B tokens of value 1 (baseline condition).

Ecology 1: five G tokens of value 1 and five B tokens of value 2.

Ecology 2: five G tokens of value 5 and five B tokens of value 10.

Ecology 3: five G tokens of value 5 and five B tokens of value 6.

Ecology 4: ten G tokens of value 1 and ten B tokens of value 2.

Aside from these manipulations, all ecological conditions were identical (see details below). For each ecology, we run two experimental tests on the individuals from the last generation:

- *Single option*: a single token of either type is placed in front of the robot, at various distances, and we measured how rapidly the individual ends up consuming it.
- *Binary choice, distancing the better option*: first two food tokens of different types are placed at the same distance (100 pixels) from the robot, then the distance of the more valuable token (B) is progressively increased and the choice is repeated.

2.2. Evolution: method and results

The robots live in a simulated environment of 1000 × 1000 pixels with randomly distributed food tokens which the robots have to eat if they want to reproduce and leave their genes to the next generation. Half of the food tokens are green and half are blue: their energetic value change across different ecological conditions, but in all cases (except in the baseline condition, Ecology 0) the blue token is more energetic than the green one. Life has the same length for all the robots (8000 cycles) but the robots differ in their ability to eat the green and blue tokens and, therefore, in the total quantity of energy that each individual robot is able to extract from the environment. The robots' behavior is under the control of a neural network, designed and evolved using the Evorobot* software (http://laral.istc.cnr.it/evorobotstar/) to control an e-puck robotic platform (http://www.e-puck.org/), a simple robot with a limited perceptual field and wheel-based locomotion in two-dimensional space (Fig. 1).

Since the robots of the first generation have random connection weights in their neural network, they are able to eat only a few food tokens during their life. The robots which, by eating, extract more energy from the environment reproduce by generating offspring that inherit the connection weights of their (single) parent with the addition of random changes in some of the weights (genetic mutations). The selective reproduction of the best robots and the constant addition of genetic mutations lead to a progressive increase in the number of food tokens eaten by the robots during their life: Fig. 2 summarizes the amount of energy obtained by the robots in a succession of 200 generations across all five ecological conditions, for both the average and best robots of each generation (here "best" means the 20% of individuals with the highest fitness). The initial population consisted of 100 randomly generated genotypes that encode the connection weights and the biases of 100 corresponding neural controllers (each parameter is encoded by 8 bits and normalized in the range [−5.0, +5.0]). The 20 best genotypes of each generation were allowed to reproduce by generating 5 copies each, with 2% of their bits replaced with a new randomly selected value.[4] The evolutionary process lasted 200 generations, and was replicated 10 times for each ecological conditions (seeds): unless stated otherwise, all results reported concern the average behavior of all robots from all seeds, except outliers.[5] Outliers were defined as individuals with a fitness level at the last generation that differed from the mean of their seed by more than two standard deviations – or, whenever this criterion generated negative values, as individuals in the worst 10% percentile of the last generation.

---

[3] We will keep referring to the environmental parameters affecting evolution as "ecology", for ease of reference: obviously, the label does not imply any approximation of the complexity of real ecologies – here each condition differs only for the amount and/or value of the food items available.

[4] Further details on the evolutionary algorithm and other simulation parameters are freely available upon request, to allow a reproducibility check of current results.

[5] We consider using 10 seeds reliable enough for the exploratory aims of the present work: however, it is worth emphasizing that the number of seeds could (and should, in future studies) be increased significantly (*e.g.*, to 30 or 100) without any special difficulty. A larger number of repetitions would provide even firmer evidence of the generality of the results discussed here.

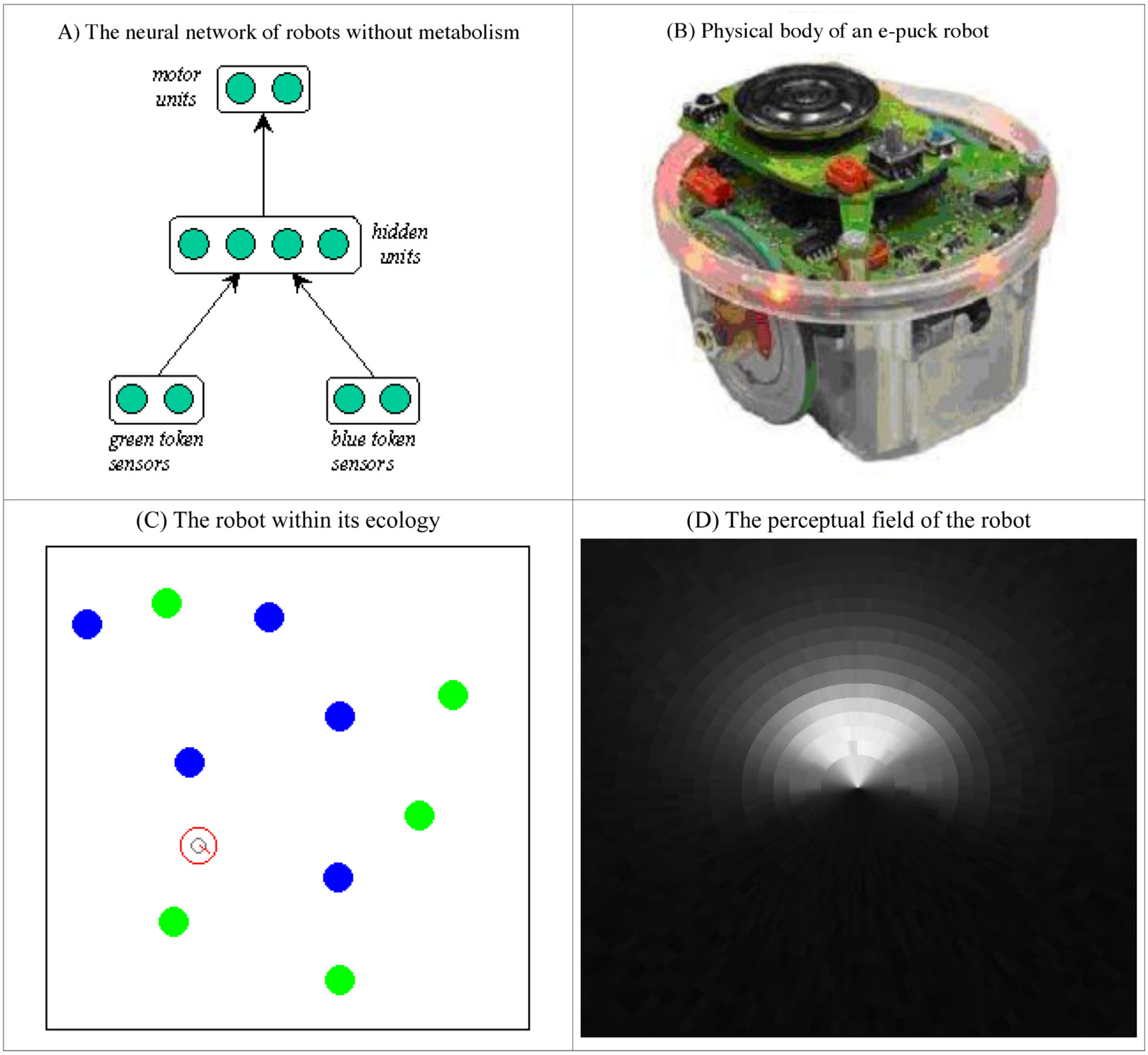


Fig. 1. The e-puck platform: control system (A), physical body (B), virtual environment (C), perceptual field (D).

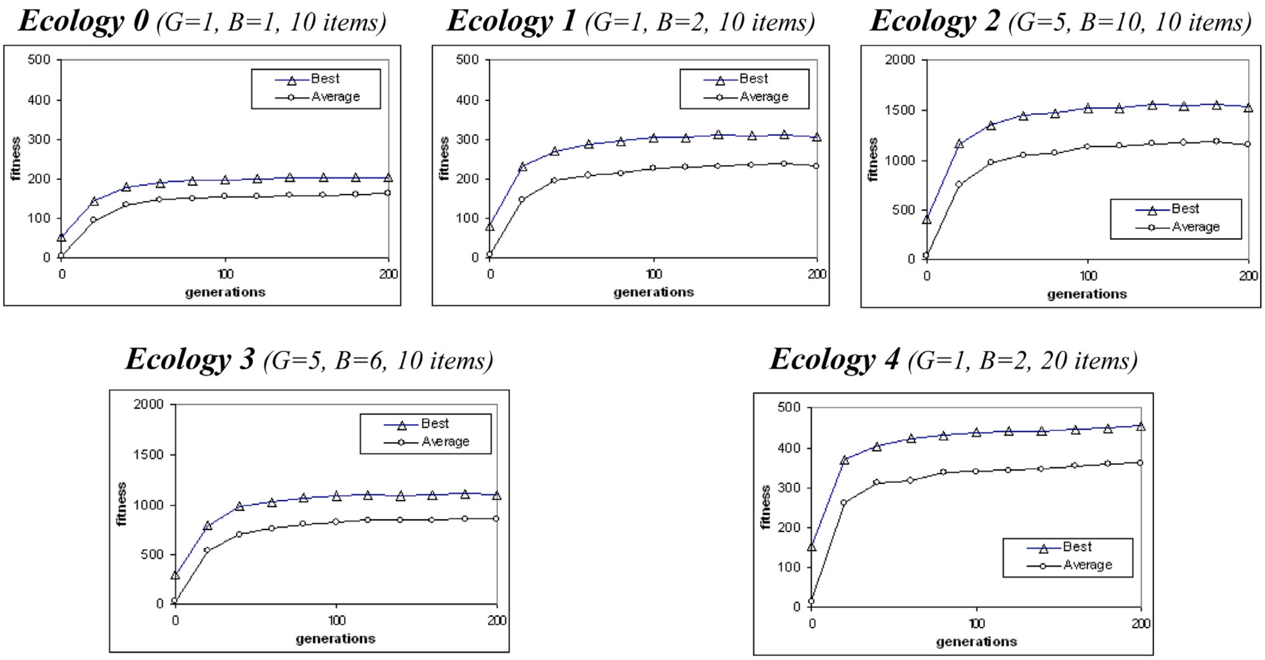
After their exclusion, the samples retained for analysis in each ecology were $N^{ec0} = 940$, $N^{ec1} = 849$, $N^{ec2} = 914$, $N^{ec3} = 935$, $N^{ec4} = 944$.

These results show that robots rapidly evolve the ability to maximize food intake in this type of environment, and that their fitness is sensitive to the overall amount of energy available in the environment. The overall amount of energy (OAE) available at any given time within an ecology is given by the number of tokens multiplied by their value: so we have that $\mathrm{OAE}_0 = 10$, $\mathrm{OAE}_1 = 15$, $\mathrm{OAE}_2 = 75$, $\mathrm{OAE}_3 = 55$, $\mathrm{OAE}_4 = 30$. As Fig. 2 shows, this measure correlates with the mean fitness value after 200 generations in each ecology, both for the average robots (Pearson: $r = 0.9929$, $p < .001$) and for the best ones (Pearson: $r = 0.9918$, $p < .001$). This result is not surprising: accumulating energy is easier in environments that are richer. Later on we will discuss more interesting results, concerning how specific forms of environmental richness (higher density of food tokens of identical value vs. more valuable tokens with equal density) impact on choice behavior in the lab.

When the two food types differ in their energetic value, robots develop a preference for the richer food: Fig. 3 summarizes the percentage of food tokens of each type consumed by the robots of the last generation of each ecology. Except when both food tokens have the same value (Ecology 0), the robot develops a robust preference for the blue token (comparison of percentages of choices for each token across seeds; $N = 10$, $p < .0001$ in Ecology 1, 2 and 4, $p < .001$ in Ecology 3, NS in Ecology 0). Moreover, this preference is sensitive to the ratio between the values of the two rewards, and not to their absolute difference: in fact, the preference for the more energetic token is stronger in Ecology 1 ($G = 1$, $B = 2$) than in Ecology 3 ($G = 5$, $B = 6$), albeit the value difference between the tokens remains the same (comparison of percent differences in preference across seeds; $N = 10$, $t(18) = 9.03$, $p < .0001$); in contrast, strength of preference does not change between Ecology 1 and Ecology 2 ($G = 5$, $B = 10$), where ratio is the same, in spite of the fact that the value difference is five times larger ($t(18) = 0.11$, $p = 0.9136$). Interestingly, also the density of food in the environment affects positively the preference for the better food: choices for the blue token are significantly more frequent in Ecology 4 than in Ecology 1 ($N = 10$, $t(18) = 5.23$, $p < .0001$), even if in both cases the token values are the same, and what changes is only the overall number of tokens.

