



OPEN ACCESS

EDITED BY

Inés Hipólito,
Humboldt University of Berlin,
Germany

REVIEWED BY

Alban Lafflaquière,
Softbank Robotics, France
David Vernon,
Carnegie Mellon University Africa,
Rwanda

*CORRESPONDENCE

Matthew D. Egbert
✉ m.egbert@auckland.ac.nz

SPECIALTY SECTION

This article was submitted to
Human-Media Interaction,
a section of the journal
Frontiers in Neurorobotics

RECEIVED 01 January 2022

ACCEPTED 29 September 2022

PUBLISHED 21 December 2022

CITATION

Egbert MD and Barandiaran XE (2022)
Using enactive robotics to think
outside of the problem-solving box:
How sensorimotor contingencies
constrain the forms of emergent
autonomous habits.
Front. Neurobot. 16:847054.
doi: 10.3389/fnbot.2022.847054

COPYRIGHT

© 2022 Egbert and Barandiaran. This is
an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction
in other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Using enactive robotics to think outside of the problem-solving box: How sensorimotor contingencies constrain the forms of emergent autonomous habits

Matthew D. Egbert^{1,2*} and Xabier E. Barandiaran³

¹School of Computer Science, University of Auckland, Auckland, New Zealand, ²Te Ao Mārama, University of Auckland, Auckland, New Zealand, ³IAS-Research Center for Life, Mind, and Society, Department of Philosophy, University of the Basque Country (UPV/EHU), Bilbao, Spain

We suggest that the influence of biology in ‘biologically inspired robotics’ can be embraced at a deeper level than is typical, if we adopt an enactive approach that moves the focus of interest from how problems are solved to how problems emerge in the first place. In addition to being inspired by mechanisms found in natural systems or by evolutionary design principles directed at solving problems posited by the environment, we can take inspiration from the precarious, self-maintaining organization of living systems to investigate forms of cognition that are also precarious and self-maintaining and that thus also, like life, have their own problems that must be addressed if they are to persist. In this vein, we use a simulation to explore precarious, self-reinforcing sensorimotor habits as a building block for a robot’s behavior. Our simulations of simple robots controlled by an Iterative Deformable Sensorimotor Medium demonstrate the spontaneous emergence of different habits, their re-enactment and the organization of an ecology of habits within each agent. The form of the emergent habits is constrained by the sensory modality of the robot such that habits formed under one modality (vision) are more similar to each other than they are to habits formed under another (audition). We discuss these results in the wider context of: (a) enactive approaches to life and mind, (b) sensorimotor contingency theory, (c) adaptationist vs. structuralist explanations in biology, and (d) the limits of functionalist problem-solving approaches to (artificial) intelligence.

KEYWORDS

enactive robotics, sensorimotor contingencies, problem-solving, habits, IDSM, sensorimotor autonomy

1. Introduction

Artificial Intelligence and the scientific approach to mind (what is known as *cognitive science*) was born (or rather raised) as a problem solving discipline (Newell et al., 1958; Putnam, 1965; Fodor, 1968). Deprived of life, the machine metaphor was one of symbol manipulation and rationality (deductive, inferential, heuristic or otherwise). The unprecedented potential of Universal Turing Machines (computers) was the driving metaphor to study the mind. The software, the mind, was the problem solving method, the hardware, the brain, its implementation.

Alternative conceptions of the mind were available at the origins of Artificial Intelligence and Robotics (Grey Walter, 1950; Ashby, 1952) but the rapid success of computer science left them aside. Over time, the limitations of the problem-solving centered computational theory of the mind became apparent and the biologically inspired, embodied and later enactive conceptions of the mind gained momentum. We are now immersed in a mesh of hybrid architectures, applied to a wide range of practices, from industrial to scientific modeling applications, and a new summer of Artificial Intelligence is rising, with robotics as a major container of social and technological expectations.

There are good reasons for why problem-solving attracts so much attention from researchers, but it is pertinent to ask: what aspects of minds are omitted or obscured by the problem-solving focused perspective? what can life teach us about what intelligence holds before and beyond problem-solving? and even when problem solving is addressed ...how is it that natural agents have and become concerned by *their own* problems?

This paper has two goals. The first is to argue that by abandoning problem solving (or at least putting it down for a time), other useful explanatory targets and ways to explain minds are given space to emerge.

The second, more specific goal is a case in point: we use simulated robots to show how sensorimotor contingencies influence the formation of self-maintaining patterns of sensorimotor activity “habits” in regular ways that depend upon sensory modality. By de-emphasizing problem solving, we are able to take a fresh look at the relationship between sensory modalities, sensorimotor contingencies and habitual behaviors. But to explain these results, we first need to provide more context.

The paper proceeds as follows. The next section explains what we mean by “problem-solving,” why it has been a popular target within the cognitive sciences, and what we see as the primary disadvantage of excessive attention being given to the topic. We then explore the intimate relationship between robotics and biologically inspired and embodied problem-solving paradigms. Section 2.3 introduces the enactivist concept of autonomy, providing an alternative framework for developing Sensorimotor Contingency Theory outside of the

problem-solving approach. The remainder of the paper presents and analyzes a simulation model that is used to explain: (i) that robots must first have their own problems instead of solving those posited by external observers; (ii) that, in doing so, they must assert a way of life whose structure and form must be taken as the object of study. We finally discuss some of the larger implications of our enactive approach in connection with wider theories of biological explanation and inspiration.

