

## **Behavioral robustness and the distributed mechanisms hypothesis: lessons from bio-inspired and theoretical biology**

Jose A. Fernandez-Leon, IEEE member\*

### **Abstract**

Theoretical discussions and computational models of bio-inspired embodied and situated agents are presented in this article capturing in simplified form the dynamical essence of robust and adaptive behavior. The general problem of how dynamical coupling between internal control (brain), body, and environment are exploited in the generation of behavior is particularly analyzed. Using evolutionary algorithms based on Evolutionary Robotics methodology to generate the appropriate neural control, four experiments are introduced to support discussions. The *first model* evolves dynamically robust engagements for goal seeking in the presence of neural noise perturbations. The *second model* develops cognitive-behavioral dependencies for minimal-cognitive behavior in dynamically limited agents. The *third one* evolves experience-dependent robust behavior in one-legged agent walking. Finally, the *last model* shows functional dependencies in a mobile-object tracking task. These experiments include a series of structural, sensorimotor, or mutational perturbations, or in the absence of them. Experimental results indicate that neural controls are not sufficient to generate robust behavior in each case, suggesting the absence of internal control ‘ensuring’ robustness. The general observation is that coupling dynamics ‘forces’ evolution to behavioral robustness in whatever dynamical form evolution cares to come up with, but relying on behavioral mechanisms

---

Fecha de recepción del original: 30/09/2010 | Fecha de evaluación del original: 13/10/2010

• Manuscript received 30 September 2010. This work was supported in part by CONICET – The National Research and Technological Council of Argentina (Type-I-Res. 38-08/01/2004), the Programme AlBan, the European Union Programme of High Level Scholarships for Latin America (No. E05D059829AR), and by ‘The Peter Carpenter CCNR DPhil Award’.

J.A.F-L is with the Centre of Computational Neuroscience and Robotics (CCNR), University of Sussex, UK. Phone: ++44 (1273) 87-2948; fax +44 (0)1273 678535; e-mail: jf76@sussex.ac.uk.

that distributes on brain, body, and environment dynamics. Experimental observations provide testable hypothesis that are likely to address in simple organisms in the biological realm, which has some implications for theoretical biology and artificial systems design.

***Index Terms:*** behavioral robustness; animat; evolutionary robotics; genetic algorithm; minimal cognition; transient dynamics.

## I. Introduction

What ‘ensures’ biological robustness? This is one of the unanswered questions facing scientists since von Neumann (1956) [43] noted the complexity of such a problem by opening debates on ‘the synthesis of reliable organisms from unreliable components’. Reliability in this context refers to the ability of artificial or biological organisms to maintain their capacities (functionalities) in a normal situation, as well as under unexpected internal or external factors (or perturbations), which associates to biological robustness research [31]. Despite the lack of a formal definition, robustness usually refers to the continuation of function in the presence of perturbations [29][32]. Robustness is a systemic property commonly attributed to living organisms [40].

Studies in neuroscience and systems biology [1] generally propose organism-centred accounts of robustness. However, the partition between organism and environment is not always helpful for thinking on organisms as ‘highly-interdependent’. This is because studying exclusively internal (e.g. brain-based) mechanisms only focuses on one-third of the potential behavioral interactions between internal control systems, body, and environment, giving special emphasis on the former component (e.g. brains or nervous systems). In fact, internal properties like modularity, decoupling, and redundancy are conventionally thought to be necessary for robustness in systems biology [33]. Structural properties like these may be required to support systemic functionality to certain perturbations between internal control systems and body, but they do not ‘ensure’ robust traits in themselves (see [29], [32], and [33] for complementary discussions). As an example of this last point, neural network models have been used to explore how modularity can lead to more efficient task management [9]. Despite recognized robust properties of most modular neural networks to noisy data [2], a considerably high amount of noise still reduces drastically their filtering capacity. Therefore, *is it enough to see robustness as generated internally in small neural-based organisms?*

Answers to what is required for robustness at behavioral level (behavioral robustness) could guide better scientific descriptions of habits, coherent experience, and adaptation to changing environments, which will follow a small step in the understanding of how the brain of simple organisms structures its internal dynamics underpinning all movements. This article promotes that only understanding organism-environment coupling, behavioral sciences and neuroscientists can understand how the brain controls organisms’ movements for robust traits so that it is possible to better understand what goes wrong after organism failures or damage in artificial or biological contexts, and to develop better ways to deal with associated outcomes.

This article describes selected studies based on Evolutionary Robotics (ER) [24], a methodology from computational neurosciences, to illustrate how the process of natural selection can lead to the evolution of robust and adaptive traits. This work proposes *experiments where agents cannot exclusively rely on internal control (i.e. neurocontrollers) for robust behavior*. As we will see later, agents can still behave coherently despite certain levels and types of perturbation by exploiting *internal noise* (section IV), *situatedness* (section V), *embodiment* (section VI), and *environmental dynamics* (section VII) for robust traits. This capacity of agents (behavioral robustness) is not only the result of their evolutionary processes (see also [20]), but also of their coupled environmental interactions.

ER is a relatively assumption-free methodology compared to other synthetic approaches based on artificial evolution [35][19]. By limiting experimental analyses to some case studies, this article highlights behavioral robustness as a dynamical process, being in any case certainly incomplete if we do not focus on engaged brain-body-environment dynamics. In fact, the described studies show us that behavioral robustness is better understood in the context of agent-environment dynamical couplings, not in terms of internal (neural) mechanisms. Such couplings, however, are not always the full determinants of robustness.

Using ER techniques, this article summarizes some observations from [13] and associated works (i.e. [14], [15], [16], and [17]) about the emergence of behavioral robustness across a broad class of minimal ER models and a variety of perturbations (e.g. mutational and sensorimotor). These models verify the power of ER technique to explore dynamical mechanisms for behavioral robustness in artificial agents as a systemic process, rather than being insured ‘from inside’. The experimental motivations in this article are because works in ER have so far paid relatively little attention to distributed cognition [46] and behavioral robustness [38] in environmentally coupled agents. In particular, this work addresses current discussions on distributed behavioral mechanisms as central to the emergence of cognitive processes and robust behaviors in ER.