2.3. Laboratory tests: method and results

The artificial laboratory was a quarter of the size of the evolution environment (500 × 500 pixels): only a limited number of food


Ecology 0 *(G=1, B=1, 10 items)*

Ecology 1 *(G=1, B=2, 10 items)*

Ecology 2 *(G=5, B=10, 10 items)*

Ecology 3 *(G=5, B=6, 10 items)*

Ecology 4 *(G=1, B=2, 20 items)*

Fig. 2. Fitness curves for all ecology, measured as the mean amount of energy obtained for each generation over 200 generations: notice the different energy scale (0–500) of the graphs for Ecologies 0, 1 and 4, in comparison with that used for Ecologies 2 and 3 (0–2000), to facilitate readability. Data refer to robots without a metabolism (Study 1).

tokens were introduced (one or two, depending on the test) and the experimenter controlled both their placement and the initial position of the robot. Every robot from the last generation in each ecology performed 10 sessions of each test. The robot was considered to have "chosen" whatever token it reached (and thereby ate) first: if after 400 cycles the robot had not yet reached any of the two food tokens, the trial was considered as null. Robots that performed null trials on all 10 sessions of an experiment (an extremely rare occurrence) were excluded from data analysis for that test. Depending on the test, two main metrics were used to assess performance: number of choices for each type of food tokens (%choice), and rapidity with which the robot reached the food token (rapidity). This last measure capture the intuition that the more valuable a token is, the surer and faster will be one's choice of it.

2.3.1. *Single option*

In this test, the robot is presented with a single food token, either a G food or a B food, and its reactions are observed: in particular, we recorded the time required to reach the food and divided it by its initial distance, thus assessing the rapidity of the robot's response. The single token was always placed directly in front of the robot, in the exact center of its perceptual field: its distance was varied across test conditions (respectively, 70, 100, 200, 300 pixels from the robot; for an example, see Fig. 4, panel A). We predicted that rapidity would decrease as a function of distance: in other words, in reaching a distant option, we expected the robot to waver and wander more that when moving toward nearer objectives. As shown in Fig. 5, this was indeed the pattern observed in our data: in all ecologies, individuals moved less promptly toward either food tokens, when the distance from it was increased.

To investigate how distance and reward type interact in this devaluation process, we performed a 4 × 2 ANOVA with repeated measures on both factors (see Table 1 for a summary of results). This confirmed that rapidity of approach decreases as a function of distance, while reward type has no main impact on rapidity, except for a weakly significant effect in Ecology 4. However, we observed

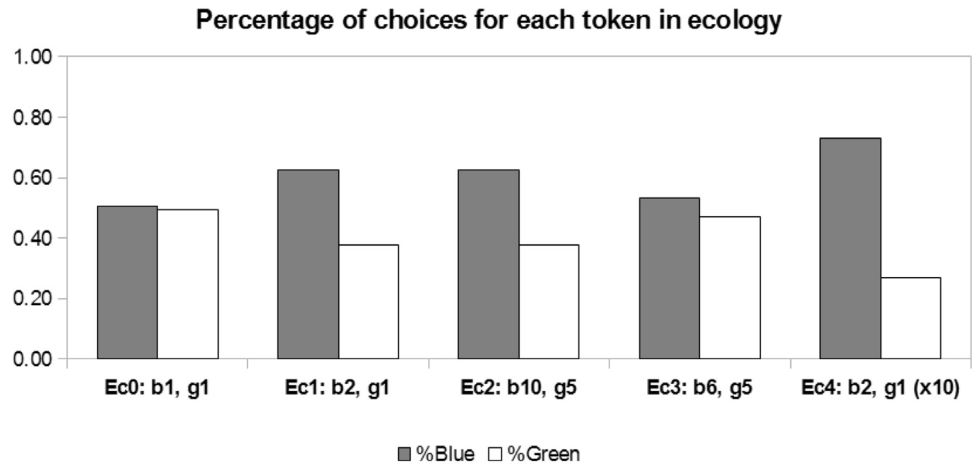


Fig. 3. Percentage of food tokens of each type consumed by robots of the last generation in their ecology. Data refer to robots without a metabolism (Study 1).

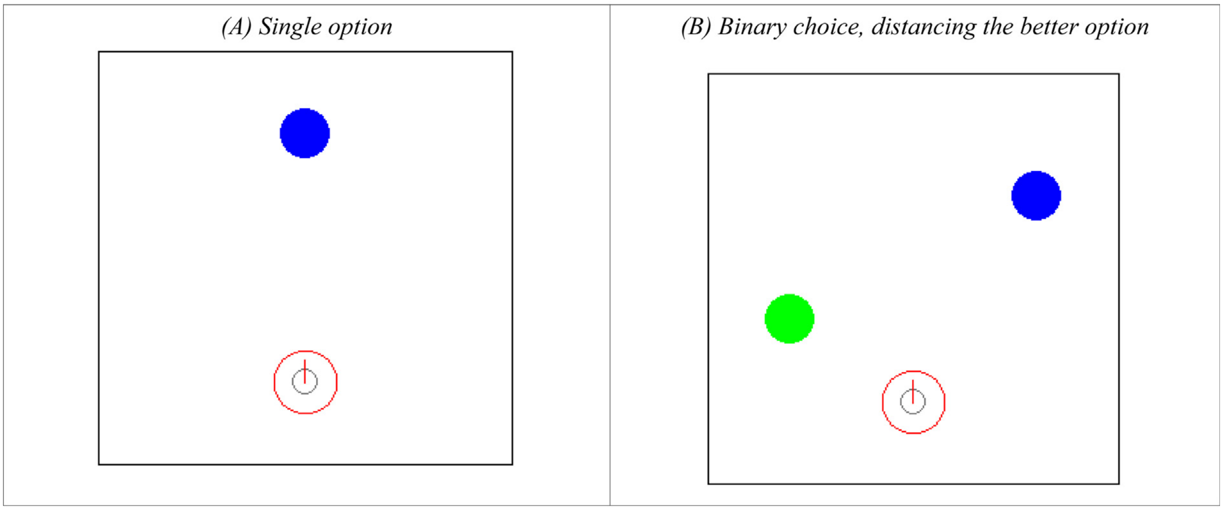


Fig. 4. Types of laboratory tasks in Study 1: single option (A) and binary choice, distancing the better option (B).

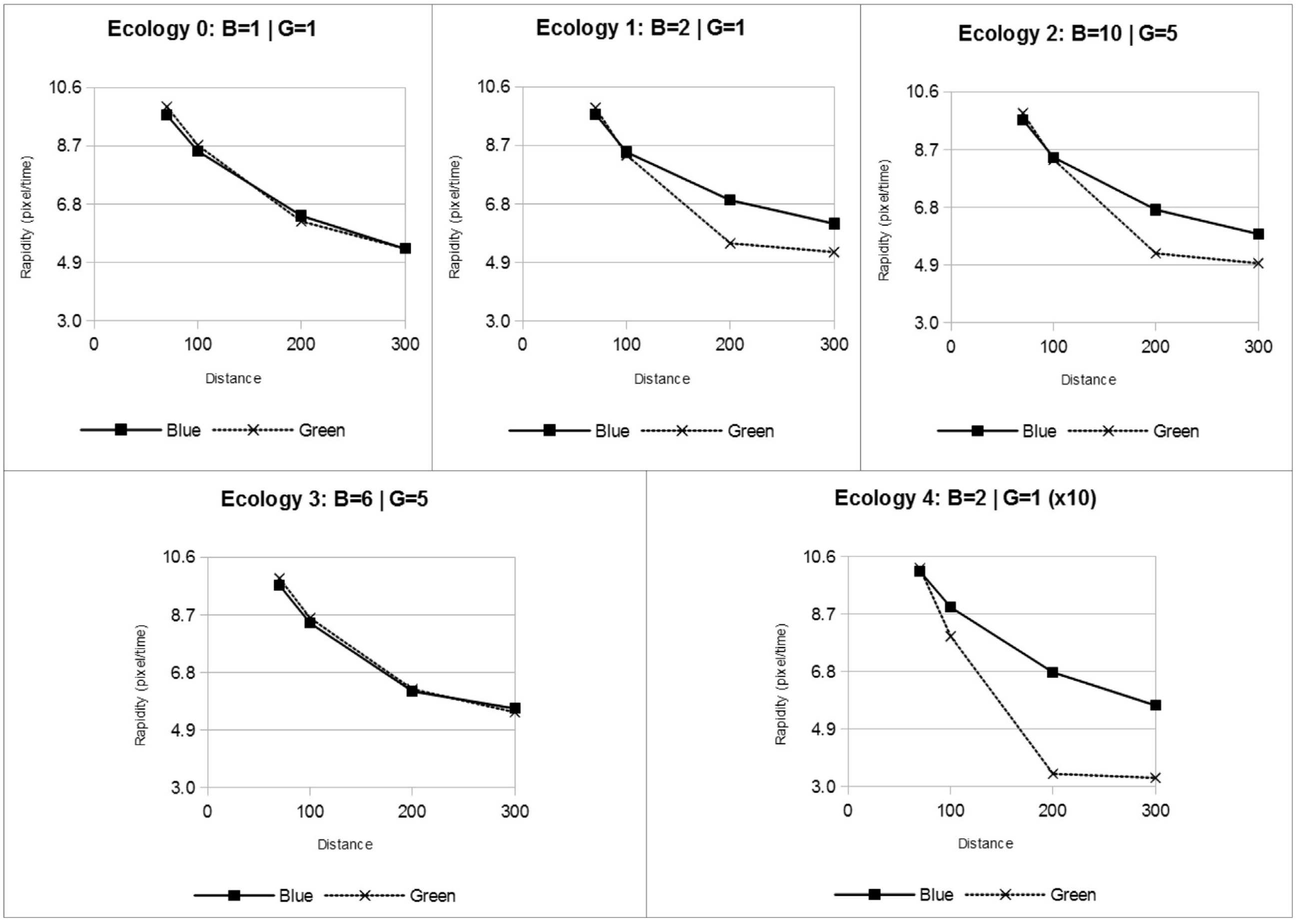


Fig. 5. Rapidity of approach (measured in pixel/time step) to a single token of each type, placed in front of the robot at various distances (70, 100, 200, 300 px). Data refer to robots without a metabolism (Study 1).

Table 1
ANOVA results for Distance (70/100/200/300 px) *X* Reward type (Green/Blue), both factors within-subjects, mean rapidity values for each seed, *N* = 10. Italics indicate statistical significance, $^{*}p<.01$, $^{**}p<.001$, $^{***}p<.0001$. Data refer to robots without a metabolism (Study 1).