2. From problem solving to enactive robot

2.1. Problem-solving in minds and machines

We use the term “problem-solving” to refer, in a broad and inclusive manner, to the kinds of things that we associate with being capable or clever. Nowadays, for many, “the ability to solve problems is not just an aspect or feature of intelligence—it is the essence of intelligence” (Hambrick et al., 2020, p.553). It is certainly not a new idea. The very birth of Artificial Intelligence owes much to it (Newell et al., 1958). As Newell and Simon later stated: “Since ability to solve problems is generally taken as a prime indicator that a system has intelligence, it is natural that much of the history of artificial intelligence is taken up with attempts to build and understand problem-solving systems” (Newell and Simon, 1976, p. 120). The task of artificial intelligence was thus to devise potential solution-generators and to design tests that could evaluate them. This assumed that the problem space was well fixed so that solutions could both be evaluated and generated. Decades later, Artificial Intelligence handbooks still devote their first central sections to problem solving (e.g., see part II of Norvig and Russell, 2021, 4th edition).

A problem is a context in which behaviors can be evaluated according to a norm of success at “solving” the problem. Problems vary from being trivial to challenging to impossible as the proportion of behaviors that are good (out of all possible behaviors) shrinks. They include “high-level” human problem-solving, such as the skills that are taught in schools or universities, as well as embodied problems such as balancing on two feet or swimming efficiently. For a system to be evaluated in terms of problem-solving, one must first have the specification of the context and of a normative evaluation so that behaviors within that context can be compared as more or less successful. Defined as such, just about anything can be seen as a problem solver. A bottle lid solves the problem of preventing spills; a car’s differential solves the problem of distributing force effectively to its wheels; a computer program solves the problem of beating a human at chess.

Herein lies both the advantage and the disadvantage of placing problem solving at the center of the cognitive sciences: *almost anything can be evaluated in terms of its problem solving*

ability. On the one hand this is a great boon. Quantifying problem solving ability is relatively straight-forward and the ability to quantify how effective a system is at solving one or more problems facilitates technoscientific progress where measurements can play an important role in defining progress. This is apparent in artificial intelligence (AI) research where benchmarks such as chess and other games (Canaan et al., 2019), hand-writing recognition (e.g., Graves and Schmidhuber, 2009), image classification, speech recognition, etc. (e.g., MLPerf benchmarks MLPerf, 2021), are used to quantify progress. Problem-solving similarly provides metrics for studying the minds in psychology and neuroscience, where problem-solving related notion of ‘tasks’ (e.g., the Simon task) are used to structure human activity and performance related metrics such as reaction speed or error rate are seen as providing key insights into how our minds operate (Simon and Wolf, 1963).

The ease of measuring problem-solving ability sometimes leads to it (problem-solving ability) becoming the *explanandum*—the thing we strive to understand. This is seen in research questions like: *How do people recognize faces so well? How do babies come to understand the motives of other people? How do we play chess? How can we make a safe self-driving car?* etc. Problem-solving also sometimes becomes the *explananda*—the terms in which we explain what minds are, how they work, or what they do. Evolutionary psychology (Pinker, 1997; Buss, 1998), for example, emphasizes the evolutionary advantage of problem-solving ability and in this context, explanations that are provided in terms of problem-solving ability are seen as complete, as evolution can be invoked to explain why such mechanisms exist. The evolutionary advantage of having a mind is in its contribution to problem solving and therefore, minds are best understood as problem solving machines.

However, it has been argued that: “The essence of intelligence is to act appropriately when there is no simple pre-definition of the problem or the space of states in which to search for a solution. Rational search within a problem space is not possible until the space itself has been created, and is useful only to the extent that the formal structure corresponds effectively to the situation” (Winograd and Flores, 1987, p.98). Thus, even from a problem solving perspective, intelligence is not really the capacity to solve a problem but to bring a situation into a fabricated frame where it can be treated as a problem to be solved.

Moreover, the problem with excessive focus upon problem-solving is that there are other unique and important features of minds that are worthy of study—features that may only indirectly relate to problem-solving ability or perhaps not at all. The problem, in a nutshell, is the conflation of (i) “problem-solving ability” with (ii) all of the other phenomena associated with “being a mindful body.”

The mainstream computational functionalist approach to the mind (Putnam, 1965; Fodor, 1968) doesn’t really help much addressing what mindful bodies are beyond problem-solving devices. For Putnam, the very definition of the mental is always in reference to a Turing machine table that works out rational transitions (e.g., computing and storing preferences over a utility function or solving problems in problem representation space). Deviations from this rationality are treated as pathological. All human mental life is, according to Putnam, not perfectly normal, thus relatively pathological. Putnam acknowledges “our model is an overly simple and overly rationalistic one in a number of respects. However, it would be easy, in principle, although perhaps impossible in practice, to complicate our model in all these respects—to make the model dynamical, to allow for irrationalities in preference, to allow for irrationalities in the inductive logic of the machine, to allow for deviations from the rule: maximize the estimated utility. But I do not believe that any of these complications would affect the philosophical conclusions reached in this paper” (Putnam, 1965, p. 43).

Deviations from the abstract rational rule are pathological. Explanation lies on the pure domain of abstract problem solving, the deviations make it all more complicated (as if dirt in the form of a set of exceptions where to be added to the pure explanation) but change fundamentally nothing. As we are about to see, embodied approaches, and particularly enactivism, bring these “pathological” expressions to the center of the explanation (of which rational thinking is the complicated achievement) turning it into the core constitution of mind. Biologically inspired robotics has a lot to contribute in this direction.