The next section discusses the approach promoted in this article and its comparison to what has been reported in associated literature. Challenges and limitations of the proposed approach are also identified for future studies.

## II. A perspective shift of behavioral robustness

The accepted understanding of robust and adaptive behavior is gradually changing from being generated by isolated control mechanisms within organisms towards dynamical process occurring over multiple and distributed systemic components (see [28][10]). However, the word ‘distributed’ in neuroscience and

Artificial Intelligence (AI), for instance, still means distributed within the brain like distributed parallel computation in neural networks. A scientific account of biological robustness in organisms with brains, bodies, and interacting with diverse environments will not arrive fully formed by only looking what happen inside an organism during coupling with the environment. What is needed is an integrative approach combining theory and experiment, and also exploiting the interchange between theoretical biology and computational neuroscience.

Much current works in computational neuroscience focuses on single levels of description of neural systems and certainly neglects the dynamical relations that connect brains, bodies, and environment. Moreover, the growing consensus about the importance of brain-body-environment couplings is still a minority view in several disciplines. These include cognitive psychology, neuroscience, a good part of AI and robotics, and indeed several areas of systems biology. It is to be expected that even for those who agree on the view promoted in this work (i.e. the distribution of mechanisms enabling behavioral robustness), the full implications of it have not fully been drawn, and this work attempts to address the proposed issue. By targeting and overcoming these limitations, this research will deliver new insights into the neural mechanisms underlying robust, yet adaptive behavior and minimal cognitive experience (see also [3][4][6]).

### **III. Artificial evolution as a tool for robustness research**

ER has increased in popularity across the computational neuroscience and robotics research fields, generally in the development of controllers for simulated organisms (agents) by means of Darwinian selection [20]. Using ER, scientists can understand that adaptive and robust behaviors, and finally cognitive experience, arise from complex (non-linear) brain-body-environment coupled interactions. In brief, ER is a method to automatically generate control systems that are comparatively simpler or more efficient than those engineered with other design methods [20]. With ER methodology, the space of solutions explored by artificial evolution can be larger and less constrained than that explored by conventional engineering methods [35]. An ER agent's neurocontrollers can be defined with a simple architecture with no special mechanisms for ensuring adaptive or robust traits (e.g. synaptic plasticity, no ontogenetic development, and no detailed modeling of spike transmission [20]).

A major issue in ER is that agents may use distinctive features of the environment in which they are selected to increase performance, hence leading to a major fitness drop in new environments where these features are lacking. Recently, Floreano and Keller (2010) [20] summarized some examples of experimental evolution 'from

silico' with robotic agents, which verifies the suitability of ER as an evidence-proof methodology. Their ER experiments were mostly based in the context of computer science and engineering (e.g. [36], [35], [23], [22], [24], and [25]). Biologists and cognitive scientists have become interested in ER as a tool to study, for instance, how phenotypes (e.g. neurocontrollers acting on agents' bodies) can be shaped by natural selection which is difficult to address with real organisms. Examples of these interests range from the evolution of cognitive traits based on environmental dynamics [46], the functional role of morphology in coevolving bodies and brains [36], to the development of a methodology to understand robustness in behaving organisms [26][38][42].

This article describes in the next sections the main aspects of four representative examples that illustrate how bio-inspired ER models can benefit our understanding of the emergence of robust behavioral control in simple organisms from a situated, embodied, and dynamical perspective [3]. Agents in the proposed models are controlled by simple neural networks can evolve the ability to show robust goal seeking in the presence of neural noise perturbations, exhibit minimal-cognitive behavior, robust walking, and mobile-object tracking task under a series of structural, sensorimotor, or mutational perturbations.

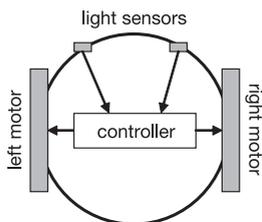
#### **IV. Agents exploiting the effects of neural noise for robust behavior**

Biological robustness is generally discussed in systems biology literature as a by-product of evolution, where robust mechanisms emerge from noisy processes [12]. In a neural context, the incidence of internally generated noise (neural noise) on minimal situated, embodied, and dynamical agents and its effects in the production of adaptive traits is investigated so far in [15][16], where some of the preliminary ideas shown in this section were also evaluated in physical robots in [18].

The model in this section (i.e. robust goal seeking under neural noise perturbations) requires appropriate processing of sensory information and coordinated activation of the agent motor system. Agents were evolved for goal seeking (phototaxis) task and their performance were tested in the presence of sensorimotor and structural perturbations and for different levels of neural noise during agent lifetime. Experiments were conducted in an unlimited (simulated) arena with a two-wheeled agent equipped with two diametrically opposed motors that differentially steer the agent with their output and two frontal light sensors positioned with a separation between sensors of  $47.75^\circ$  (Figure 1). The two sensors were connected to two interneurons that were linked to two output (motor) neurons, each controlling the direction and speed of rotation of one of the wheels. Full

connectivity is used for linking neurons, but only output neurons include self-connections. Left/right symmetry in synaptic weights is not enforced.

The genotypic expression of agents' internal control (neurocontrollers) consisted of a sequence of real-valued parameters encoding all synaptic weights between internal and output neurons, and connections with sensors. Mutations to these evolved parameters were induced as a small vector displacement selected from a Gaussian distributed value in each gene. Artificial evolution was experimented in five independent populations each consisting of 60 individuals [15]. The performance of each agent was evaluated with a fitness function describing the ability of the robot to efficiently seek and approach a light attracting goal. Although the fitness function *did not* specify in what direction the agents could rotate, it can be done in both directions. The best-evolved agents moved in the direction toward targets after evolution despite high, medium, low levels of induced neural noise, or in noiseless conditions. These results and tests in [15] indicate that agents evolved with noise remained robust to sensorimotor and structural perturbations even when neural noise was removed during tests.



**Fig. 1** | Robot layout for exploring robust goal seeking under neural noise perturbations. The agent's circular body, showing motors, and sensors layout which light impinges on the sensors.