	Distance	Reward type	Interaction
Ec0: g1, b1	$^{***}F(3,27)=937.52$, $p<.0001$	$F(1,9)=1.0463$, $p=0.333$	*, $F(3,27)=4.553$, $p=0.0105$
Ec1: g1, b2	$^{***}F(3,27)=1996.94$, $p<.0001$	$F(1,9)=0.6417$, $p=0.4437$	$^{***}F(3,27)=20.33$, $p<.0001$
Ec2: g5, b10	$^{***}F(3,27)=1859.49$, $p<.0001$	$F(1,9)=0.9384$, $p=0.358$	$^{***}F(3,27)=20$, $p<.0001$
Ec3: g5, b6	$^{***}F(3,27)=2389.06$, $p<.0001$	$F(1,9)=0.919$, $p=0.3628$	$^{**}F(3,27)=5.333$, $p=0.0051$
Ec4: g1, b2 (× 10)	$^{***}F(3,27)=249.51$, $p<.0001$	*, $F(1,9)=6.5779$, $p=0.0304$	$^{***}F(3,27)=23.14$, $p<.0001$

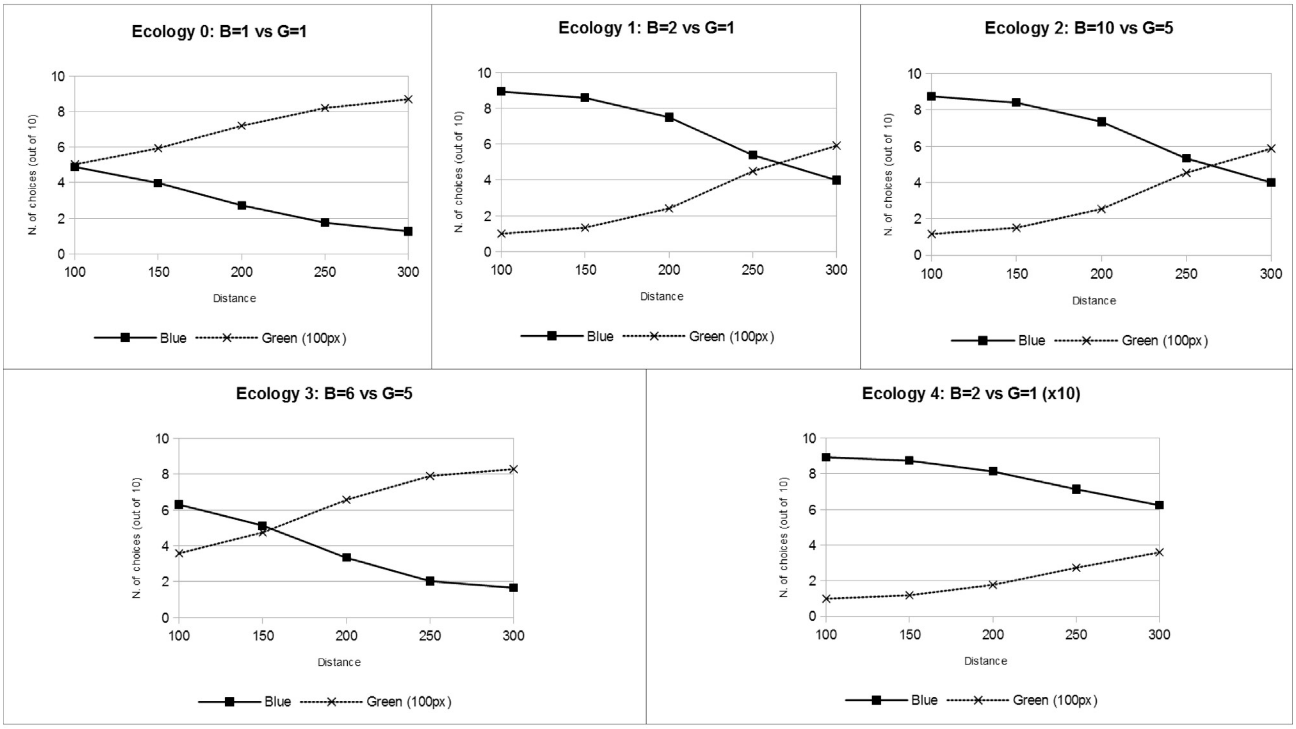


Fig. 6. Choice patterns in a binary choice when the distance from the blue token progressively increases and the green token is always at 100 pixels from the individual: mean number of tokens chosen by each individual out of 10 trials. Data refer to robots without a metabolism (Study 1).

a significant interaction between distance and reward type in all ecologies, and most markedly in Ecology 1, 2, and 4: while rapidity of approach is nearly identical across tokens on short distances, on longer distances rapidity decreases more sharply for the less valuable token (green), as opposed to the blue one (see Fig. 5). This interaction is affected more by the ratio of, than by the difference between, token values, as a comparison between Ecology 1, 2 and 3 immediately show. Moreover, also the density of food in the environment seems to affect this interaction, which is most pronounced in Ecology 4.

2.3.2. Binary choice, distancing the better option

In this test, the robot is first presented with two options, one green token vs. one blue token, both placed at the same distance (100 pixels), on two sides of its visual field, and with an inter-token distance of 100 pixels. Then the distance of the blue token from the robot is progressively increased (150, 200, 250, 300 pixels), while the green token remains at 100 pixels from the robot (for an example, see Fig. 4, panel B). Fig. 6 summarizes the choice patterns of the robots in this test: in general, distancing the blue token progressively reduces its appeal, until the individual becomes indifferent between the two options (indifference point) or develops a preference for the green token, even though it delivers less energy.

These data show how the indifference point between the two options shifts across ecologies. In Ecology 0, where both tokens have the same value, individuals are indifferent between equidistant options, and they start preferring the green option as soon as the blue one is moved further away – not surprisingly. In contrast, in Ecology 1 and 2 the mean indifference point is approximately 270 px, which means that individuals on average are indifferent between a token of value *X* (1 or 5) distant 100 px and a token of value 2*X* (2 or 10) distant 270 px; whereas in Ecology 3 the indifference point is significantly closer, around 160 px. Taken together, these findings show that indifference between two options of different value depends on the ratio of the values, rather than on their difference. Interestingly, in Ecology 4 we observe a marked effect of food density on choice behavior: in an environment where food tokens are more numerous, individuals are much more inclined to persist in pursuing a larger token, even when it is much further than a smaller one.

2.4. Discussion and comparison with natural organisms

This first study was mostly intended to test the general feasibility of our methodology. Given the extremely unnatural properties of these individuals (*e.g.*, lack of a metabolism), we did not expect to replicate natural data to any meaningful extent. However, we did predict that robots would in the binary choice task (Fig. 6) reveal a clear process of reward devaluation as a function of distance. The more distant the reward is, the less keen is the robot to reach it or to choose it over a less energetic prize. Current data do not yet permit to assess what mathematical model (exponential, hyperbolic, quasi-hyperbolic, sub-additive, *etc.*) best describes this form of spatial (hence temporal) discounting, since rapidity data cannot account for differences in the absolute value of rewards (the maximum rapidity depends on the robot specification, not on the features of the reward), whereas for binary choice more indifference points would be needed, to try plotting a discounting function. However, it is easy to see how these additional data could be gathered, and doing so will be part of our future work.

On the other hand, analyzing laboratory behavior *via* discounting functions might not be the most profitable approach, with respect to experimental evolutionary robotics. Perhaps the most notable added value of this methodology is to deliver rich data on evolutionary patterns, which can thus be put into contact with laboratory evidence. In this respect, there are two main findings worth emphasizing in this first study: the emergence of discounting processes even in the absence of metabolism, and the role of reward values and numerosity in determining behavior.

Regarding the first point, it is important to emphasize once again that the robots in this study do not have a metabolism: that is, they do not need energy to function, cannot die of hunger, or even experience it, and do not consume any energy while moving in the environment. Thus, tokens are relevant for them only as a means to maximize the chances of reproducing their genes in the next generation. As a consequence, many of the normal reasons for being averse to distance (or delay) do not apply to them: no extra consumption of energy is required to reach a distant reward, nor there is any life-threatening urgency of getting to the food before it is too late. In fact, the only costs positively associated with distance here are opportunity costs: given their limited life-span (8000 cycles), natural selection favors those robots who learn how to "make the most of it" – that is, robots that do not forsake valuable alternatives, just because they are fixated on reaching a distant objective. The fact that discounting emerges under such conditions prove that opportunity costs are sufficient to produce an evolutionary pressure toward it. Of course, this does not imply that discounting in natural organisms is the product of opportunity costs alone. But it does vouch for the possibility that opportunity costs might be a crucial factor in the evolution of temporal discounting – a suggestion already made by a variety of studies, and also backed by some empirical evidence (Stephens and Anderson, 2001; Stevens and Stephens, 2010; Paglieri, 2013).

As for the effects of reward values and numerosity on behavior, both under ecological conditions and in laboratory tasks, several interesting patterns are apparent in the data. First, changing reward magnitude while keeping their ratio constant does not produce any change in behavior: indeed, data from Ecology 1 ($G = 1$ vs. $B = 2$) and Ecology 2 ($G = 5$ vs. $B = 10$) are virtually indistinguishable. This might seem in contrast with the magnitude effect observed in intertemporal choices in humans (Kirby, 1997; Green et al., 1997, 1999), but it is not, because here we are comparing different magnitudes across ecologies, rather than within them, and the absence of a metabolism makes absolute magnitude adaptively indifferent: a blue-eating robot gains twice more energy than a green-eating robot both in Ecology 1 and in Ecology 2, and that is all that matters for reproductive success, in the absence of a metabolism. More precisely, the ratio between reward values matters more than their absolute difference, in terms of behavioral effects: this is why, in all relevant measures (see Figs. 3, 5, and 6), robots from Ecology 1 ($G = 1$ vs. $B = 2$) show sharper discrimination between reward types than robots from Ecology 3 ($G = 5$ vs. $B = 6$), even though the value difference in both ecologies is exactly the same. Finally, also the numerosity of food tokens has a significant impact on behavior, even when their values are kept constant across ecologies. Here the relevant contrast is between Ecology 1 and 4, in which blue and green tokens have the same values ($B = 2$ and $G = 1$), but are more numerous in the Ecology 4 (10 blue and 10 green) than in Ecology 1 (5 blue and 5 green). This results in sharper discrimination between reward types in all relevant measures: in particular, robots from Ecology 4 show a stronger ecological preference for blue food than robots in Ecology 1 (Fig. 3), which is further reflected by their greater tolerance for distance in the binary choice task (Fig. 6). This effect is probably due to the fact that, in an ecology more densely packed with food, the advantages of even a slight tendency to favor the better food are likely to pay off, thus increasing the rate at which that adaptive trait is selected. This seems confirmed by the fact that the fitness curve is steeper for Ecology 4 than for Ecology 1, especially in the first 50 generations or so (Fig. 2).

3. Study 2: hungry robots

3.1. Basic experimental design

In this second study, we introduced a simplified metabolism in the robots. We used robots that consume energy as a function of time (0.021 energy units consumed per time step) and die when their energy level reaches zero, so that they need food to acquire the energy necessary for their survival. Crucially, their energy level also has an upper boundary (set to 1), beyond which further food consumption no longer increases the robot's energy: this captures the fact that biological organisms cannot eat indefinitely. Moreover, the robots' brain is informed of the current energy level in the robots' body, to simulate the sensation of hunger of real animals. All other parameters of the robots were kept identical to Study 1, in order to assess how having a metabolism would impact on the behavioral patterns observed in the previous experiments.

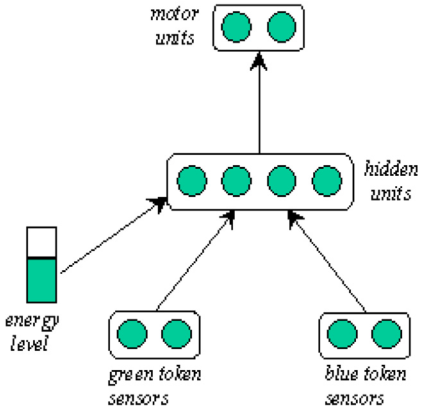


Fig. 7. Control system for hungry robots: the neural network encodes also the current energy level.

We evolved these robots under three ecological conditions, manipulating again the magnitude of the rewards (*i.e.*, energetic values assigned to G and B). With respect to Study 1, we had to use different numerical values, due to the introduction of a bounded energy level (from 0 to 1) for each individual. Concerning reward magnitude, we simulated three different ecology, as follows:

Ecology 5: five G tokens of value 0.1 and five B tokens of value 0.1 ($OAE_5 = 1.0$).

Ecology 6: five G tokens of value 0.1 and five B tokens of value 0.2 ($OAE_6 = 1.5$).

Ecology 7: five G tokens of value 0.3 and five B tokens of value 0.6 ($OAE_7 = 4.5$).

Aside from these manipulations, all ecological conditions were identical, and for each of them we run the same laboratory tests used in Study 1 on the individuals from the last generation. However, in addition we also manipulated the level of hunger of the individual, to see if and how this variable would affect behavior (see details below).

3.2. Evolution: method and results

The robotic platform, the virtual environment, and the evolutionary algorithm were exactly the same as in Study 1, except that now the neural network controlling the robot's behavior also included a sensory unit encoding the current energy level (see Fig. 7) and that fitness was measured as the length of a robot's life rather than as the quantity of food eaten by the robot (albeit, for obvious reasons, the two measures are strongly correlated).[6] This implies that, for the robots with a metabolism, it is not only important how much they eat but also when they eat: eating a food

---

[6] Notice that population size was kept constant (100 individuals) across generations: this is standard practice in evolutionary robotics, mostly to facilitate the evolutionary process and avoid risks of mass extinction during the first generations, when behavioral strategies are not yet well adapted to environmental constraints. However, this is also a significant abstraction with respect to natural selection, where extinction is a very concrete (and necessary) occurrence.