2.2. From biologically inspired problem-solving to enactive robotics

Enactivism was born under the conviction that robotics, as a field, would require, or even force, cognitive science to move beyond the problem-solving framework:

The assumption in CS [Cognitive Science] has all along been that the world can be divided into regions of discrete elements and tasks to which the cognitive system addresses itself, acting within a given “domain” of problems: vision, language, movement. Although it is relatively easy to define all possible states in the “domain” of the game of chess, it has proven less productive to carry this approach over into, say, the “domain” of mobile robots. Of course, here too one can single out discrete items (such as steel frames, wheels and windows in a car assembly). But it is also clear that while the chess world ends neatly at some point, the world of movement amongst objects does not. It requires our continuous use of common sense to configure our world of objects. (Varela, 1992, p.251)

Yet, moving away from the problem-solving paradigm has taken a long path, most of which, rather than abandoning problem-solving has deeply transformed the way we understand how nature solves those problems. In a sense, biologically inspired robotics has mostly followed the problem-solving approach and biological inspiration has focused on picking up biological mechanisms to solve problems: from the internal neuronal inspiration of artificial neural networks since its early conception (Rosenblatt, 1958) to their later development (Rumelhart et al., 1988) to the embodied strategies that either transform the problems to be solved by their “brains” or have outsourced the computational load of the problem solving to body and world (Pfeifer and Scheier, 2001). What radically distinguished biologically inspired robotics from GOFAI (Good Old Fashioned Artificial Intelligence) was a change of focus from the abstract to the concrete, from the symbolic to the sensorimotor and from the rational to the practical know-how of situated action. Despite the emphasis on self-organization, agent-environment emergence of behavioral functioning (Steels, 1990) embodiment and situated action (Maes, 1990), etc. the main goal was still to build robots capable of solving behavioral problems. After all, to put it with biologically-inspired roboticist Barbara Webb: “The sensorimotor problems faced by animals and by robots have much in common” (Webb, 1995, p. 117) and, not only can animals help us devise robots that solve problems in a biologically inspired manner, but also solving a sensorimotor problem with the robot could help us understand how the animal solves it; like “[d]etecting which ear is closer to the sound” which “is a non-trivial problem for the cricket” (Webb, 1995, p. 120).

Other trends of biological inspiration have built upon evolutionary theory itself and artificially evolved brains or brains and bodies to solve the problems (encoded as fitness function) in what is commonly known as evolutionary robotics (Cliff et al., 1993; Nolfi et al., 2016). Random variations to the parameters of robotic brains and bodies are selected against a fitness function that operates as the benchmark of the problem to be solved. Despite the problem-solving focus that is almost inherent in artificial evolutionary optimization techniques, evolutionary robotics served to disclose a number of principles of behavioral self-organization that non-linear, fine grained agent-environment coupling display when artificial evolution can freely explore the solution space without the prejudices inherent to the human design (Harvey et al., 1997).

Some of these approaches entail radical departures from core assumptions of the computational functionalist theory of the mind: cognitive processing does not only occur in the head and the body must be integrated as a key feature of cognitive problem solving (not simply as an executioner of the solution representation worked out in the head or a sensory transmitter of the problem into it); agent-environment interactions can self-organize with little if any representations; material bodily and

interactive constitution are not mere implementation details of abstract capacities but intrinsic part of the problems and solutions that cut across them.

But *enactive robotics* moves yet further into biological inspiration. On the one hand there are enactive approaches that have attempted to introduce more of the living body of natural intelligence into robotics by including self-organized mechanical, soft bodies (Man and Damasio, 2019) or even chemical bodies (Damiano and Stano, 2021). But also, and perhaps most relevant for enactive theory transferring to the robot what metabolism has to offer to anchor intrinsic needs or emotional feedback (Ziemke and Lowe, 2009). There is however another enactive path that brings into robotics some principles of living organization and, more specifically, the autonomy of behavior, a *way of life* for robots, not as somehow transferred from the biological body, but enacted at the scale of brain-body-world dynamics (Barandiaran and Moreno, 2006). Some forerunners of this inspiration are no doubt Ross Ashby on the organism-centered inspiration of adaptive controllers as machines capable to remain homeostatic in the face of perturbations (Ashby, 1952) and Grey Walter’s “life imitating” robots “designed to illustrate the uncertainty, randomness, free will or independence so strikingly absent in most well-designed machines” (Grey Walter, 1950, p. 44).

More recent development of this line of inspiration on natural and biological principles for the design of robots came hand in hand with the development of a theory of autonomous behavior and agency (Smithers, 1997), organismically inspired robotics (Di Paolo, 2003), and habit-centered enactive robotics (Egbert and Barandiaran, 2014).

2.3. Enactivism and the autonomy of sensorimotor life

Varela, Thompson and Rosch opened up a new way of thinking in 1991. Their enactive approach conceives that “cognitive structures emerge from the recurrent sensorimotor patterns that enable action to be perceptually guided” (Varela et al., 1991, p.173). They later stated that: “[C]ognition is no longer seen as problem solving on the basis of representations; instead, cognition in its most encompassing sense consists in the enactment or a bringing forth of a world by a viable history of structural coupling” (Varela et al., 1991, p.205). The enactive approach thus emphasizes sensorimotor coupling and the recurrent patterns that emerge from agent-environment interactions.

This way Varela overcame the operational (en)closure of the nervous system that served as his main analogy with the organization of the living, captured (together with Humberto Maturana) within the theory of autopoiesis. Ever since, the relationship between the self-organizing nature of nervous

activity and that of behavior became to some extent problematic (see [Barandiaran, 2017](#) for a discussion).

Inspired by Maturana and Varela, Tim Smithers re-states the need for biologically inspired autonomy in robotics, in the context of the impossibility to design robot agents from a problem-solving stand point:

Designing and building autonomous agents thus becomes the problem of designing and building processes that can support and maintain this kind of identity formation through interaction: processes that, through interaction, are continuously forming the laws of interaction that can sustain and maintain the interaction needed to form them. In other words, we need interaction processes that can support the self-construction and maintenance of interaction processes through interaction, in essentially the same way that the material and energy interaction processes of single cells can be understood as being involved in the continual forming of the mechanisms that support this interaction. Such systems will thus be self-law making as well as self-regulating, in essentially the same way as we can understand biological systems and autonomous city states. ([Smithers, 1997](#), p.102)

This analogy between metabolic autonomy and the autonomy of behavior was further explored in [Di Paolo \(2003\)](#). According to this view, enactivism needs not be committed to build bio-chemically living robots (provided that this is possible or even desirable) but to endow a robot with a *way of life*. This intuition was further explored in [Barandiaran \(2007, 2008\)](#).