Behavioral and perturbation tests for the described model show that behavioral robustness emerges as a by-product of evolution that takes into account not only the environmental characteristics, but also the morphological and mechanical properties of the agents. *Behavioral robustness shown to arise from environmentally coupled dynamics that shape agent behavioral mechanisms during evolution combined with dynamical features of neurocontrollers that promote robust behaviors.* This is possible based on an artificial evolutionary process due to ER generating behavioral adaptation [20]. In particular, the implemented process implicitly selected neural systems that operate in noise-resistant landscapes that were resistant to bifurcation and/or contain dynamical bifurcations [41] that retain goal approaching functionality. In other words, the implemented evolution selects neural control that enables goal approaching despite induced perturbations. This experimental evidence also serves as a transition into the next experiments, which begin discussions on the relation between distributed mechanisms and behavioral robustness.

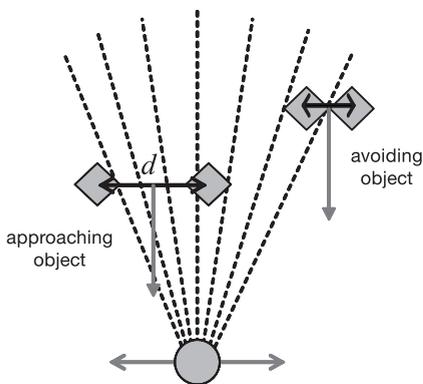
## V. Agents exploiting situatedness for robust behavior

The idea of cognition as enabled through simple internal neural control is investigated by Ziemke et al. (2004) [46], Chiel and Beer (1997) [11], Beer (1995) [3], and Gallagher (2005) [23], among other works, emphasizing the importance of the environment for cognition (see also [6]). The experiment in this section addresses discussions on how situated agents develop coupled engagements and use neurocontroller properties as well as environmental dynamics to sustain behaviors in the presence of sensorimotor perturbations. ‘Situatedness’ (or being situated in the environment) means that agents (artificial organisms or robots) use spatio-temporal situations in their environment directly influence the future behavior of agents [8].

This section continues investigating the emergence of behavioral mechanisms but evolving agents with dynamically limited controllers (monostable agents) and compares them to less limited ones (bistable agents). ‘Dynamically limited’ (i.e., having one or two internal states) here relates to a reduced quantity of steady states that an agent’s controller exhibits when it does not receive stimulus from the environment (see also [41] for formal definitions of mono and bistability). The proposed model requires that agents, despite their limitations, exploit opportunities from brain, body, and environment. This requirement is imposed by the task (a minimal cognitive task [5]) and the evolutionary selecting criteria or fitness function. The fitness function discriminates between apertures wide enough for the agent to pass through (approaching objects) and when the aperture is smaller than agent’s diameter (avoiding objects). This fitness measure assigns near-zero value to incorrect actions and linearly penalizes by misses. Since making the correct decision without hitting the diamonds results in a significantly higher score, this measure also rewards accuracy in agent’s movements. The model has been designed so that the final sensory state of agents is qualitatively the same for both approaching and avoiding behaviors, i.e. without sensory stimuli from the environment. In this way, the model challenges evolution to exploit opportunities from coupled dynamics, given a restricted amount of internal dynamical resources (see for [13] further details). The categorical perception task for which agents should adapt during the evolutionary process requires an agent capacity to differentiate between ‘objects to catch’ and ‘objects to avoid’. The categorization of approaching or avoiding objects is to sense two diamonds that are separated enough to enable the agent body to pass through diamonds or joined, respectively (Figure 2). Each type of object is dropped from above a certain number of trials (63 trials), and the object separation is changed between trials. In every trial, we place the agent at the centre of arena (horizontal axis), and objects are initially located from different starting horizontal places in a fix range relative to agent’s initial position. In each trial, the state of neurons is initialized

to zero. Objects fall down from the top side of the arena to lower positions were the agent moves horizontally (the bottom part of the arena). The distance from agent initial position and the final horizontal separation between the agent and the object describes the agent's behavior (to catch or avoid objects).

The agent has a circular body receiving stimulus from an array of seven sensor rays equally placed from the centre of the agent over an angle of  $\pi/4$  on the agent's top side. Sensors take binary values representing the intersection (incoming signal 1) and no-intersection (incoming signal 0) between a ray and an object. This intersection causes a binary signal from the corresponding sensory node to interneurons. Seven sensory nodes project their outputs to 5 interneurons. These neurons connect to 2 motor neurons that control the horizontal displacement of the agent.



**Fig. 2 |** Schematic representation for exploring a minimal-cognitive behavior in dynamically limited agents. Experimental setup for the categorical perception task. The agent (circle) can move horizontally while objects composed by two diamonds with an adjustable aperture 'd' fall from above.

The network architecture of the neurocontroller is defined as bilaterally symmetric in the connection weights, biases, and time constants. While imposing such symmetry often made trials involving nearly centred objects difficult, it reflects the symmetry of the agent and the task and halves the number of parameters to evolve (see also [4][5]). All neural relevant parameters and synaptic connections are genetically determined. The ER evolutionary process works for this experiment using a population size of 300 neurocontrollers was conducted. The average performance of each individual is maximized by considering the fitness of agents in each test (i.e. trials for each object position), discriminating between apertures wide enough for the agent to pass through (approaching objects) and when the aperture is smaller than agent's diameter (avoiding objects).

This fitness measure assigns near-zero value to incorrect actions and linearly penalizes by misses. Since making the correct decision without hitting the objects results in a significantly higher score, this measure also rewards accuracy in

agent's movements and is similar to the work of Slocum et al. (2000) [39]. The fitness for a trial ( $n$  trials in total) is determined by the absolute horizontal distance between the center of an agent and the falling object when the object is at a random distance in a fix range from the agent's final (vertical) position. By setting the right fitness reward for approaching and avoiding objects, neurocontrollers show monostability or bistability in their autonomous dynamics in the long run. The performance of mono- and bistable populations (20 successfully evolved agents each) was analyzed in [13]. These best-evolved agents were obtained in independent evolutionary experiments.