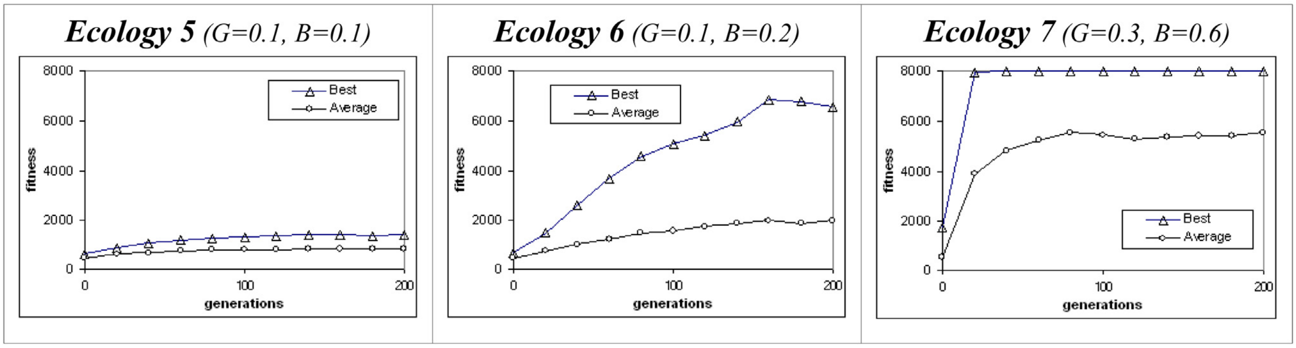


Fig. 8. Fitness curves for all ecologies, measured as the mean life length (max. 8000 cycles) for each generation over 200 generations. Data refer to robots with a metabolism (Study 2).

token when one is already satiated does not improve one's chances of survival, whereas eating the same food token is crucial when one is close to starvation.

In the face of these more complex evolutionary pressures, the robots adaptive success is tied even more closely to the amount of food present in the environment, as Fig. 8 shows (data refers to mean life length for each generation). Applying the exclusion criteria for outliers used in Study 1, the samples retained for analysis in each ecology were $N^{ec5} = 882$, $N^{ec6} = 845$, $N^{ec7} = 819$.

In the baseline condition (Ecology 5), where at any given time the food present in the environment is barely sufficient to sustain the robots ($\text{OAE}_6 = 1.0$), the fitness curve grows very slowly across generations, for both the average and best individuals. However, an increase of 50% in OAE (Ecology 6) dramatically improves adaptive success, and evolving optimal responses becomes fairly trivial when the OAE is especially abundant (Ecology 7).

In fact, it is important to assess how much easier it is to reach evolutionary success in richer environments, and how this might impact on the quality of the behavioral strategies evolved in these environments. We used data on life length and intake rates to gather some insights in that direction, as summarized in Fig. 9. Looking at mean values for each seed ($N = 30$), richer ecologies result in a longer life-span ($F(2, 27) = 167.72$, $p < .0001$, all pairwise comparisons $p < .01$, except Ec5 vs. Ec6 $p < .05$) and in a superior energy-based intake rate, that is, the amount of energy obtained per time unit ($F(2, 27) = 31.28$, $p < .0001$, all pairwise comparisons $p < .01$). But if we look at the token-based intake rate, that is, the number of food tokens eaten per time unit, we find that robots from the richer ecology (Ecology 7) are less effective in finding food than robots from poorer ecologies ($F(2, 27) = 29.41$, $p < .0001$, all pairwise comparisons $p < .01$). This shows that their longer life-span and higher energetic gains are very much due to the better quality of their ecology, rather than to superior behavioral strategies.

In looking at how much energy robots of the last generation are able to consume during their life, it is worth observing them when the energy level is kept fixed (at 1.0, 0.5, or 0.1), as opposed to the standard condition in which they evolved, where energy decreases as a function of time and increases whenever food is consumed. This manipulation is roughly equivalent to performing a field experiment (one that would be impossible to realize in natural organisms), as opposed to the controlled experiments performed in the lab. It provides valuable information on specific behavioral strategies evolved to cope with different levels of energy, which correspond to different levels of need for food: extreme when the energy level is low, negligible or null when the energy level is high or maximal.

As Fig. 10 shows, the amount of food robots are able to consume positively correlates with the abundance of food in their ecology, which of course is no surprise; nor it is surprising that, by keeping fixed their energy level, robots are able to consume more food, since they are *de facto* immortal. But what is interesting is the interaction between these two effects: robots that have no need for food (fixed energy level at 1.0) are less inclined to eat than robots with lower energy level in all ecologies, and most clearly in Ecology 7, where food tokens have greater value. In that ecology, satiated robots do not just express a mild lack of interest for food, as it happens in other ecologies, but actually practice a deliberate avoidance strategy: whenever they see food, they tend to avoid consuming it – witness the fact that 21% of robots in that ecology managed to consume zero tokens of food over 8000 cycles. Notice that this cannot be interpreted as inefficient behavior, since the same robots are very good at consuming food whenever their energy level is lower

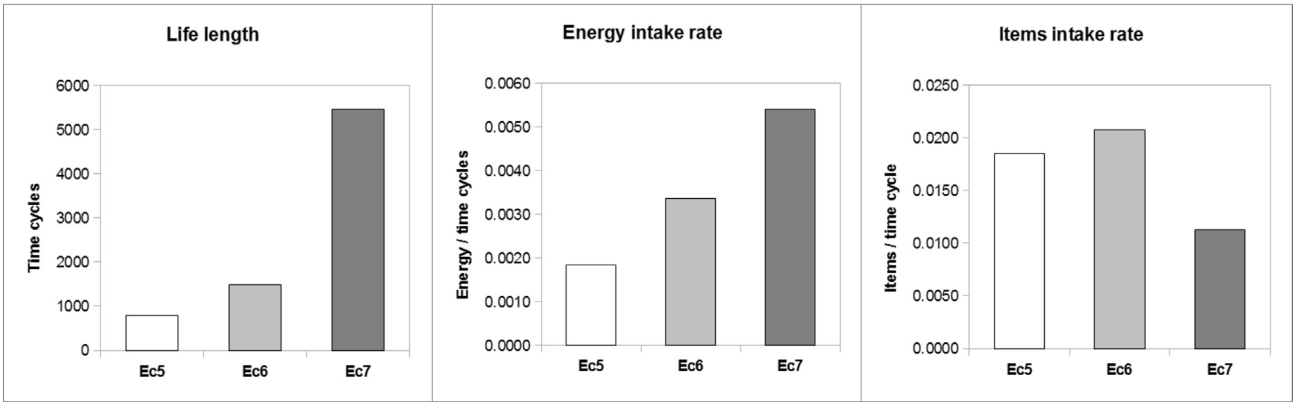


Fig. 9. Mean life length (left panel), energy intake rate (center panel), and token intake rate (right panel) in robots with a metabolism (Study 2), measured at the last generation with free energy levels.

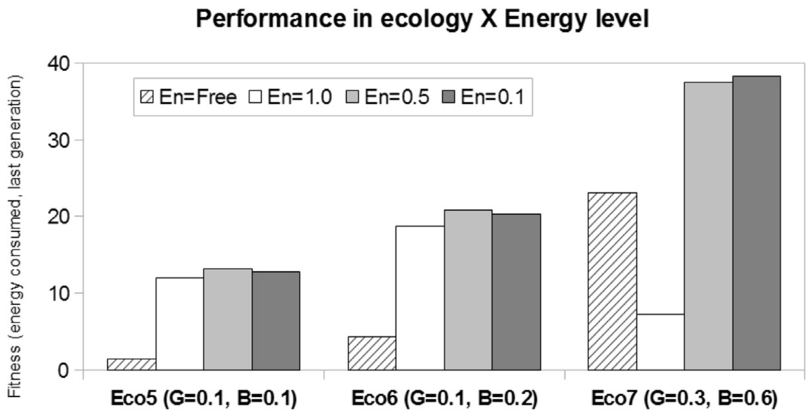


Fig. 10. Overall energy consumed in the last generation for each ecology, either allowing the energy level to vary as usual (En = Free) or keeping it fixed (to 1.0, 0.5, or 0.1). Data refer to robots with a metabolism (Study 2).

than 1. A mixed 3×4 ANOVA performed on mean amount of energy consumed at last generation for each seed ($N = 30$), with Ecology varying between-subjects and energy level (free/1.0/0.5/0.1) as within-subjects variable, confirmed the statistical significance of both main effects and their interaction ($p < .0001$ in all cases).

In terms of choices between different food types, robots with a metabolism develop the same preference for the richer food already observed in the absence of energetic consumption (see Fig. 11): whereas robots in Ecology 5 ($G = 0.1$, $B = 0.1$) are substantially indifferent between tokens, a clear preference for the better food is developed when blue tokens are worth twice as much as the green ones (Ecologies 6 and 7). A between-subjects ANOVA over the mean percentage of blue tokens consumed at the last generation across seeds ($N = 30$) confirms the significance of the effect: $F(2, 18) = 5.48$, $p = .01$. Pairwise comparisons (Tukey HSD test) shows that, while the preference for blue tokens is significantly higher in Ecologies 6 and 7 than in Ecology 5 ($p < .05$), there is no significant difference in preference between Ecologies 6 and 7. This is also consistent with what was observed in robots without a metabolism (Study 1), and suggests that what determines preference between food types is the ratio of their values, not their difference – even in robots with a metabolism. Finally, no substantial interaction between food preference and energy level was observed (robots maintain stable preferences across all energy levels), except in Ecology 7: here a within-subjects ANOVA ($N = 10$) shows a weakly significant effect of energy level on food preference ($F(3, 27) = 3.83$, $p < .05$), with the preference for the better food being lower for satiated robots, as opposed to both robots with free energy levels and robots with fixed energy at 0.5 ($p < .05$ in both cases). This suggests that, when food tokens are especially large, satiated robots not only try to avoid eating them, but also weaken their preference for the larger ones.

3.3. Laboratory tests: method and results

The laboratory tasks were the same as in Study 1 (single option, and binary choice varying the distance of the blue reward), but for each task robots were tested either with a free energy level (set to 1 at the beginning of testing, then let free to decrease as usual), or with their energy level kept constant during testing at one of three values: 1, 0.5, or 0.1. We predicted that this manipulation would reveal energy-dependent behavioral strategies, determined by the ecology in which robots were evolved.

3.3.1. Single option

Using rapidity of approach as a measure of the value robots assign to rewards, we observed in all ecologies the same pattern of devaluation noted in Study 1 (see Fig. 12, data with free energy levels). However, while in Ecologies 5 and 6 this effect was virtually identical to the one observed in robots without metabolism, in Ecology 7 the presence of a metabolism determined the emergence of a novel pattern, showing that metabolic robots evolved in a rich environment move much more slowly toward food than robots evolved in less wealthy ecologies.

To better appreciate the significance of these findings, we performed a 4×2 ANOVA with repeated measures on both factors (see Table 2 for a summary of results). This confirmed that rapidity of approach decreases as a function of distance, while reward type has no impact on rapidity only when both options are equally

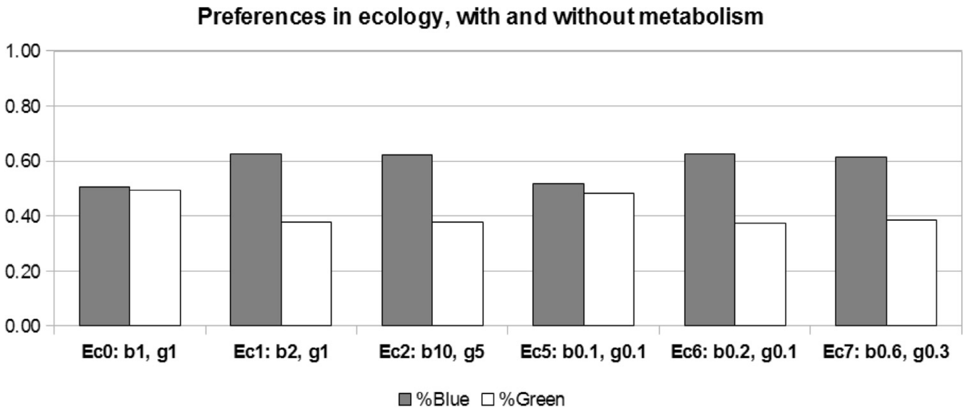


Fig. 11. Percentage of food tokens of each type consumed by robots of the last generation in their ecology, with free energy levels. The comparison shows similar patterns both for robots without a metabolism (Ecologies 0–2, Study 1) and for robots with a metabolism (Ecologies 5–7, Study 2).