Sensorimotor Contingency Theory can further enrich this approach. It uses regularities in the ways that motor activity affect sensory activity (sensorimotor contingencies) to explain the qualities of perceptual experience ([O'Regan and Noë, 2001](#)). Empirical research involving sensory substitution, sensory modification, psychophysics has informed the development of sensorimotor contingency theory (SMCT), which attempts to explain diverse aspects of perceptual experience, including why certain sensory modalities have a particular “feel” to them; how it is possible to make one experience one sensory modality (e.g., touch) in a way that feels more like another (e.g., sight); and the conditions in which subjects are (or are not) capable of adapting to major transformations to their sensorium. The key idea in SMCT is the role that action plays in perception: a classical enactive theme ([Noë, 2004](#)).

Using variations of the basic theme of how motor activity modifies sensory input, a set of robotic architectures were made using sensorimotor contingencies as building blocks for robotic design ([Maye and Engel, 2013](#); [Jacquey et al., 2019](#)). But these are hardly enactive in the sense of the deep biological inspiration that the enactive approach can offer.

Perhaps the best way to explore such potential is to bring forth the concept of sensorimotor autonomy (updated and

refined from a previous proposal of Mental Life and also explored in more detail on the concept of Sensorimotor Life): the capacity of an agent to sustain and regulate the structures that generate behavior. This definition echoes the metabolism-based definitions of life as far-from-thermodynamic equilibrium chemical systems capable of maintaining the network of chemical reactions that constitute it ([Gánti, 1975](#); [Maturana and Varela, 1980](#); [Rosen, 1991](#); [Ruiz-Mirazo et al., 2004](#); [Luisi, 2006](#)).

The basic constituent of sensorimotor autonomy is a sensorimotor structure (a behavioral scheme or habit) made possible by both a set embodied-neural pathways and a set of sensorimotor contingency relationships. Think of it as a coordination pattern that emerges out of environmental, sensorimotor and neural (or behavior generating) mechanisms. Now, if this structure is far-from-equilibrium or, said differently, if let alone it tends to extinguish or vanish, and if the very enactment of the sensorimotor scheme reinforces itself by repetition or by satisfying certain conditions that feed-back into its supporting structure (e.g., reinforcement of synaptic connections by Hebbian learning or reward reinforcement), then we have first sense of self-maintenance that is characteristic of habits. The more the habit is enacted the more it is strengthened, the stronger it is the more likely it is to be enacted. That is the virtuous (or vicious!) self-sustaining nature of the habit.

We can now go back to the original, albeit obscure and self-referential, intuition of roboticist Tim Smithers and his idea of the autonomy of behavior based on the “processes that, through interaction, are continuously forming the laws of interaction that can sustain and maintain the interaction needed to form them.” These laws or norms are nothing other than the very conditions under which the habit is viable, that it can persist and sustain itself. This is a strong analogy with (metabolic) life that opens up the very possibility of *having a problem of your own* and *having to solve it*. The problem for the precarious, self-maintaining autonomous cell is persistence, avoiding decay and disintegration (see [Barandiaran and Egbert, 2013](#) for a more detailed analysis). The same goes for the precarious, self-maintaining nature of a habit, and ultimately of autonomous sensorimotor life (see [Barandiaran, 2008](#); [Di Paolo et al., 2017](#)). This approach also opens up a new mode of explanation that is characteristic of biological thinking and can be applied to sensorimotor dynamics: focusing on the nature and structure of constraints rather than the problems they are suppose to be adapted to solve. We return to these themes in the discussion section.

In what follows we introduce habit-based enactive robotics to illustrate and further develop the points we have briefly outlined. Inspired by SMC and enactive principles we build robots that are capable of generating spontaneously a complete ecology of habits that display structural constraints within the sensorimotor space. The results will help us discuss how enactive robotics can contribute to a new understanding of mind

and cognition with a deeper biological inspiration than what problem-solving can provide.

3. Model

3.1. Overview

The computational model simulates a two-wheeled robot that moves around a two-dimensional environment. The robot has two independently controlled motors; one for each of its wheels, allowing it to move forwards or backwards in a straight line or to turn in a variety of arcs, or on the spot. The robot's motors are controlled by an iterant deformable sensorimotor medium (IDSM) (Egbert and Barandiaran, 2014)—a habit-forming controller that is described in detail below.

The robot's environment is periodic in the sense that when the robot moves off one side of the environment, it appears on the opposite side. A stimulus source is also located in the environment. This source, which moves around the environment, can be thought of as simultaneously emitting a sound tone and a source of light, but in any given trial the robot is only capable of perceiving one of those sensory modalities (light or sound). We now present each element of the simulation in detail.

The robot and its environment are simulated using Euler forward integration with a time step of $\Delta t = 0.01$. Thus, a single time-unit consists of 100 iterations, and the IDSM is updated every iteration.

3.2. Stimulus

The stimulus source moves around the environment in a circle. Each rotation, it slows to a stop at its left-most position, before accelerating again to complete another rotation. This trajectory is specified by the following equations which describe the stimulus's position (s_x, s_y) as a function of time (t):

$$s_x = \frac{1}{2} + \cos\left(\frac{t}{10} + \sin\left(\frac{t}{10}\right)\right) \quad (1)$$

$$s_y = \frac{1}{2} + \sin\left(\frac{t}{10} + \sin\left(\frac{t}{10}\right)\right). \quad (2)$$