Results obtained for this experiment suggest that monostable controllers tend to function by actively using body and environmental dynamics more than bistable controllers, which by definition implies a more distributed control system for producing categorical perception. By actively using environmental dynamics, a best-fit monostable agent exhibited approaching or avoiding behavior based on continuously sensing falling objects. The bistable agent, however, moved independently to sensing an avoiding object. This was a result of dynamical switches between autonomous attractors, which was not observed in a best-fit analyzed monostable agent [13]. In this respect, the monostable agent's own position regarding a falling object represented a simplified external memory to exhibit coupled approaching or avoiding behavior that overcomes the absence of further internal states.

The single-state characteristic of monostable agents enabled them to better cope with unpredictable environmental perturbations to which they where not evolved to deal with (e.g. variable amplitude between objects implying ambiguity between objects to approach or to avoid). This robustness was presented but less-observed for bistable agents. Note then that *it is not the neural controller's dynamics itself what sustained behaviors despite perturbations, but the coupled dynamics*. This conclusion reinforces the idea introduced in the first model that behavioral robustness cannot be deduced directly from internal mechanisms themselves, a common belief in literature; rather it can also be observed in dynamics arising from the coupled brain-body-environment.

Summarizing, statistical and dynamical analyses from the proposed experiment suggest a small but significant tendency of better performance by monostable in contrast to bistable agents in the presence of sensorimotor, structural, and mutational perturbations. The difference is based on greater environmental dependence of monostable neurocontrollers to coupled dynamics. However, further studies of the interaction between distributed dynamics and robustness are necessary. This experiment lays the foundation for further experimental work inducing coupled (environmental or body-based) dependencies of behavioral mechanisms as explained in the next experiment.

## VI. Agents exploiting embodiment for robust behavior

Computational simulations with one-legged embodied, situated, and dynamical agents have also been implemented to study the emergence of behavioral mechanism shaped by coupled body interactions (see [14] and [13]). Here, agents have the capacity to change feedback from their bodies to accomplish a one-legged walking task by implementing the evolution of ‘sensory offsets’ [34] (i.e. evolutionary-defined biases that modify incoming body signals). Discussions in this experiment remark the role of embodiment [44][45] in the production of behaviors. This experiment examines whether agents show further dependence to their coupled dynamics than others with no sensory feedback control. The ability to maintain behaviors functional is tested during lifetime with mutational and sensory perturbations after evolution. By ‘functional’ in this context, refers to the capacity of agents to produce rhythmic stepping for the expected walking behavior.

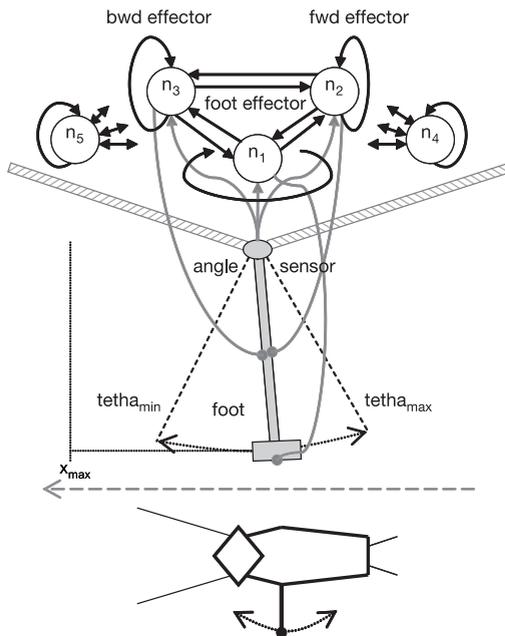
The leg’s set-up is as follows. The leg was only able to generate force over a limited angular range of motion (Figure 3). When a stretched stancing leg lifts its foot (i.e. the position while standing), the leg immediately snaps back to the swing angular limits of  $[-\theta_{\min}, \theta_{\max}]$ . When a stancing leg reached these limits, forward motion comes to an abrupt stop, which according to Beer’s [7] descriptions it produces a loss of postural stability.

During the stance phase, the leg stretches between the body joint and the stationary foot as the body moves with a horizontal distance between the joint and the foot. A stancing leg exceeding the angular range of motion did still provide support, but only within vertical limits. Torque was controlled by two motor neurons (forward or backward neuron effectors onwards). When the foot was up (swing phase), torque produced by effectors serves to swing the leg along an arc relative to the body [7]. For this movement applied a limit constraint with a maximum angular acceleration.

The agent’s leg was controlled by a fully connected five-neuron controller, where three of these neurons were effectors creating the force applied to the agent’s body that generates translational motion. One effector (n1) governed the state of the foot, and the other two generate (n2) clockwise and (n3) counter clockwise torques to the leg’s single joint producing forward and leftward movements. The remaining two units were interneurons with no-specified role in the agent’s leg control. Only effector neurons received a weighted sensory input from the leg’s angle sensor that measures the leg’s angular position in radians. The angle sensor was proportional to the angular deviation of the leg from the perpendicular axis to the long one of the body. The neurocontroller supplied signals specifying what torques should be applied to the leg-body joint. These signals were summed, and depending on the state of the leg’s foot will either move the body (foot down) or rotate the leg about its joint (foot up).

The embedded controller defined none, one, or three mutable sensory attributes or offsets to the agent’s genotype. These offsets were either all the same or different for every sensor-interneuron connection and were added to every sensory signal. The use of offsets means that we can no longer observe a signal of zero to neurons if we disrupt the angle sensor. The model defined a population of 40 genotypes during evolution. The number of trials, run duration, and number of generations was defined as 100, 500 and 1000, respectively. A generation, in this context, was the time it takes to generate the same number of new individuals (or a new population).

Results indicate that an evolved agent with the same evolvable sensory offset for all sensors (or in absence of them) does not necessarily offer an advantage in resilience to sensory perturbations. However, agents using multiple sensory offsets were not considerably affected by mutational perturbations. Results in addition show that agents sensing the environment directly (in the absence of sensory offsets) in a noiseless environment evolve controllers that were highly independent to sensory stimuli, while those agents using multiple sensory offsets develop further dependence to sensory signals.



**Fig. 3 |** Schematic representation for exploring robust walking in one-legged agents. (top) Representation of the agent’s leg configuration for one-leg walking behaviour. Neurons are fully connected including self-connections. Three effectors controls the forward and backward force applied to the leg and the foot for walking. Effectors receive sensory stimuli of the leg angle during the ongoing task. (bottom) The leg model of a simulated insect where the leg can swing about their single joint with the body (figure based on [3]).