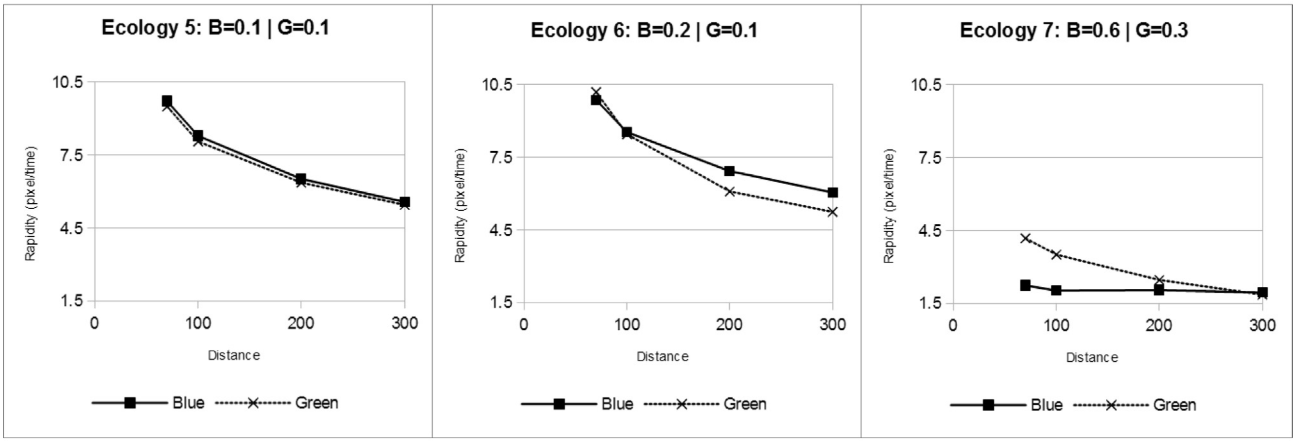


Fig. 12. Rapidity of approach (measured in pixel/time step) to a single token of each type, placed in front of the robot at 70, 100, 200, or 300 px**.** Data refer to robots with a metabolism (Study 2), tested with free energy levels.

valuable (Ecology 5). But when the value of the rewards is different, its effect on rapidity depends on the evolutionary environment: robots evolved in a poorer environment (Ecology 6) move faster toward larger rewards, when these are placed far away; on the contrary, robots evolved in a richer environment (Ecology 7) move faster toward smaller rewards, when these are placed near them.

The behavior exhibited by the robots of Ecology 7 may seem puzzling, but it has a natural explanation. As we discussed in the previous section, robots from a rich ecology do not exhibit a very high intake rate in terms of number of tokens eaten, because they do not need to: they live long, successful lives without having to look for food all the time, thanks to the relative abundance of resources provided in their environment. This might well result in a tendency to move more slowly toward food: there is literally no need to run, if resources are plentiful. However, it is an open question whether this ecological abundance dulls the selection process and thus produces less sharp adaptions, or if evolution results anyway in very good strategies, which simply do not need to be employed constantly. The evidence seems to point toward the latter interpretation: since robots start testing at full energy, when rewards as large as those present in Ecology 7 ($B = 0.6$ and $G = 0.3$) are close to them, consuming such rewards would be nearly useless, because most of their energetic value would be lost (*i.e.*, it would exceed the maximum capacity of the robot's energy level). So these data could be taken to suggest that well adapted robots refrain from consuming food until its energetic value is fully needed: since in Ecology 7 the larger reward is fully needed only with an energy level of 0.4 or lower, whereas the green reward is fully needed already at an energy level of 0.7 or less, robots will take more time to reach the larger reward than the smaller ones, especially when they are close to their current location.

This interpretation is confirmed by looking at how the devaluation process is affected by keeping fixed the energy level in Ecology 7 – and only there, because no effect of energy level is observed on the single option performance in Ecologies 5 and 6. In contrast, robots from the richest ecology have very different performances depending on whether their energy level is kept fixed at 1, 0.5, or 0.1; as Fig. 13 shows, their rapidity on all distances is inversely proportional to their energy level, and the faster reaction times to the lesser reward also tend to disappear as the energy level drops; moreover, whereas robots at full energy make a lot of null trials in the single option task, and more so for the better reward, null trials are very rare and evenly distributed between rewards with lower energy levels (factorial ANOVA on mean % of null trials for all seeds, Energy levels *X* Reward type: effect of energy level, $F(2, 18) = 62.15$, $p < .0001$, no significant effect of reward type, significant interaction, $F(2, 18) = 8.93$, $p < .01$). All of this suggests that in Ecology 7 motivation to move quickly toward a reward, and especially a large one, is low only when there is in fact no need to do so, because the robot is already satiated and would not benefit from eating more (highly energetic) food.

3.3.2. Binary choice, distancing the better option

In this test, as it was done in Study 1, the robot was first presented with two options, one green token vs. one blue token, both placed at the same distance (100 pixels), on two sides of its visual field, and with an inter-token distance of 100 pixels. Again we observed a preference shift from the larger reward to the smaller one, when the distance to the former was increased – except in Ecology 5, where both rewards have the same value. However, while in robots without metabolism the magnitude of the rewards did not affect at all this process (see Fig. 6, in particular the panels for Ecology 1 and 2), this factor has a very significant effect in robots with metabolic constraints (see Fig. 14, in particular the panels for Ecology 6 and 7).

Keeping the ratio between rewards constant, we observe that robots evolved in a more energetic environment (Ecology 7) become indifferent between the two rewards sooner than robots adapted to less energetic food tokens (Ecology 6). Whereas a blue token loses half of its value after approximately 170 px in Ecology 7, it takes about 280 px of distance to replicate the same devaluation effect in Ecology 6. Translated in discounting terms, this means that the discounting rate of the larger reward is much higher in rich ecologies than in poor ones. This might seem counter-intuitive, but

Table 2
ANOVA results for Distance (70/100/200/300 px) *X* Reward type (Green/Blue), both factors within-subjects, mean rapidity values for each seed, $N = 10$. Italics indicate statistical significance, $^{*}p < .01$, $^{**}p < .001$, $^{***}p < .0001$. Data refer to robots with a metabolism (Study 2), tested with free energy levels.

	Distance	Reward type	Interaction
Ec5: g0.1, b0.1	$^{***}F(3,27) = 308.61$, $p < .0001$	$F(1,9) = 1.17$, $p = 0.309$	$F(3,27) = 0.28$, $p = 0.835$
Ec6: g0.1, b0.2	$^{***}F(3,27) = 374.28$, $p < .0001$	*, $F(1,9) = 21.56$, $p = 0.0012$	$^{***}F(3,27) = 28.17$, $p < .0001$
Ec7: g0.3, b0.6	$^{***}F(3,27) = 19.83$, $p < .0001$	*, $F(1,9) = 13.36$, $p = 0.0053$	$^{***}F(3,27) = 14.41$, $p < .0001$

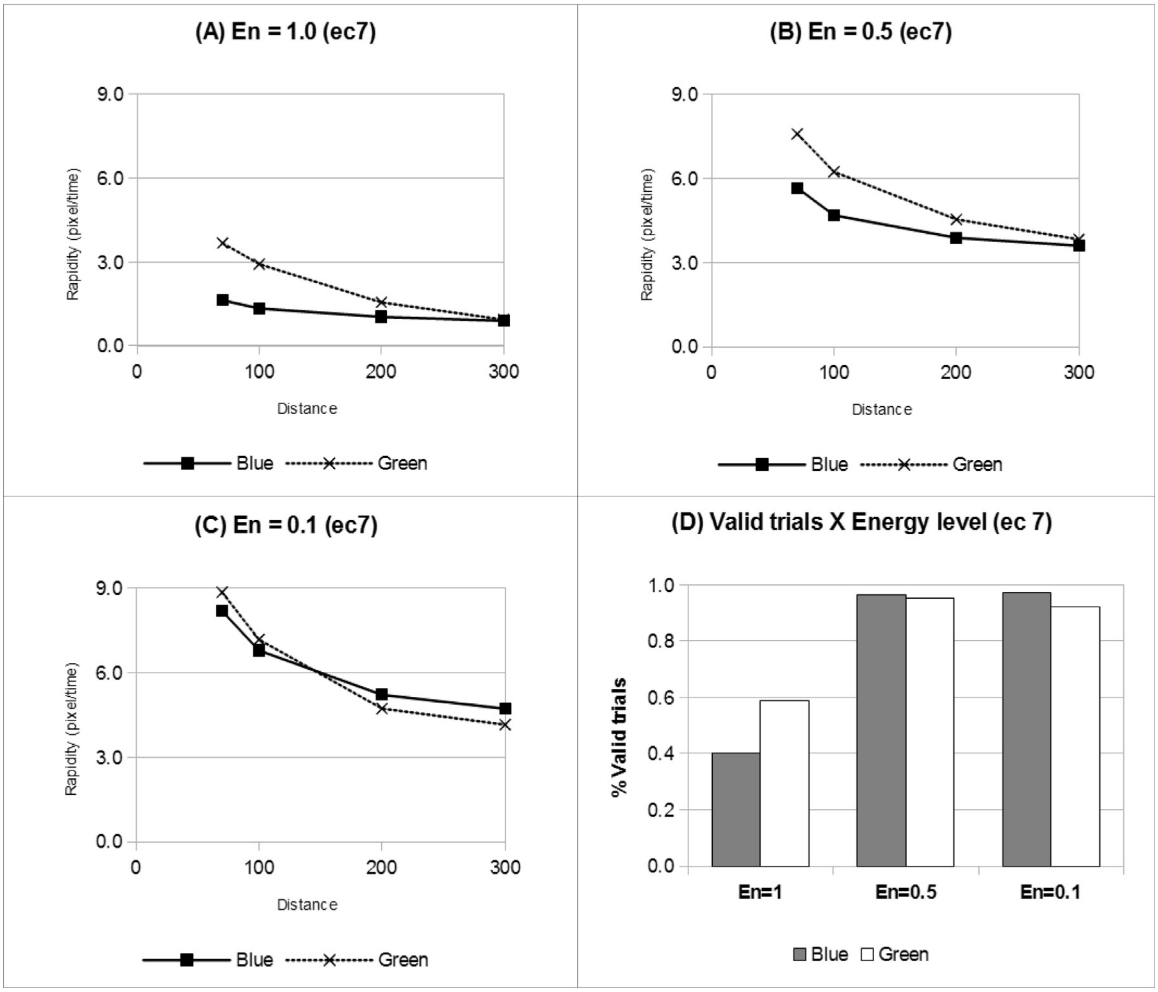


Fig. 13. Single choice task: effects of energy levels on the devaluation process in terms of rapidity of approach (panels A–C) and on percentage of valid trials (panel D) in Ecology 7 ($G = 0.3$, $B = 0.6$), robots with metabolism.

in fact it is just a further manifestation of the emergence of energy-specific strategies in Ecology 7. While in all other ecologies robots maintain in this task the same choice behavior regardless of the energetic level to which they are set during testing, we notice a clear effect of the energy level on choice performance in Ecology 7, as shown in Fig. 15.