3.3. Robot

The simulated robot (Figure 1) has two independently controlled motorized wheels. It's position changes according to the following differential equations,

$$\frac{dx}{dt} = k_s \cos(\alpha)(m_l + m_r) \quad (3)$$

$$\frac{dy}{dt} = k_s \sin(\alpha)(m_l + m_r) \quad (4)$$

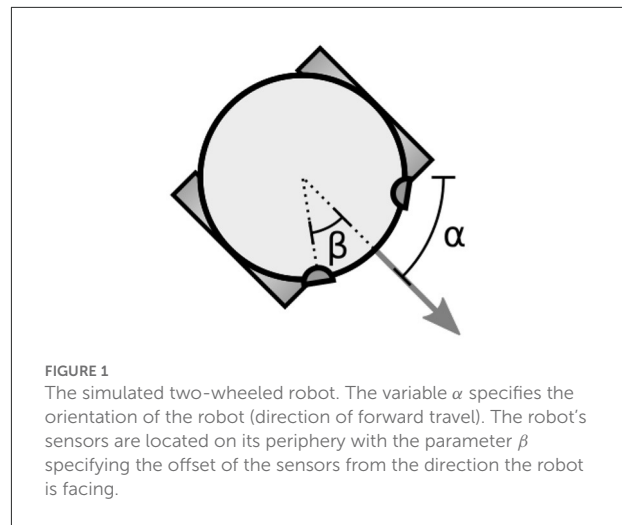


FIGURE 1
The simulated two-wheeled robot. The variable α specifies the orientation of the robot (direction of forward travel). The robot's sensors are located on its periphery with the parameter β specifying the offset of the sensors from the direction the robot is facing.

$$\frac{d\alpha}{dt} = k_s \frac{(m_r - m_l)}{2R}, \quad (5)$$

where x and y are the robot's position in the environment; α is its heading; $k_s = 0.25$ scales the speed of the motors; $R = 0.05$ is the robot's radius; and the variables m_l and m_r represent the velocities of the robot's left and right wheel motors.

We consider two robot sensor configurations. "Visual" robots have two directional sensors. The excitation of each, V , is the product of an attenuation factor due to distance from that sensor to the stimulus, and an attenuation factor due to misalignment between the orientation of the sensor and the relative direction of the stimulus. This second attenuation factor is calculated by taking the scalar product of a unit vector that points from the sensor to the stimulus and \vec{o} , a unit vector that specifies the direction that the sensor is facing.

Formally,

$$V = \frac{0.25}{0.25 + \|\vec{r}\|^2} \underbrace{\left(\frac{\vec{r}}{\|\vec{r}\|} \cdot \vec{o} \right)^+}_{\text{orientation}}, \quad (6)$$

where \vec{r} is a vector that describes the position of the stimulus relative to the sensor; $\|\vec{r}\|$ is the magnitude of that vector; and the $+$ superscript indicates that negative values within the parentheses are truncated to zero. The excitation of these sensors is thus highest when the sensor is close to the stimulus and directly facing it.

"Auditory" robots have two sensors that respond to the rate at which the sensor is approaching or moving away from the stimulus source. This is analogous to the Doppler effect whereby the perceived frequency of a sound when approaching is higher than when moving away from the listener. The excitation of an auditory sensor, A , is given by:

$$A = \frac{1}{2} + k \frac{d\|\vec{r}\|}{dt} \quad (7)$$

where the first term can be thought of as the tones natural pitch which is offset by the relative speed of the sensor and the stimulus scaled by $k = \frac{3}{4}$ to keep the magnitude of the sensor (given the relative speeds of the robot and the stimulus) within a similar range of excitation values as simulated for the visual sensor.

To address the periodic boundaries of the environment, auditory sensors always use the nearest stimulus source as defined by the minimum image convention and visual sensors calculate the combined effect of 5 stimuli: one in the simulated space and four virtual copies of the sensor offset north, south, west and east of the simulated space arena by one arena width. This means that if, for example, a visual robot is close to the north boundary of the arena and facing north it can still see stimuli.

Visual sensors are offset from the direction the robot is facing by $\beta_v = \pi/5$ (see Figure 1). Auditory sensors are offset by $\beta_a = \pi/2$. The orientation of visual sensors is $\alpha \pm \beta_v$, i.e., perpendicular to the tangent of the robot's circular body at that position, facing outwards. Auditory sensors have no orientation.

3.4. IDSM

3.4.1. Overview

The IDSM is a robot controller intended to capture the idea of a self-maintaining pattern of sensorimotor activity. Inspired by the habitual behavior of people and by the enactivist concept of “autonomous” self-sustaining sensorimotor systems (see e.g., Di Paolo et al., 2017), the IDSM was designed such that patterns of sensorimotor activity reinforce the mechanisms that produce them (Egbert and Barandiaran, 2014; Egbert and Cañamero, 2014; Egbert, 2018). The IDSM has been used to explore how a habit-based individual can be trained to perform different tasks (Egbert and Barandiaran, 2014); how different forms of motor babbling can bias the subsequent formation of habits (Zarco and Egbert, 2019); how the essential variables of a biological autonomous system can be shared with the essential variables of a sensorimotor autonomous system (Egbert and Cañamero, 2014) and the extent to which IDSM-based sensorimotor autonomous systems can be considered to be adaptive (Egbert, 2018). In the present paper, we use the IDSM in a new way: to show, in a formal model, how sensorimotor-contingencies can play an essential role of sculpting the form of habits without themselves being explicitly internalized or represented by the “brain” or “controller” of an embodied agent.

The IDSM works by recording trajectories taken through sensorimotor space, i.e., how the sensorimotor state changes for various experienced sensorimotor states. When a sensorimotor state is experienced that is similar to one that has been experienced in the past, the motors of the robot are actuated in a way similar to how they were actuated in that previous experience. Memories of previous trajectories are gradually

forgotten, unless they are reinforced by re-enactment and so the only patterns of behavior that can persist for long periods of time are those that are re-enacted. When the sensorimotor state is in an unfamiliar (or forgotten) state, motor activity is random. The IDSM used in this paper is very similar to that described in Egbert (2018). Any differences between the model here and that in Egbert (2018) are explicitly highlighted below.