In [14] and [13], it is discussed that by evolving with sensory offsets agents would engage more with the environment. Reported experiments with reductions of sensory capacity have shown that agents with offsets develop neurocontrollers

with considerable dependency to incoming feedback. The interactions between the controller, agent's body, and environment give then sufficient conditions to exploit distinct regions of autonomous internal dynamics for solving the one-leg walking task. Walking behavior was consequently not the result of the decoupled internal dynamics of the controller, attributed to a dynamical attractor, or attained to a basin of attraction in the agent's internal milieu. In fact, for this model, *behavior emerged because of the interactions between multiple attractors and basins of attraction at neurocontroller level in a sensor-dependant coupled dynamics* (further details in [14]).

Experimental evidence here leads us to suggest a dynamical systems perspective on behavioral robustness that goes beyond attractors in phase space. The behavior of agents that are evolved with sensory offsets depends not only on where in neural space the state of the neural system operates, but also on the transients to which the internal-system was being driven by sensory signals from its interactions with the environment, nervous system, and agent body.

## **VII. Agents exploiting coupled environmental dynamics for robust behavior**

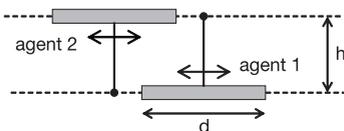
Experiments in previous sections enable discussions on behavioral mechanisms that emerged from agent-environment coupling and agent interaction experience. Experimental evolution has also been used to explore in [17] the role of a similar dynamical process but in a historically dependent task under sensorimotor perturbations. This last experiment proposes a mobile object-tracking task, also known as two-agent interaction task. The basic setup can be described as follows (Figure 4).

Two structurally identical agents face each other in a 1-D circular environment, in which they can move left or right and detect each other by means of a single touch sensor placed at the centre of their body. The task of the agents was to move in the same direction together for as far as possible while continually interacting with each other. This task was made non-trivial by the fact that each agent has to locate the other, coordinate a common direction of movement, and then to move in the same direction, while having minimal sensory input and being perturbed by sensorimotor noise. Tests for behavioral robustness in two kinds of model agents were also proposed in [21]. First, the experiment evolved a population of 10 standard three-node controllers obtained employing the usual ER methodology (see [35] for a description of ER). For a second population (also of size 10) the experiment introduced some additional constraints into the artificial evolution of the neurocontrollers. The idea was to induce local (structural) instability into

the neurocontroller by explicitly requiring that all nodes have excitatory self-connections, while connections between nodes must be inhibitory. Such local instability might facilitate the emergence of behavioral robustness (see [37]).

Imposing structural restrictions into neurocontrollers is not a common practice in ER, but some exploratory investigations revealed that we thereby increase the chances of evolving agents with internal transient trajectories, at least in their decoupled dynamics (e.g. heteroclinic trajectories [41]). These dynamics were relatively stable against high frequency perturbations (e.g. environmental noise of the overall spatio-temporal network), while simultaneously retaining high sensitivity to low frequency change (e.g. the same dynamical pattern may become different over time) [37]. The addition of the described constraints is a necessary but not sufficient precondition for inducing such transient dynamics. This is because the evolutionary process can still lead to a variety of dynamical strategies to produce behaviors in the described model (e.g. using multiple fixed-point attractors). For each type of controller, the model evolved populations composed of 10 genotypes. The number of trials, trial duration, and number of generations were set to 25, 500 and 1000, respectively.

The initial spatial distance between the agents was a control parameter in the described simulations, where agent ‘down’ (or agent 2) always was placed at position 0 and agent ‘up’ (or agent 1) starts at a different position for each trial (25 randomly distributed different positions across a fix range). Since the two agents were started in opposite orientation (‘up’ and ‘down’), it was not possible for the evolutionary algorithm to ‘hard code’ any trivial solution such as having the agents always move to the same direction. The model also added a small perturbation to the motor outputs at each time step drawn from a Gaussian distribution with zero mean and deviation 0.05. The motor noise was applied to the outputs of motor neurons before calculating the difference between them, and before the application of motor gains.



**Fig. 4 |** Schematic representation for exploring mobile-object tracking task. A schematic representation to the two-agent interaction task. The two identical agents are only able to move horizontally facing each other in an unlimited continuous 1-D simulated space. Agents are equipped with a single on/off sensor at the center of their simulated bodies with a separation between agents of ‘h’ arbitrary units.

For every trial, the model computed the fitness as the distance the agents managed to travel together by comparing the location of their final sensory contact with their starting position. The overall fitness for a solution was set to the smallest

score obtained for any of the trial runs during the evaluation. In other words, in order to increase the agents' behavioral robustness against motor noise, sensory switches, and the variations in initial conditions, only the lowest score achieved in any of the trials was chosen as the overall score.

The experimental analysis indicated that agents' coupled dynamics remained in internal transients [41] to maintain the expected tracking behavior in both types of agents. The results also indicated that a highly distributed realization of behavior can be (i) detrimental, if it is mostly based on factors that are necessary for the behavior, or (ii) beneficial, if it is chiefly rooted on factors that are sufficient for the behavior [17]. This difference depends on the effects of perturbations on such a set of factors (e.g. the presence or absence of responsive partners). Accordingly, results suggest that future discussions of distributed cognition should take into account that there are at least two different possible modes of interpreting distributed behavior and that these have a qualitatively different effect on behavioral robustness [17].

Summarizing, this experiment demonstrates that successfully evolved agents can use a combination of both internal dynamics and their history of interaction with the environment to exhibit coupled behavior. Results show that successfully evolved agents can emerge with at least two different modes of exhibiting behavior, one in which inner controls necessarily depend on a set of specific environmental factors (i.e. the presence of a responsive agent) and another for which these factors are equally sufficient for behaviors (i.e. a responsive 'or' non-responsive agent). The described experiment suggests that the first mode is more vulnerable to perturbations on that set, showing that robust agents will exploit behavioral opportunities rooted in external factors only when available. In other words, this experiment provides *a concrete example of how behavioral robustness is a capacity of the coupled agent-environment system, rather than warranted exclusively by agents' internal mechanisms* (see [17] for further details and dynamical analyses).