In Ecology 7, satiated robots clearly tend to avoid choosing either option (about 40% of null trials), and if they do take one food, they are indifferent between them – rightly so, since a satiated robot does not gain any energy from either option. In contrast, robots with a lower energetic level pick one of the reward almost always (less than 1.3% of null trials), and exhibit a clear preference for the larger one, until it is moved away by a considerable distance – either 200 or 260 px, depending on the energetic level. Interestingly, the preference for the larger reward is less marked but more stable for robots with very low energy (0.1), which again is understandable:

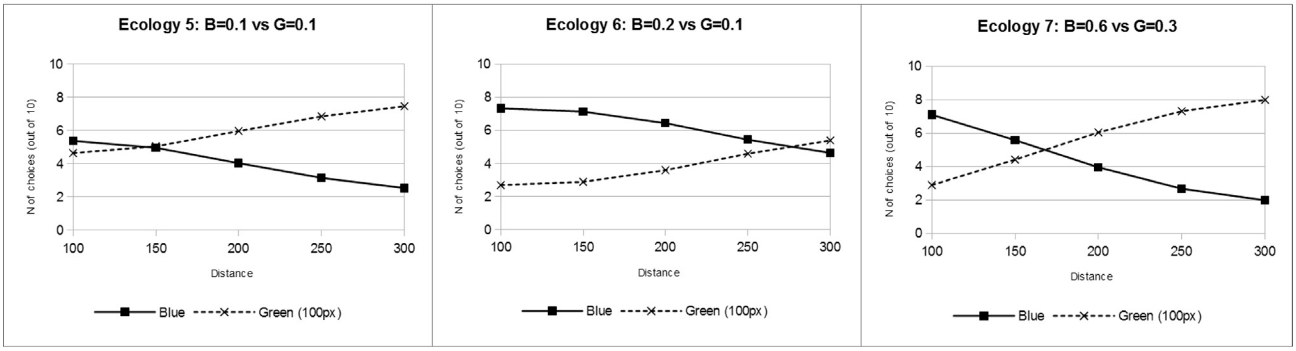


Fig. 14. Choice patterns in a binary choice when the distance from the blue token progressively increases and the green token is always at 100 pixels from the individual: mean number of tokens chosen by each individual out of 10 trials. Data refer to the robots with metabolism (Study 2), tested with free energy levels.

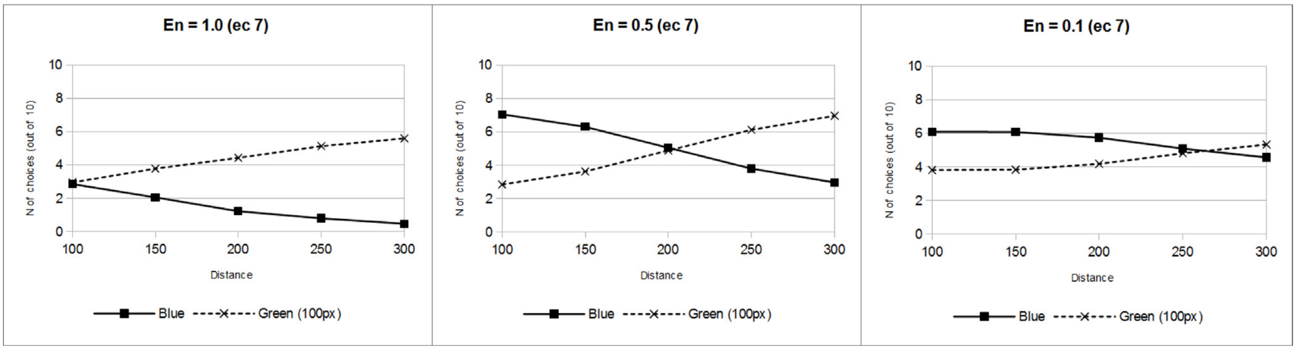


Fig. 15. Effects of energy levels (1.0, 0.5, or 0.1) on choice patterns in a binary choice in Ecology 7 ($G = 0.3$, $B = 0.6$), robots with metabolism. Data refer to mean number of token chosen by each individual out of 10 trials.

an individual facing the prospect of starvation cannot afford to be too choosy.

3.4. Discussion and comparison with natural organisms

The main finding in this second study concerns the emergence of clear differences between adaptive responses evolved in different ecologies, as soon as metabolism becomes a relevant variable. In particular, we notice that living in an environment with very large food rewards (Ecology 7) produces energy-dependent behavioral strategies, whereas individuals adapted to smaller food rewards (Ecologies 5 and 6) tend to apply a uniform choice strategy, regardless of their current energy level: go for the larger and closer option as quickly as possible. The lack of differentiation observed in poorer ecologies is understandable: after all, over-eating is rarely an issue there, whereas being very quick in gathering food is a highly viable trait.

What is remarkable, however, is the emergence of thrift in richer ecologies: in Ecology 7, satiated robots clearly refrain from consuming food, take their time in moving toward it (plausibly, to allow their energy level to drop up to the point when eating it will no longer be wasteful), and even exhibit a preference for the smaller option in these circumstances (that is, for the choice that minimizes waste). This pattern is so striking because here there are very weak ecological pressures toward it: over-eating is not harmful for the individual (contrary to what happens in most natural organisms), and resources are re-introduced in the ecology as soon as they are consumed – so scarcity is not really an issue. In fact, the only scarce commodity that robots might be trying to preserve with this strategy is information: if a robot manages to refrain from consuming a food until there is need for its full energetic value, meanwhile keeping that food in sight (since robots have no memory, in these studies), then that robot will be able to eat the food at just the right time – as if it was looking for some sort of "optimal bout". Indeed, individual analysis of behavior (trajectories) provides tentative confirmation of this hypothesis: satiated robots in Ecology 7 tend to circle around food tokens, until it is time to eat them. The fact that thrifty tendencies evolve even in the absence of specific pressures against over-indulgence might suggest a preadaptation pattern (Bock, 1959): a trait originally evolved to optimize search efficiency might later become crucial to avoid over-eating and squandering, once these behaviors start affecting negatively one's fitness.

Another relevant result concerns the comparison of strategies evolved in different ecologies, especially in terms of effectiveness. Obviously, each ecology produces different behavioral patterns, each of them well adapted to the relevant environmental constraints. Still, comparison across ecologies raises two intriguing issues: (i) How resilient is the strategy evolved in that particular ecology? That is, how effective would that behavior be, if the ecological parameters were to change more or less suddenly (*e.g.*, due to famine, population increase, colonization)? (ii) How much is evolutionary success an indication of the evolution of truly smart strategies, and how much is instead due to relative "environmental abundance"? Obviously, richer ecologies are easier to navigate, but this might have negative effects too, by promoting sloppier behavioral strategies.

Our data indeed suggest that robots evolved in richer ecologies are less effective in their ability to consume food (token intake rate, Fig. 9, right panel), in spite of their longer lives and superior energy consumption. This in turn does not depend on any general "dumbness", but rather derives from their energy-dependent behavior: their thrifty strategies avoid consuming food on a full stomach, and their energy-dense environment allows them to remain constantly satiated without consuming a particular high number of food tokens. This is perfectly fine in their native environment, but it is likely to limit the resilience of their behavioral strategy. What would happen to these robots, if their ecology abruptly became less rich in highly energetic food? Their refusal to look for food while satiated would no longer be viable: even though they would start effectively looking for food as soon as their energy level dropped, this might still be too late–especially if they were facing the competition of another "tribe" used to more demanding environments, like the robots from Ecology 5 or 6. This suggests that robots from such "tougher neighborhoods" might outperform their more privileged cousins, if they were to colonize their environment. The fact that experimental evolutionary robotics allows to test similar predictions (*e.g.*, testing robots in ecologies different from those where they evolved) is a further advantage of this methodology.

Finally, in terms of temporal preferences, we observed the same form of delay aversion already noted in Study 1, again measured as distance aversion. However, the presence of a metabolism produced a novel result also in this respect: whereas in Study 1 no difference in discounting was observed, whenever the reward ratio was kept constant (and excluding food density manipulation, as in Ecology 4), here we see clear differences between Ecologies 6 and 7, although they share the same reward ratio ($0.2/01 = 0.6/03 = 2$) and ecological food density (10 tokens). In particular, robots move much more slowly towards rewards, and become indifferent between them much sooner in Ecology 7 rather than in Ecology 6: both results suggest that food tokens loose their value more rapidly in the richer ecology, as a function of distance. This is not surprising: where abundance of food is the rule, distance becomes a lesser concern, especially if fitness is assessed in terms of life length. Even without running toward food, a robot is likely to eat enough to live long, thus maximizing its chances of reproductive success. Again, this should not necessarily be taken as a violation of the magnitude effect frequently reported for humans, where larger rewards tend to be discounted less steeply than smaller ones (Kirby, 1997; Green et al., 1997, 1999). Whereas that result refers to different magnitudes of reward within the same ecology, here we are comparing magnitudes across ecologies, and we are not aware of any empirical data on magnitude effects of the latter type.

4. General discussion and conclusions

With respect to intertemporal choices, we succeeded in replicating two basic behavioral patterns observed in all species tested so far: a decay of reward value as a function of distance (to wit, delay), and the emergence of a preference for the smaller and closer option, if the larger one is placed too far away from the robot. Both behaviors are indicative of delay aversion. Current data do not yet allow to establish what reward devaluation function (*e.g.*, exponential, hyperbolic, quasi-hyperbolic, sub-additive) or choice strategy (*e.g.*, the similarity based procedure proposed by Rubinstein, 2003) would better account for reward devaluation in these robots, but it would be relatively easy to gather further evidence to settle this point: for instance, by determining further indifference points in the asymmetric choice task, or by measuring rapidity of approach at different distances in the single option task. Moreover, it would be interesting to check whether time-dependent preference reversal (Strotz, 1956; Ainslie, 2001) is observed in these robots, and under what ecological conditions. This would require taking a choice problem where the robot has a preference for the smaller option (*e.g.*, in Ecology 7, "blue at 200 px vs. green at 100 px"), and then progressively increase the distance from both options (*e.g.*, in Ecology 7, "blue at 250 px vs. green at 150 px", "blue at 300 px vs. green at 200 px", *etc.*), making sure that both options remain within the perceptual field of the robot: if at some point the robot switches its choice to the larger reward, this would reveal a time-dependent preference reversal, and thus would suggest that an exponential model cannot account for reward devaluation in that population. The fact that even extremely simple robots show clear signs of delay aversion, and thus can be further tested in more sophisticated tasks, indicates the universality of delay aversion, and the fecundity of this method.

Among the potential advantages of using experimental evolutionary robotics for the analysis of intertemporal choice (see Section 1), this study mainly highlighted two of them: full observability of behaviors, both in the ecology and in the lab, and full control on all relevant parameters. We also made use of non-deterministic responses, although without putting too much emphasis on this point: in all laboratory tests, individual responses were averaged over 10 trials for each condition, and the results clearly show that robots vary in how they respond to the same set of stimuli – as natural organisms also do. Two other advantages of experimental evolutionary robotics, however, have not been exploited in this study: neurocomputational transparency, *i.e.*, looking at what activation patterns in the neural control system are responsible for the observed behaviors, and individual differences, *i.e.*, analyzing how different individuals cope with the same ecological or experimental problem. Length constraints prevented us from dealing with these results in this paper: but a very convenient feature of this methodology is that all data are stored in the system, and could be accessed and analyzed at any time – by us, as well as by other researchers.[7]

As expected (and intended), our work also showcased all the various limitations of this methodology: most notably, the high level of abstraction from naturalistic data, the ensuing arbitrariness of several design decisions, and the need to start small in exploring new issues. Regarding abstraction, there is no doubt that our robots are much simpler than any living organism tested so far in intertemporal choice tasks, from every point of view: sophistication of the control system (their neural network is closer to the nervous system of very simple invertebrates, than to the brain of any mammal, not to mention primates), complexity of the environment (a bidimensional square with only two salient objects and no competition from conspecifics or predators), reproductive system (without any kind of sexual reproduction), learning capabilities (null, as opposed to those of most animals), action repertoire (moving around on two wheels, as opposed to the complex locomotion of most organisms), *etc.* However, as noted above, the very fact that such naive organisms evolve clear patterns of delay aversion, even in the absence of a metabolism (Study 1), indicates how universal this phenomenon is, and points to the key role of opportunity costs in determining its evolution.

As for arbitrariness, some design choices were based on a mixture of practical considerations and naturalistic evidence (*e.g.*, using distance as a proxy of delay), some were guided by specific research questions (*e.g.*, manipulating both ratio and difference of rewards across different ecologies), some were grounded on past experience with this technology (*e.g.*, the number of internal units in the neural network, the size of the environment in relation to the limited perceptual field of the robot, the number of generations used for evolution), and some others were frankly random (*e.g.*, the number of tokens to be placed in the ecology, the maximum life length for each generation, the metabolic rate in Study 2). Did these factors affect our results? Almost certainly so. Could this be avoided? Partially, yes, but not completely. Using naturalistic data to fix as many parameters as possible would greatly help, and this preliminary work was intended to lay the groundwork for such further attempts. Alternatively, testing how variation in a given parameter (*e.g.*, food density) affects behavior would also justify greater confidence in those results that remain constant across such variation. Both methods are worthy of being pursued in future work. However, arbitrariness cannot be completely removed from experimental evolutionary robotics, nor it needs to: as argued at the onset, this method provides optimal results in combination with naturalistic approaches, rather than in competition with them. And arbitrariness, all things considered, is just the (unavoidable) downside of having full control on all the relevant parameters, which is one of the key advantages of experimental evolutionary robotics.