A useful metaphor for understanding how the IDSM works is the paths that form on university campuses, where paths taken by students crossing a grassy field between academic buildings trample and kill the grass. The emergent dirt paths influence the trajectories taken by subsequent students, but the grass also regrows, so only emergent paths that are regularly traveled can persist in the long-term. This is essentially how the IDSM operates, but the trajectories taken and the paths that form are in sensorimotor space, rather than on a university campus. The dynamic also relates to the self-reinforcing nature of habitual behavior, where repeated performance of patterns of behavior (e.g., the direction you look when crossing the street; smoking a cigarette; or a tendency to worry) increases the likelihood of similar behavior being performed in the future. And to reiterate: the IDSM was designed to capture the enactivist concept of autonomy (a precarious self-maintaining system) in a sensorimotor system—see Di Paolo et al. (2017).

More formally, the IDSM can be thought of as a function, f , that transforms the robot's current sensorimotor state into an “output,” i.e., the next moment's motor state: $f_t(S_t, M_t) \rightarrow M_{t+1}$. As this function is applied, the function itself also changes as a function of the current state of sensors and motors and the current state of the function: $\frac{\Delta f}{\Delta t} = g(f, S, M, \frac{\Delta S}{\Delta t}, \frac{\Delta M}{\Delta t})$. This change, which we shall now describe, was engineered so that sensorimotor state trajectories would bias the system to increase the likelihood that similar sensorimotor trajectories will be repeated in the future.

3.4.2. Tracking sensorimotor trajectories

As the robot's sensorimotor state changes, the IDSM maintains a set of records called “nodes.” Each node describes the sensorimotor-velocity (i.e., the rate of change in all sensors and motors) for a particular sensorimotor-state at the moment that the node was created. Each node, N , is a tuple, $N = \langle \mathbf{p}, \mathbf{v}, w \rangle$, where \mathbf{p} represents the sensorimotor state associated with the node (referred to as the node's ‘position in sensorimotor space’); \mathbf{v} indicates the sensorimotor velocity when the node was created; and w indicates the weight of the node, a value that changes dynamically and is used to scale the overall influence of the node. We shall refer to these components using a subscript notation, where the position, SM-velocity vector, and weight of node N are written as N_p and N_v and N_w , respectively.

As a robot controlled by the IDSM moves through sensorimotor states, new nodes are created recording the sensorimotor velocities experienced at different sensorimotor states. Specifically, when a new node is created, its N_p is set

to the current sensorimotor state; its N_v is set to the current rate of change in each sensorimotor dimension, and its initial weight, N_w^0 .

The two vector terms (N_p and N_v) can be thought of as existing within an abstract sensorimotor space (ASMS). A many to one mapping transforms any given position in the ASMS to a specific sensorimotor state, according to:

$$\begin{bmatrix} \Lambda(\mu_l) \\ \Lambda(\mu_r) \\ \sigma_l \\ \sigma_r \end{bmatrix} = \begin{bmatrix} m_l \\ m_r \\ s_l \\ s_r \end{bmatrix} \quad (8)$$

where μ_l, μ_r, σ_l , and σ_r indicate a position in the four dimensional ASMS; the vector on the right indicates the sensorimotor state (i.e., the actual state of left-motor, right-motor, left-sensor, and right-sensor) associated with that ASMS position; and $\Lambda(x) = \sin[(4+x)2\pi x^3]$ is the non-linear function plotted in Figure 2. The purpose of this mapping is to avoid prescribing a characteristic rate of motor change, and instead to allow the IDSM to autonomously find habits with rates of motor change that are neither too fast nor too slow. Different regions of ASMS correspond to different rates of motor change. For instance, when abstract motor state $x \lesssim 0.5$, a small change in that state variable corresponds to a small change in $\Lambda(x)$, the actual motor state (see the difference between the red circle and red X in Figure 2), where an equivalent change in x when $x \gtrsim 0.7$ corresponds to a greater change in the actual motor state (blue circle and X in Figure 2). When the sensorimotor state of the robot is unfamiliar, changes in motor activity are driven randomly (as explained below) and this Λ mapping allows these random changes in sensorimotor space to correspond to slow or fast changes in motor state. By exploring different parts of the abstract sensorimotor state, the IDSM can experiment with sensorimotor patterns with different rates of change until ones that are self-reinforcing emerge. The ASMS is also treated as periodic so as to avoid the IDSM getting stuck at motor boundaries (as discussed in Egbert, 2018). ASMS state variables, $\mu_l, \mu_r, \sigma_l, \sigma_r$ all $\in [0, 1]$ and the ASMS distance functions (described below) adhere to the minimum-image convention.

New nodes are added when the density of nodes near the current sensorimotor state is less than a threshold value, i.e., when $\psi(x) < k_t$. Loosely speaking, ψ is a measure of how ‘familiar’ the current sensorimotor state is, as estimated by summing a non-linear function of the distance (Equation 11) from every node with a positive weight to the current sensorimotor state. Formally,

$$\psi(x) = \sum_N \omega(N_w)d(N_p, x) \quad (9)$$

$$\omega(N_w) = \begin{cases} 1 & \text{if } N_w > 0 \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

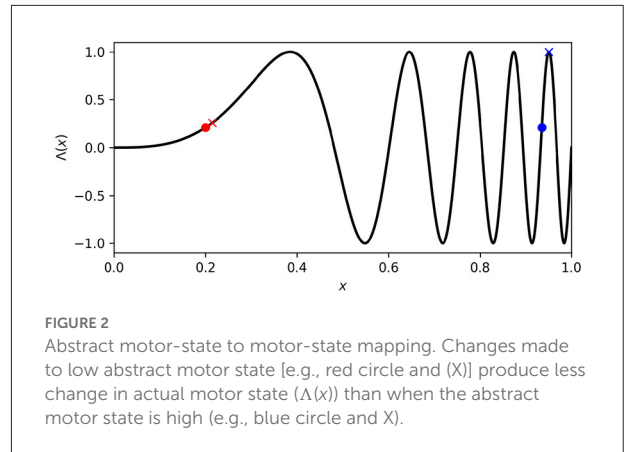
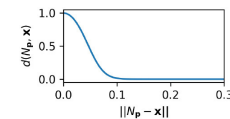