## VIII. Experimental lessons

The promoted holistic view of robust traits raises difficulties for future experimental studies: *the dynamical mechanisms that produce robustness cannot be studied as isolated parts in most biological systems*. Taking the dynamics of the environment (including body) into account generally makes the study of robustness, a hard problem, even more difficult. Reported experiments in this article consequently have concentrated on minimal models and analyze dynamically contributions of agent-environment to robust and adaptive behavior.

The general observations and the ‘take-home-messages’ proposed in these experiments are listed as follows.

*Not all distribution of agents’ cognitive mechanisms are equally helpful in ensuring behavioral robustness.* The last reported experiment (mobile-object tracking task section) showed us that analyzed agents (one dynamically restricted and other non-restricted) presented behavioral mechanisms distributed differently among the brain-body-environment. Both agents require sensory signals to sustain interaction behavior despite sensorimotor perturbations. However, the dynamically restricted agent exhibits less dependence to responsive partners in comparison to a non-restricted agent. In the presence of sensorimotor and environmental perturbations affecting brain-body-environment interdependencies, a wider distribution (further systemic dependencies) of the behavioral mechanisms seems to be beneficial for robust behavior. This was observed only if not all of the necessary conditions to produce behaviors are affected by significant levels of perturbation. Experiments have also analyzed a non-restricted agent that necessitates the presence of a responsive partner to produce interaction, but does not require other factors to sustain interactive behavior (e.g. sensory or motor noise). Importantly, the conclusion is that *it is not the wider distribution, as such, which is beneficial for behavioral robustness, it is the particular kind of distribution*—a collection of individually sufficient conditions (rather than necessary ones)—*that produces agent’s vulnerability to sensory perturbations* (see [17]). The observation of particular types of environmental dependencies also applies to the monostable agent analyzed in the second reported experiment in this article.

In the second experiment (minimal-cognitive behavior in dynamically limited agents section), the analyzed agent showed a necessary dependence on objects in the environment to produce categorical perception behavior. The requirement of an object in the environment was enforced conceptually: monostable agents cannot produce the necessary internal dynamics without the continuous stimuli from objects in the environment. This stimulus enabled monostable agents to work far from their unique autonomous attractor to produce controls to catch or avoid falling objects. The comparison to bistable agents has showed us that monostable agents were slightly but significantly more robust to a wide set of sensorimotor, morphological, and mutational perturbations. This is mainly because perturbations can affect the ‘internal state’ that bistable agent’s neurocontrollers are evolved to maintain after sensing the environment.

*The relative increase in internal complexity of small agents has both virtues and disadvantages for exhibiting coherent behaviors under perturbations.* A long these same lines, the autonomy that encompasses behaviors in bistable agents—those actions that do not require the continuous presence/absence of certain object in

the environment— manifests as ultimately ‘wrong decisions’ when the state of the agent is affected by perturbations. Mono- and bistable agents in the second experiment both respond to cues from the environment, but when we perturbed *the internal state of bistable agents their relative dynamical richness (i.e. having more autonomous attractors) is not necessarily an advantage in producing the ‘right behavior’*. When perturbations are big enough to produce a change in the agent-internal’s state, the bistable agent can show approaching rather than avoiding behavior in the presence of avoiding objects. This indicates behavior being affected by the effects of perturbations on dynamical trajectories in agents’ internal milieu. When perturbations were small enough, agents maintain their current states by ‘tolerating’ perturbations. The important lesson from those results is that systemic robustness can be also manifested as requiring a robust transition to an appropriate (internal) steady state in order to exhibit coherent behaviors.

*The tuning of sensory mechanisms based on agent-environment interactions is not always beneficial for behavioral robustness, which depends on the complexity of incoming signals.* The proposed third experiment (robust walking in one-legged agents section) demonstrates that the fine-tuning of sensory mechanisms through the evolution of sensory offsets has some negative and positive influence on agent’s behavioral robustness: it depends on the complexity of incoming signals that agents should process and the effects of perturbations (see [14]). When evolving with multiple sensory offsets, we have observed that agents perform a relatively more robust one-leg walking behavior with noisy signals than agents evolved with one or zero sensory offsets. For example, results demonstrate that the use of a simple (one) sensory offset created internal dynamical control in agents that necessarily depend on sensory feedback from a leg’s angle in noiseless environments. In turn, perturbations in sensory feedback generated a considerable decay in performance. Robust behavior also emerged after inducing a relatively low degree of mutational perturbations in medium or highly noisy environments.

The lesson from the third experiment is that agents showed robustness by using the experience of interacting with the environment to tune their sensory mechanisms, but the level depends on the effects of perturbations on the feedback from agent’s leg. Agents’ sensory capacities of self-tuning their body senses are thus not always beneficial for sustaining one-leg walking behavior, because the effect of sensorimotor perturbations on these dependencies. The importance of these results is that such a demonstration, because of the simplicity, lays the foundation for further work on agent environmental dependency to exhibit behaviors.

Finally, from an evolutionary perspective, *neural noise seems to bias selection toward neural systems that are resistant to the effects of bifurcations during*

*internal transient dynamics*. In the first experiment reported in this article (robust goal seeking under neural noise perturbations section), the analyzed agents had dynamic landscapes that remained functionally the same during coupled dynamics (phototaxis behavior) in the presence of sensorimotor and morphological perturbations. Robust behaviors arose in agents as a combination of locating the neurocontroller's dynamics in regions of parameter space where bifurcations produced by neural noise are unlikely to occur and where, if bifurcations occurred, they remained in functional balance (see [15] and [16]).