This dovetails nicely on the need for starting small in similar studies: in the Introduction, we listed this as a limitation of our method, because it has severe drawbacks – first and foremost, it is time-consuming. However, this study demonstrates also the benefits of a step-by-step approach: if we had begun by studying robots with a metabolism, it would have been very natural to interpret their delay aversion in terms of energetic concerns – robots devalue distant rewards because reaching them costs more energy. But this would have been misleading, since we now know (thanks to Study 1, with its "unrealistic" non-metabolic robots) that, even in the absence of any energetic constraints, robots develop strong delay aversion – and they do so in ways similar to the patterns exhibited by metabolic robots (except for the interaction with energy-based behavioral strategies). This is an important result, all the more so because it is hardly replicable in naturalistic studies: all known animals do have a metabolism. And yes, of course animals devalue rewards also due to the direct costs of reaching them; but these have to be added to the opportunity costs, and the latter, according to our findings, might be the dominant force behind delay aversion – albeit more evidence is needed to fully substantiate this claim.

In the end, the most exciting thing in presenting a new methodological approach is, obviously, what comes next: that is, what other manipulations such approach allows, once the basic groundwork has been laid out. Even confining our attention to

---

[7] Indeed, requests for our data set are welcome and should be addressed to the corresponding author.

intertemporal choice (but similar considerations apply to any other behavior), several opportunities for future studies come to mind. For instance, Study 2 showed how having a metabolism affects the evolution of delay aversion, but we kept the metabolic rate constant across all ecologies (0.021 energy units consumed per time step). However, it has been suggested (Tobin and Logue, 1994) that metabolic rate might have a direct effect on delay discounting, leading species that consume energy more rapidly to adopt higher discount rates, that is, less willingness to wait for food rewards. Our method permits to gather evidence on that hypothesis: in a first phase, keeping constant all other ecological parameters, it would be easy to verify whether robots with higher metabolic rates in fact evolve stronger delay aversion; in a second phase, this effect could be studied in interaction with other ecological factors, *e.g.*, food density. Indeed, so far we discussed “ecological abundance” as if it depended only on the number, density and value of food rewards, but clearly it is also relative to the metabolic rate of the species inhabiting that environment: “slow-burning” robots might find more than enough energy even in relatively poor ecologies, whereas “fast-burning” robots would be hard pressed for food even in relatively rich contexts.

Another relevant topic easily available for study is the magnitude effect (Kirby, 1997; Green et al., 1997), that is, the tendency to devalue larger rewards less steeply than smaller ones, as a function of time (in our setting, distance). As already noted, the failure of observing this effect in the current study is neither surprising nor worrisome, since here we manipulated reward magnitudes between ecologies, rather than within them. Besides, in all ecologies where rewards sharply differed in value, we did observe a faster devaluation process, in terms of rapidity of approach, for the smaller one (Ecologies 1, 2, 4 and 6), except when energy-dependent behavioral strategies interfered with it (Ecology 7). However, proper study of the magnitude effect in robots would require evolving them in an ecology where more than two food magnitudes are present, *e.g.*, an environment with green food (0.1), blue food (0.2), and red food (0.4): then the prediction would be that (i) in the single option task, rapidity of approach would decrease more steeply for green than for blue than for red food; and (ii) in the asymmetric choice task, robots would exhibit indifference between a closer green food and a farther blue food (0.1 at 100 px vs. 0.2 at N px) at a shorter distance than what is required to become indifferent between a closer blue food and a farther red food (0.2 at 100 px vs. 0.4 at M px, with $N > M$).

Future work should also aim to increase, step by step, the similarity between robots and natural organisms. Two key factors come to mind here: learning and sociality. Several techniques are available for learning in neural networks, both supervised and unsupervised (for a survey, see Hagan et al., 1996; Haykin, 1999), and plugging them in these robots would be unproblematic. This would allow two major advancements: on the one hand, the evolutionary process would now be affected also by learning, so that selection would favor not only individuals that are good at food gathering, but also those that are better at learning useful strategies during their lifetime (on the interaction between evolution and learning in evolutionary robots, see Parisi et al., 1992; Parisi and Nolfi, 1993, 1996; Nolfi et al., 1994; Nolfi and Parisi, 1996); on the other hand, robots would learn also during laboratory tests, exactly as animals do, and this would permit both training manipulations (*e.g.*, conditioning robots to respond to stimuli they never encountered in their ecology, like the operant keys or arbitrary tokens used in intertemporal choices studies with animals; see for instance Ainslie, 1974; Evans et al., 2012) and analysis of how choices change across repeated trials and/or sessions. As for social interaction, in this study it was completely absent, since each robot evolved in its own private world and reproduced asexually. But of course it would be highly interesting to see what happens when robots live socially and are sexually differentiated (for some insight on how gender-specific behaviors evolve under such conditions, see Da Rold et al., 2011). The ecology would then become much more similar to the natural environment of most species, but also much more complex to analyze: among other things, conspecifics would become both competitors (for food) and a scarce resource (for reproduction), not to mention the possibility of strategic interaction (*e.g.*, avoiding over-crowded area to minimize competition, or follow the best feeders). Further complications, such as predators or other species competing for food, would also be easy to introduce, provided one proceeds one step at a time, to avoid loss of explanatory power.

Whatever manipulation one favors, experimental evolutionary robotics always offers a wealth of data, and in this study we only scratched the surface of such a bounty: not only we ignored the activation patterns of the neural network and the fine-grained behavioral strategies of individual robots, as noted, but also focused exclusively on fully evolved robots – that is, individuals of the last generation. But individuals from previous generations are not lost, contrary to what happens with natural organisms: we still have their ecological data, as well as their genotypes, *i.e.*, the weights of their neural network at any given generation, so we can “resurrect” them at will. This allows studying their evolution more minutely, not only in terms of fitness and ecological behavior, but also testing in the lab individuals from previous generations – as if we could test and compare both contemporary and ancestral individuals of a certain species (something which is impossible to do with natural organisms, for obvious reasons). In this way, we will be able to track down the evolution of specific behavioral patterns (preference for the better reward, delay aversion, *etc.*), and whether they were an early adaptation, present since the very first generations, or a relatively late “discovery”. Moreover, the evolutionary history itself can be manipulated, by changing the ecological parameters during evolution: *e.g.*, to simulate some sudden environmental change. This would show how evolution is affected by “ecological shocks”, both in terms of survival rate and fitness, and for the emergence and extinction of individual behaviors and traits.

Finally, more effort should be put in developing unified theories of multiple behaviors: with respect to evolutionary robotics, this is well reflected in the motto “one robot, many phenomena” (Parisi, 2014). Given the inherent complexity of robotics research, the current tendency is to focus only on one phenomenon at a time: after building a convincing robotic model of it, then a different model is proposed to account for another behavior or trait. Thus, the robots evolved to study intertemporal choice are typically not the same evolved to deal with uncertain decisions, seasonal foraging, cooperation, courtship, cultural transmission, etc. Indeed, theoretical fragmentation is widespread also outside of evolutionary robotics: in cognitive science, for instance, we currently have an abundance of detailed models of specific phenomena, while overarching theories of cognition are comparatively scarce and not always well integrated with those models. This situation is at best a stepping stone toward a more comprehensive understanding of behavior, at worst a methodological dead end. Obviously, natural organisms deal with behavioral complexity in a unified fashion: even assuming some form of cognitive modularity, the control system responsible for all behaviors is still one and the same. Evolutionary robotics could do much to approximate a similar situation with artificial organism, but this requires a methodological revolution: instead of building *ad hoc* systems to deal with only a limited set of behavioral problems (as we did in this study), the same evolutionary robots should have to face a variety of behavioral challenges. Granted, this will further complicate the analysis of the interaction between evolution, learning and behavior, but it will also make it much more informative.

In sum, this study, albeit preliminary, demonstrates the viability and fecundity of experimental evolutionary robotics for the study of decision making in general, and intertemporal choice in particular. The fact that several improvements are possible and many opportunities for future research are apparent only adds, rather than detracts, to the interest for this methodology. Still, it is worth insisting on a key epistemological concern: the very power and flexibility of this method also paves the way to its most critical downfall – that is, loosing touch with the natural phenomena that it is supposed to illuminate. Given the high level of freedom and control on a multitude of relevant parameters, it is easy to fall prey of questionable scientific practices, such as making too many arbitrary design decisions and thus ending up with an overwhelming amount of (potentially irrelevant) data. It is precisely to avoid a similar degeneration that experimental evolutionary robotics should strive to take its place among the natural and social sciences, rather than being secluded in relatively isolated and highly specialistic sectors, like artificial life and cognitive robotics. This will require a mutual effort of interdisciplinary adjustment: evolutionary roboticists will have to adapt their methods and present their findings in ways consistent with standard scientific practices outside their own field; and natural and social scientists will have to keep an open mind toward a radically different method of studying exactly the same phenomena they are interested to, with a keen eye for any potential synergy with their own approaches.

Acknowledgments

A preliminary version of this paper was presented at the 1st Conference of the Herbert Simon Society, “Bounded Rationality Updated” (New York, 8–10 April 2013), at the 10th national conference of the Italian Association of Cognitive Sciences (Naples, 19–23 November 2013), and at the APA symposium “Worth Waiting For—The Evolutionary and Developmental Foundations of Self-Control” (Washington, 9 August 2014): we are indebted to the organizers and participants of these events for providing valuable feedback on our work. We are also grateful to the editors and one anonymous reviewer of Behavioural Processes, who offered constructive and cogent criticism on a previous version of the article.

References

Addessi, E., Bellagamba, F., Delfino, A., De Petrillo, F., Focaroli, V., Macchitella, L., Maggiorelli, V., Pace, B., Pecora, G., Rossi, S., Sbaffi, A., Tasselli, M.I., Paglieri, F., 2014. Waiting by mistake: symbolic representation of rewards modulates intertemporal choice in capuchin monkeys preschool children and adult humans. Cognition 130, 428–441.

Addessi, E., Paglieri, F., Focaroli, V., 2011. The ecological rationality of delay tolerance: insights from capuchin monkeys. Cognition 119, 142–147.

Ainslie, G., 1974. Impulse control in pigeons. J. Exp. Anal Behav. 21 (3), 485–489.

Ainslie, G., 1975. Specious reward: a behavioral theory of impulsiveness and impulse control. Psychol. Bull. 82 (4), 463–496.

Ainslie, G., 2001. Breakdown of Will. Cambridge University Press, New York.

Amici, F., Aureli, F., Call, J., 2008. Fission-fusion dynamics, behavioral flexibility: and inhibitory control in primates. Curr. Biol. 18, 1415–1419.

Arbilly, M., Motro, U., Feldman, M.W., Lotem, A., 2011. Evolution of social learning when high expected payoffs are associated with high risk of failure. J. R. Soc. Interface 8, 1604–1615.

Axelrod, R., 1997. The Complexity of Cooperation: Agent-based Models of Competition and Collaboration. Princeton University Press, Princeton.

Berns, Laibson, D., Loewenstein, G., 2007. Intertemporal choice – toward an integrative framework. Trends Cognit. Sci. 11, 482–488.

Bock, W.J., 1959. Preadaptation and multiple evolutionary pathways. Evolution 13 (2), 194–211.

Bongard, J., 2013. Evolutionary robotics. Commun. ACM 56 (8), 74–83.

Brooks, R., 1999. Cambrian Intelligence: The Early History of the New AI. Cambridge MIT Press.

Caligiore, D., Borghi, A.M., Parisi, D., Baldassarre, G., 2010. TRoPICALS: a computational embodied neuroscience model of compatibility effects. Psychol. Rev. 117, 1188–1228.

Chelonis, J.J., King, G., Logue, A.W., Tobin, H., 1994. The effect of variable delays on self control. J. Exp. Anal. Behav. 62, 33–43.

Cheng, K., Peña, J., Porter, M., Irwin, J., 2002. Self-control in honeybees. Psychonomic Bull. Rev. 9, 259–263.

Clark, A., 1997. Being There: Putting Brain, Body and World Together again. MIT Press, Cambridge.

Da Rold, F., Petrosino, G., Parisi, D., 2011. Male and female robots. Adapt. Behav. 19 (5), 317–334.