FIGURE 2 Abstract motor-state to motor-state mapping. Changes made to low abstract motor state [e.g., red circle and (X)] produce less change in actual motor state $\Lambda(x)$ than when the abstract motor state is high (e.g., blue circle and X).

where x represents the current ASMS position, k_t is a threshold parameter describing maximum node-density at which new nodes will be created and $d()$ is the following non-linear ASMS-distance function.

$$d(N_p, x) = \frac{2}{1 + \exp(k_d|N_p - x|^2)} \quad (11)$$



3.4.3. Nodes influence the sensorimotor state

One time unit after a node has been created, it is added to the set of ‘activated’ nodes that influence the sensorimotor state according to:

$$m = \frac{1}{\phi(x)} \sum_N \omega(N_w)d(N_p, x) \underbrace{\left(\underbrace{N_v}_{\text{vel.}} + \underbrace{A(N_p - x, N_v)}_{\text{attraction}} \right)^\mu \quad (12)$$

This equation describes a weighted average of the influence of all of the nodes. The influence of each individual node is the sum of its ‘velocity’ factor and its ‘attraction’ factor. The velocity factor is simply the N_v vector (i.e., the sensorimotor velocity recorded when the node was created). The attraction factor is defined by

$$A(a, N_v) = a - \left(a \cdot \frac{N_v}{||N_v||} \right) \frac{N_v}{||N_v||} \quad (13)$$

and it causes the sensorimotor state to move toward the node. The attraction term is included to cause the system to move toward more familiar regions of sensorimotor space and to help stabilize patterns of repeated behavior (see Egbert and Barandiaran, 2014, 2015). The μ superscript in Equation (12) expresses that the IDSM only (directly) controls the motor components of the sensorimotor state. The sensory components are the result of the robot’s relation to its environment and so are not directly controlled by the IDSM, but are, of course, influenced by the motor dynamics, indirectly through

the sensorimotor contingencies determined by the robot's environment and body.

The influence of each node is attenuated by a non-linear function of the distance between the node and the current sensorimotor state. This attenuation is expressed by the term $d(N_p, \mathbf{x})$ and it means that nearby nodes affect the sensorimotor state and farther away nodes have little influence. The influence is also attenuated by a threshold function of weight, $[\omega(N_w)]$ such that only positively weighted nodes affect the motor state. Previous versions of the IDSM had a more complicated sigmoidal function in place of the simpler threshold function used here (Equation 10). Note that the degradation of the nodes and the threshold function of Equation (10) mean that when nodes are not visited for a long period of time they cease to have any influence whatsoever. Nodes that have degraded to this point essentially cease to exist.

After a node is created, its weight changes according to:

$$\frac{dN_w}{dt} = -k_{\downarrow} + k_{\uparrow}d(N_p, \mathbf{x}) \quad (14)$$

In this equation, the first term represents a steady degradation of the node's influence and the second term represents a strengthening of the node that occurs when the current sensorimotor state is close to the node's position.

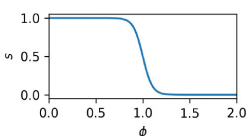
The influence of all nodes is summed and then scaled by the local density of nodes,

$$\phi(\mathbf{x}) = \sum_N \omega(N_w)d(N_p, \mathbf{x}) \Big|_{\text{activated nodes}} \quad (15)$$

This equation looks similar to that used to calculate ψ (Equation 9), but is different in that ψ describes the local density of all nodes, where as ϕ describes the local density of activated nodes *only*.

3.4.4. Random motor activity

When the local density of activated nodes is low, motor behavior is random. This is accomplished by defining of a "switch" value, s , which specifies when the behavior is to be random and when it is to be controlled by the influence of the IDSM's nodes. The following equation expresses that s is 1 when ϕ is low and 0 when ϕ is high; and that it moves between these values in a smooth, sigmoidal manner:

$$s = \frac{1}{1 + \exp(R_g(\phi(\mathbf{x}) - R_t))} \quad (16)$$


Here, R_g and R_t are parameters that specify the threshold of familiarity and the discreteness of the transition between random and non-random motor activity. This value is then used to switch between random motor activity (r), and IDSM's output (m), thus:

$$\frac{d\mu}{dt} = (1 - s)m + sr. \quad (17)$$

where the random motor activity vector, r is varied over time to produce a random walk in motor space as follows: every iteration, there is a R_p chance that the components of r will be assigned random values selected from a normal distribution with a standard deviation of R_σ . The value of the R_σ and other parameters can be found in Table 1.

4. Experiments and results

We now present two computational simulations of this model where we vary the sensory modality of the robot to explore how sensorimotor contingencies constrain the forms of the sensorimotor habits that can emerge and self-stabilize. The simulated robots and their environments are identical except that one robot's sensors are visual (as described above) and the other's are auditory. Formally, the only difference between these simulations is the equations that describe the stimulation of the robot's left and right sensors (Equations 6, 7).

To generate the data presented below, we ran 10 trials of each condition (i.e., we simulated 10 auditory robots and 10 visual robots) and for each condition, we selected a trial that displayed a wide variety of habits, and for which the simulated agent returned to one or more habits that it performed earlier but had stopped performing for some period of time. Not all trials did this—some instead rapidly fell into a pattern of behavior that was stable for the duration of the simulation. Data for all of the simulations is available at [DATA STORE LOCATION]. The analysis below covers visual simulation #9 and auditory simulation #0.

We now present an overview of the behaviors demonstrated by these two robots.