These findings suggest that robustness to perturbations in the face of neural noise may be a by-product of locating such regions of parameter space. In presence of certain systematic variations of parameters, the evolutionary process guarantees behavioral robustness to these variations, but it cannot ensure robustness to variations that were not given during evolution. The results also indicate that the evolution of agents under certain levels of neural noise produces robustness to this condition. Under the right parameter definition, every agent evolved in a minimal scenario and in the presence of certain level of additive neural noise has the capacity to succeed at goal approaching. Evolutionarily emerged robustness is a necessary property of agents if they are to overcome the failings of induced neural noise and perturbations during lifetime (which agrees with Jakobi's (1998) [27] seminal work in ER on robustness). This last observation confirms Jakobi's proposal in how to induce the emergence of behavioral robustness in artificial agents. In other words, results indicate that noise 'forces' evolution to this type of robustness in whatever form evolution cares to come up with. However, when agents evolved with certain level of neural noise are evaluated with higher levels of it during tests after evolution, agents show a decay in performance. If the level of neural noise was lower during evolution, agents instead exhibit at least similar fitness than after evolution.

## **IX. Discussions**

The experimental models presented in this work are far from emulating realistic examples of biological robustness. They provide theoretical evidence toward particular kinds of robust behaviors that may exist in the biological world for categorical perception, goal seeking, mobile object tracking, and walking behaviors. The key assumption is that biological organisms have evolved in coupled, controller-body-environment conditions. Simulated models of biological mechanisms with this coupling allow systematic tests that are not currently

amenable to experimental techniques in the biological realm. The methods proposed in this article (further developed in [13] and associated publications) are taken as a starting point to develop such work. Even in minimal cases, it is not trivial to evaluate the robustness of coupled systems, at least in an absolute sense. However, we can always ask: ‘robustness as compared to what?’ and ‘how does it emerge?’ in order to investigate the idea of robustness in any particular context. From an experimental viewpoint, a more interesting formulation of the study of robustness is: ‘Is there a common dynamical mechanism producing behavioral robustness in a set of related bio-inspired models?’ A universal mechanism is difficult to conceive fully formed (see [28][29][30][31][32]).

Unfortunately, heuristic answers to these questions are not enough because discussions around systems’ dynamics require no less than analysis of mathematical models. For this reason, arguments in this article are based on concrete experimental scenarios proposing comparisons that relate to each other. This work also suggests the idea that additional theoretical directions will come from studies of behavioral robustness in connected, but different domains. *Experimental evidence and theoretical discussions here demonstrate the importance of understanding behavioral robustness as a dynamical and systemic process.*

The prime motivation for the described models can be expressed as a search for the simplest interpretation of what affects systemic robustness in terms of an interface between the internal and external world of organisms. Aside from the practicality of finding such an interface in biological organisms, by understanding emergent dynamics at an organism-to-environment systemic level, this article serves as a baseline from which to understand the causally connected interplay between structure and behavior in organisms. Proposed experimental results from ‘silico’ demonstrate clearly that such a dynamical interface is possible, as distributed processes in a coupled system. Nevertheless, after this experience it seems likely that even more elegant and useful interpretations can be developed with further work. The theoretical perspective on systemic robustness provided here can effectively guide the understanding of robust phenomena in the real realm (where the holistic study is often impractical). Work here is a small contribution toward that ambitious goal, through the creation of a theoretical and algorithmic bridge between dynamics and robustness at coupled system level.

## **Acknowledgements**

The author would like to thank to Lucas Wilkins, Renan Moioli, and other members of CCNR (University of Sussex) for preliminary comments of this article.

## References

- [1] Alon, U. (2006). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Mathematical & Computational Biology. Chapman & Hall.
- [2] Arbib, M. (Ed.) (1995). *The Handbook of Brain Theory and Neural Networks*. The MIT Press.
- [3] Beer, R. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence* 72, 173–215.
- [4] Beer, R. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences* 4(3), 91-99.
- [5] Beer, R. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior* 11(4), 209–243.
- [6] Beer, R. (2004). Autopoiesis and cognition in the game of Life. *Artificial Life* 10(3), 309-326.
- [7] Beer, R. (in press). Dynamical analysis of evolved agents: a premier. In P. Vargas, E. Di Paolo, I. Harvey, & P. Husbands, *The Horizons for Evolutionary Robotics*. MIT Press.
- [8] Brooks, R. (1991d). New approaches to robotics. *Science* 253 (5025), 1227–1232.
- [9] Calabretta, R., Nolfi, S., Parisi, D., & Wagner, G. (1998). Emergence of functional modularity in robots. In *From Animals to Animats*, (ed. R. Pfeifer B. Blumberg J.-A. Meyer and S. W. Wilson), 497-504. Cambridge, Mass.
- [10] Calcott, B. (2010). Wimsatt and the robustness family: Review of Wimsatt's Re-engineering Philosophy for Limited Beings. *Journal Biology and Philosophy: Review Essay*, Springer Netherlands. DOI 10.1007/s10539-010-9202-x.
- [11] Chiel, H., & Beer, R. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences* 20, 553-557.
- [12] Félix, M., & Wagner, A. (2008). Robustness and evolution: concepts, insights and challenges from a developmental model system. *Heredity* 100, 132–140.
- [13] Fernandez-Leon, J.A. (2010). Behavioural Robustness and the Distributed Mechanisms Hypothesis. PhD Thesis. University of Sussex, Brighton, United Kingdom.
- [14] Fernandez-Leon, J.A. (2010). Evolving experience-dependent robust behaviour in embodied agents. *BioSystems*, doi:10.1016/j.biosystems.2010.09.010.