Doncieux, S., Bredeche, N., Mouret, J.-B., 2011. New Horizons in Evolutionary Robotics. Berlin/Heidelberg, Springer-Verlag.

Duckworth, A., Seligman, M., 2005. Self-discipline outdoes IQ in predicting academic performance of adolescents. Psychol. Sci. 16 (12), 939–944.

Eiben, A.E., Smith, J.E., 2007. Introduction to evolutionary computing. Springer-Verlag, Berlin.

Evans, T., Beran, M., Paglieri, F., Addessi, E., 2012. Delaying gratification for food and tokens in capuchin monkeys (*Cebus apella*) and chimpanzees (*Pan troglodytes*): when quantity is salient symbolic stimuli do not improve performance. Anim. Cognit. 15, 539–548.

Fawcett, T., Fallenstein, B., Higginson, A., Houston, A., Mallpress, D., Trimmer, P., McNamara, J., 2014. The evolution of decision rules in complex environments. Trends Cognit. Sci. 18 (3), 153–161.

Fawcett, T., McNamara, J., Houston, A., 2012. When is it adaptive to be patient?: A general framework for evaluating delayed rewards. Behav. Processes 89, 128–136.

Floreano, D., Keller, L., 2010. Evolution of adaptive behavior in robots by means of Darwinian selection. PLoS Biol. 8 (1), e1000292, http://dx.doi.org/10.1371/journal.pbio.1000292.

Floreano, D., Husbands, P., Nolfi, S., 2008. Evolutionary robotics. In: Siciliano, B., Khatib, O. (Eds.), Springer Handbook of Robotics. Springer, Berlin, pp. 1423–1451.

Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time discounting and time preference: a critical review. J. Econ. Lit. 40, 351–401.

Garon, N.M., Longard, J., Bryson, S.E., Moore, C., 2012. Making decisions about now and later: development of future-oriented self-control. Cognit. Dev. 27 (3), 314–322.

Gigerenzer, G., Goldstein, D., 1996. Reasoning the fast and frugal way: models of bounded rationality. Psychol. Rev. 103 (4), 650–669.

Green, L., Estle, S.J., 2003. Preference reversals with food and water reinforcers in rats. J. Exp. Anal. Behav. 79, 233–242.

Green, L., Fry, A.F., Myerson, J., 1994. Discounting of delayed rewards: a life-span comparison. Psychol. Sci. 5, 33–36.

Green, L., Myerson, J., McFadden, E., 1997. Rate of temporal discounting decreases with amount of reward. Memory Cognit. 25, 715–723.

Green, L., Myerson, J., Ostaszewski, P., 1999. Amount of reward has opposite effects on the discounting of delayed and probabilistic outcomes. J. Exp. Psychol. Learn. Memory Cognit. 25, 418–427.

Hagan, M., Demuth, H., Beale, M., 1996. Neural Network Design. PWS Publishing, Boston.

Hammerstein, P., Stevens, J. (Eds.), 2012. Strüngmann Forum Reports. Cambridge, MIT Press.

Haslam, M., Hernandez-Aguilar, A., Ling, V., Carvalho, S. de la Torre, I., DeStefano, A., Du, A., Hardy, B., Harris, J., Marchant, L., Matsuzawa, T., McGrew, W., Mercader, J., Mora, R., Petraglia, M., Roche, H., Visalberghi, E., Warren, R., 2009. Primate archaeology. Nature 460, 339–344.

Haykin, S., 1999. Neural networks: A comprehensive foundation, 2nd Ed. Prentice-Hall, Upper Saddle River.

Kirby, K., 1997. Bidding on the future: Evidence against normative discounting of delayed rewards. J. Exp. Psychol. Gen. 126, 54–70.

Laibson, D., 1997. Golden eggs and hyperbolic discounting. Quarterly J. Econ. 112, 443–477.

Lawyer, S.R., Williams, S.A., Prihodova, T., Rollins, J.D., Lester, A.C., 2010. Probability and delay discounting of hypothetical sexual outcomes. Behav. Processes 84, 687–692.

Loewenstein, G., 1999. Experimental economics from the vantage-point of behavioural economics. Econ. J. 109, F25–F34.

Logue, A., 1988. Research on self-control: an integrating framework. Behav. Brain Sci. 11, 665–709.

Logue, A.W., Peña-Correal, T.E., 1985. The effect of food deprivation on self control. Behav. Processes 10, 355–368.

Logue, A.W., Forzano, L.B., Ackerman, K.T., 1996. Self-control in children: age preference for reinforcer amount and delay: and language ability. Learn. Motivation 27, 260–277.

Long, J., 2012. Darwin's devices: what evolving robots can teach us about the history of life and the future of technology. Basic Books, New York.

Macy, M.W., Willer, R., 2002. From factors to actors: computational sociology and agent-based modeling. Ann. Rev. Soc. 28, 143–166.

Mazur, J.E., 2007. Species differences between rats and pigeons in choices with probabilistic and delayed reinforcers. Behav. Processes 75, 220–224.

Mischel, W., 1974. Processes in delay of gratification. In: Berkowitz, L. (Ed.), Advances in Experimental Social Psychology, Vol. 7. Academic Press, New York, pp. 249–292.

Mischel, W., Shoda, Y., Rodriguez, M.L., 1989. Delay of gratification in children. Science 244, 933–938.

Moffitt, T., Arseneault, L., Belsky, D., Dickson, N., Hancox, R., Harrington, H., Houts, R., Poulton, R., Roberts, B., Ross, S., Sears, M., Thomson, M., Caspi, A., 2011. A gradient of childhood self-control predicts health, wealth, and public safety. Proc. Natl. Acad. Sci. 108, 2693–2698.

Mondada, F., Bonani, M., Raemy, X., Pugh, J., Cianci, C., Klaptocz, A., Magnenat, S., Zufferey, J.C., Floreano, D., Martinoli, A., 2009. The epuck, a robot designed for education in engineering. Proc. 9th Conf. Autonom. Robot Syst. Comp. Vol. 1, 59–65, Castelo Branco: IPCB.

Mühlhoff, N., Stevens, J.R., Reader, S.M., 2011. Spatial discounting of food and social rewards in guppies (*Poecilia reticulata*). Front. Psychol. 2, 68.

Niv, Y., Joel, D., Meilijson, I., Ruppin, E., 2002. Evolution of reinforcement learning in uncertain environments: a simple explanation for complex foraging behaviors. Adapt. Behav. 10, 5–24.

Nolfi, S., Floreano, D., 2000. Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-organizing Machines. MIT Press, Cambridge.

Nolfi, S., Parisi, D., 1996. Learning to adapt to changing environments in evolving neural networks. Adapt. Behav. 5, 75–98.

Nolfi, S., Elman, J.L., Parisi, D., 1994. Learning and evolution in neural networks. Adapt. Behav. 3, 5–28.

Noser, R., Byrne, R., 2007. Travel routes and planning of visits to out-of- sight resources in wild chacma baboons *Papio ursinus*. Anim. Behav. 73, 257–266.

Paglieri, F., 2013. The costs of delay: waiting versus postponing in intertemporal choice. J. Exp. Anal. Behav. 99 (3), 362–377.

Paglieri, F., Addessi, E., De Petrillo, F., Laviola, G., Mirolli, M., Parisi, D., Petrosino, G., Ventricelli, M., Zoratto, F., Adriani, W., 2014. Nonhuman gamblers: lessons from rodents, primates, and robots. Front. Behav. Neurosci. 8, 33, http://dx.doi.org/10.3389/fnbeh.2014.00033.

Paglieri, F., Addessi, E., Sbaffi, A., Tasselli, M.I., Delfino, A., 2015. Is it patience or motivation? On motivational confounds in intertemporal choice tasks. J. Exp. Anal. Behav. 103 (1), 196–217.

Paglieri, F., Borghi, A., Colzato, L., Hommel, B., Scorolli, C., 2013a. Heaven can wait: how religion modulates temporal discounting. Psychol. Res. 77, 738–747.

Paglieri, F., Focaroli, V., Bramlett, J., Tierno, V., McIntyre, J., Addessi, E., Evans, T., Beran, M., 2013b. The hybrid delay task: Can capuchin monkeys (*Cebus apella*) sustain a delay after an initial choice to do so? Behav. Processes 94, 45–54.

Parisi, D., 2014. Future Robots: Towards a Robotic Science of Human Beings. Benjamins, Amsterdam.

Parisi, D., Nolfi, S., 1996. The influence of learning on evolution. In: Belew, R.K., Mitchell, M. (Eds.), Adaptive Individuals in Evolving Populations. Reading: Addison-Wesley, pp. 419–428.

Pfeifer, R., Bongard, J., 2006. How the Body Shapes the Way We Think: A New View of Intelligence. MIT Press, Cambridge.

Rosati, A., Stevens, J., Hare, B., Hauser, M., 2007. The evolutionary origins of human patience: Temporal preferences in chimpanzees, bonobos and human adults. Curr. Biol. 17, 1663–1668.

Rubinstein, A., 2003. Economics and psychology? The case of hyperbolic discounting. Int. Econ. Rev. 44 (4), 1207–1216.

Saglimbeni, F., Parisi, D., 2011. Input from the external environment and input from within the body. In: Kampis, I. (Ed.), Advances in Artificial Life. Darwin meets von Neumann, Part I. Springer, Berlin, pp. 148–155.

Schwarz, J.C., Schrager, J.B., Lyons, A.E., 1983. Delay of gratification by preschoolers: evidence for the validity of the choice paradigm. Child Dev. 54, 620–625.

Simon, H.A., 1956. Rational choice and the structure of the environment. Psychol. Rev. 63, 129–138.

Squazzoni, F., 2012. Agent-based Computational Sociology. John Wiley & Sons, Chichester.

Starmer, C., 1999. Experiments in economics: should we trust the dismal scientists in white coats? J. Econ. Method. 6 (1), 1–30.

Stephens, D., Anderson, D., 2001. The adaptive value of preference for immediacy: when shortsighted rules have farsighted consequences. Behav. Ecol. 12, 330–339.

Stevens, J., Stephens, D., 2010. The adaptive nature of impulsivity. In: Madden, G., Bickel, W. (Eds.), Impulsivity: The Behavioral and Neurological Science of Discounting. APA, Washington, pp. 361–387.

Stevens, J., Hallinan, E., Hauser, M., 2005a. The ecology and evolution of patience in two New World primates. Biol. Lett. 1, 223–226.

Stevens, J., Rosati, A., Ross, K., Hauser, M., 2005b. Will travel for food: Spatial discounting in two New World monkeys. Curr. Biol. 15, 1855–1860.

Stevens, J.R., Mühlhoff, N., 2012. Intertemporal choice in lemurs. Behav. Processes 89, 121–127.

Strotz, R., 1956. Myopia and inconsistency in dynamic utility maximization. Rev. Econ. Stud. 23, 165–180.

Tangney, J.P., Baumeister, R.F., Boone, A.L., 2004. High self-control predicts good adjustment, less pathology, better grades: and interpersonal success. J. Personality 72, 271–324.

Thompson, C., Barresi, J., Moore, C., 1997. The development of future-oriented prudence and altruism in preschool children. Cognit. Dev. 12, 199–212.

Tobin, H., Logue, A., 1994. Self-control across species (*Columba livia*, *Homo sapiens:* and *Rattus norvegicus*). J. Comp. Psychol. 108, 126–133.

Tobin, H., Chelonis, J.J., Logue, A.W., 1993. Choice in self control paradigms using rats. Psychol. Rec. 43, 441–453.

Tobin, H., Logue, A.W., Chelonis, J.J., Ackerman, K.T., May, J.G., 1996. Self-control in the monkey *Macaca fascicularis*. Anim. Learn. Behav. 24, 168–174.

Todd, P., Gigerenzer, G., 2012. the ABC Research Group. In: Ecological Rationality: Intelligence in the World. Oxford University Press, New York.

van Haaren, F., van Hest, A., van de Poll, N.E., 1988. Self-control in male and female rats. J. Exp. Anal. Behav. 49, 201–211.

Varela, F., Thompson, E., Rosch, E., 1992. The embodied mind: Cognitive science and human experience. MIT Press, Cambridge.

Waibel, M., Floreano, D., Keller, L., 2011. A quantitative test of Hamilton's rule for the evolution of altruism. PLoS Biol. 9 (5), e1000615, http://dx.doi.org/10.1371/journal.pbio.1000615.