- [15] Fernandez-Leon, J.A., & Di Paolo, E. (2007). Neural Uncertainty and Sensorimotor Robustness. Almeida e Costa, F., Rocha, L.M., Costa, E., Harvey, I. and Coutinho, A., editors, *Advances in Artificial Life. Proceedings of the 9th European Conference on Artificial life ECAL 2007* (pp. 786-795). Berlin/Heidelberg: Springer.
- [16] Fernandez-Leon, J.A., & Di Paolo, E. (2008). Neural noise induces the evolution of robust behaviour by avoiding non-functional bifurcations. Asada, M., Hallam, J.C.T., Meyer, J.-A. and Tani, J., editors, *From Animals to Animats 10. 10th International Conference on the Simulation of Adaptive Behavior* (pp. 32-41). Springer.
- [17] Fernandez-Leon, J.A., & Froese, T. (2010). What is the relationship between behavioural robustness and distributed mechanisms of cognitive behaviour? *Special Session on Evolutionary Robotics. Proceedings of the 2010 IEEE World Congress on Computational Intelligence*. July, 18-23, 2010 - CCIB, Barcelona, Spain. IEEE Society, 4645-4652.
- [18] Fernandez-Leon, J.A., Acosta, G., & Mayosky, M. (2009). Behavioral control through evolutionary neurocontrollers for autonomous mobile robot navigation. *Journal of Robotics & Autonomous Systems, Vol. 57, Issue 4, Elsevier*, 411-419.
- [19] Fine, P., Di Paolo, E., & Izquierdo, E. (2007). Adapting to your body. Almeida e Costa, F., Rocha, L.M., Costa, E., Harvey, I. and Coutinho, A., editors, *Proceedings of the 9th European Conference on Artificial Life* (pp. 203-212). Springer.
- [20] Floreano, D., & Keller, L. (2010). Evolution of Adaptive Behaviour in Robots by Means of Darwinian Selection. *PLoS Biol* 8(1): e1000292. doi:10.1371/journal.pbio.1000292
- [21] Froese, T., & Di Paolo, E. (2008). Stability of coordination requires mutuality of interaction in a model of embodied agents. M. Asada, J. C. T. Hallam, J.-A. Meyer & J. Tani (eds.), *From Animals to Animats 10: Proc. of the 10th Int. Conf. on Simulation of Adaptive Behavior* (pp. 52-61). Berlin, Germany: Springer-Verlag.
- [22] Gallagher, J., Beer, R., Espenschied, K., & Quinn, R. (1996). Application of evolved locomotion controllers to a hexapod robot. *Robotics and Autonomous Systems* 19, 95-103.
- [23] Gallagher, S. (2005). *How the Body Shapes the Mind*. NY: Oxford University Press.
- [24] Harvey, I. (1992). The SAGA cross: The mechanics of recombination for species with variable length genotypes. Manner, R. and Manderick, B., editors, *Parallel problem solving from nature 2*, (pp. 269-278). Amsterdam. North-Holland.
- [25] Harvey, I., Di Paolo, E., Wood, R., Quinn, M., & Tuci, E. (2005). Evolutionary robotics: A new scientific tool for studying cognition. *Artificial Life* 11(1-2), 79-98.

- [26] Hubert, J., Matsuda, E., Silverman, E., & Ikegami, T. (2009). A Robotic Approach to Understanding Robustness. *The 3rd International Symposium on Mobiligence*, (pp. 361-366).
- [27] Jakobi, N. (1998). Evolutionary robotics and the radical envelope of noise hypothesis. *Adaptive Behavior* 6(2), 325–368.
- [28] Kitano, H. (2002). Systems Biology: a brief overview. *Science* 295, 1662-1664.
- [29] Kitano, H. (2004). Biological Robustness. *Nature Reviews: Genetics* 5, Nature Publishing Group, 826-837.
- [30] Kitano, H. (2004). Cancer as a robust system: implications for anticancer therapy. *Nat. Rev. Cancer* 4, 227–235.
- [31] Kitano, H. (2006). Robustness from top to bottom. Book review of Wagner (2005): Robustness and Evolvability in Living Systems. *Nature Genetics* 38, 133.
- [32] Kitano, H. (2007). Towards a theory of biological robustness. *Molecular Systems Biology* 3:137. EMBO and Nature Publishing Group.
- [33] Krakauer, D. (2005). Robustness in biological systems: a provisional taxonomy. *T.S. Deisboeck and Y. Kresh, Editors, Complex Systems Science in Biomedicine, Plenum*, (pp. 185–207).
- [34] Macinnes, I., & Di Paolo, E. (2006). The advantages of evolving perceptual cues. *Adaptive Behavior* 14(2), 147–156.
- [35] Nolfi, S., & Floreano, D. (2000). *Evolutionary Robotics. The Biology, Technology, and Intelligence of Self-Organizing Machines*. MIT Press.
- [36] Pfeifer, R., & Bongard, J. (2006). *How the Body Shapes the Way We Think: A New View of Intelligence*. Bradford Books, MIT Press.
- [37] Rabinovich, M., Varona, P., Selverston, A., & Abarbanel, H. (2006). Dynamical principles in neuroscience. *Reviews of Modern Physics* 78, 1213-1265.
- [38] Silverman, E., & Ikegami, T. (2010). Robustness in Artificial Life, *Int. J. Bio-Inspired Computation, Vol. 2, No. 3*, 197–212.
- [39] Slocum, A., Downey, D., & Beer, R. (2000). Further Experiments in the Evolution of Minimally Cognitive Behavior. From Perceiving Affordances to Selective Attention. Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior (pp. 430-439). MIT Press.
- [40] Stelling, J., Sauer, U., Szallasi, Z., Doyle III, F., & Doyle, J. (2004). Robustness of cellular functions. *Cell* 118, 675–685.

[41] Strogatz, S. (1994). *Nonlinear Dynamics & Chaos* (2000, First paperback printing). Reading, MA: Addison-Wesley.

[42] Teo, J. (2004). Robustness of Artificially Evolved Robots: What's Beyond the Evolutionary Window? *The 2nd International Conference on Artificial Intelligence in Engineering and Technology (ICAIET 2004)*, volume 1, (pp. 14-20).

[43] von Neumann, J. (1956). Probabilistic logics and the synthesis of reliable organisms from unreliable components. In E. C. Shannon and J. McCarthy, *Automata Studies*. Princeton: Princeton University Press.

[44] Ziemke, T. (2001). Disentangling notions of embodiment. Pfeifer, R., Lungarella, M., & Westermann, G. (Eds.), *Proceedings of Workshop on Developmental and Embodied Cognition*, Edinburgh, UK, (pp. 4-8).

[45] Ziemke, T. (2003). What's that thing called embodiment? Alterman, R., Kirsh, D. (Eds), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 1134-1139). NJ: Lawrence Erlbaum, Mahwah.

[46] Ziemke, T., Bergfeldt, N., Buason, G., Susi, T., & Svensson, H. (2004). Evolving cognitive scaffolding and environment adaptation: a new research direction for evolutionary robotics. *Connection Science* 16 (4), 339-350.