4.1. Visual sensors

The path taken by the visual robot through its environment is displayed in Figure 3A. This is a complicated and difficult to visualize trajectory, but it in fact involves several distinct repeated patterns of sensorimotor behavior (Figures 3B–H). To identify these, we first plotted the proximity of each of the IDSM's nodes to the robot's sensorimotor state over the course of the simulation (Figure 4). In this time series, it is easy to observe segments of time in which a particular set of nodes is repeatedly visited. For instance, when $750 \lesssim t \lesssim 1,200$, there are a few

TABLE 1 Parameters.

Parameter	Value	Description
ROBOT		
R_s	0.25	Robot speed
R_d	0.1	Robot diameter
R_o	$\pi/5$	Offset of sensors from center line of robot
R_f	0.25	Falloff of sensor excitation with distance
LIGHT		
L_v	0.1	Light velocity scale constant
L_r	1/3	The radius of the light's circling motion
IDSM		
k_d	500	Fall-off of non-linear ASMS distance measure
k_{\downarrow}	5	Node weight decay rate
k_{\uparrow}	2000	Node-reinforcement rate constant
k_t	1	Node-density threshold parameter, influencing when new nodes are to be added
N_w^0	1000	Node weight upon creation
MOTOR BABBLING		
R_p	0.1	Chance of random reassignment of random-motor walk direction and velocity
R_σ	0.25	Standard deviation of random motor components
R_t	1	Node-density threshold parameter, influencing when motor activity is random
R_g	20	Steepness of transition between random and IDSM driven motor activity

nodes with indices close to 5,000 that are repeatedly visited (tan horizontal sequence of dots), and the same set of nodes are briefly revisited three times in the final 500 time units of the simulation. We manually identified these repeated patterns of sensorimotor activity, and assigned each pattern a label and a color (the example just provided is labeled “D” and colored tan). Times when the robots behavior is not clearly repeated were not assigned a label and are marked with a light gray color. To be clear, repetitions were identified in sensorimotor space (*via* the identification of repeatedly visited nodes), not in physical space. Returning to [Figure 3](#), we can see that when the robot is performing the D pattern, it is moving around the environment, regularly turning in loops with squarish corners. [Figure 5](#) shows how the state of the sensors, motors and the distance between the stimulus and the robot as the trial progresses. We can see that each of the colored regions tend to occupy particular regions of the sensorimotor space (the sensorimotor ‘habitat’ of the habit—see [Buhmann et al., 2013](#)) and certain patterns of behavior involve the robot pursuing the moving light (e.g., F-behavior) while others avoid it (e.g., D-behavior).

Each of patterns B–H can be construed as a potential habit, i.e., a pattern of autonomous, self-maintaining sensorimotor activity. To justify this claim, we observe that

1. A node can persist for an extended time if and only if it is regularly “visited,” i.e., if and only if the robot’s sensorimotor state regularly comes sufficiently close to the node such that the node’s weight always kept above zero.
2. The robot’s sensorimotor trajectory depends upon the nodes in that (a) the nodes directly control how the motor components of the sensorimotor state change, and (b) the nodes indirectly constrain and influence the sensory components *via* their effect upon the motors.

From these observations it follows that the nodes enable the very thing that they depend upon to persist and reflexively, the enactment of the pattern of behavior enables the very thing that it depends upon (the nodes) for its persistence. Similar to a model of biological autopoiesis ([Varela et al., 1974](#)), where inherently unstable components such as metabolism and membrane are mutually enabling, the components of these patterns of behavior in this model are in a relationship of mutual interdependence and support, and thus can be considered autonomous under some readings of the enactivist literature. The behavior itself is operationally closed entity (a “unity”), constituted by inherently unstable components (the nodes and the sensorimotor trajectory) and yet persists thanks to its own enactment or performance.

More detailed analysis of the operational closure of these systems is outside the scope of the present paper (but see Egbert,

2018 for initial analysis of the precarious autonomy of a simple IDSM-based habit). We can see the basic idea however, when we consider how the weight of the nodes changes as time passes in the simulation (see Figure 6). Recalling that when the weight of a node reaches zero, that node ceases to exist, it is clear that only nodes that are regularly visited can persist in the long term. For a collection of nodes to be visited, the sensorimotor trajectory must move in a particular trajectory and the sensorimotor trajectory is largely determined by the activity of nodes. It follows from this that the only way that a pattern of behavior can persist is if it is one that causes the repeated revisitation of its constituent nodes. We can also note that this self-reinforcement of a habit need not be constant or contiguous. For example, the B-habit (red) is established early in the simulation [$t \approx 250$] and then is not visited again until $t \approx 1,250$, where it is enacted a few times and the nodes are reinforced such that they survive until close to the end of the simulation. The parameters that prescribe the rates of node weight reinforcement (k_{\uparrow}) and degradation (k_{\downarrow}) determine how regularly patterns of behavior must be enacted if they are to persist.

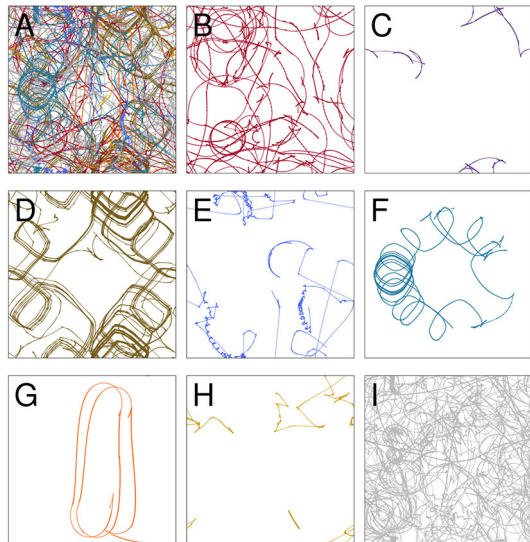


FIGURE 3
The trajectory taken by the visual robot (A), broken down into 8 patterns of repeated sensorimotor activity, i.e., habits (B–H). The portions of the trajectory where the robot's behavior is not clearly repetitious are included in plot (I). Each of the nine plots shows the full 1×1 arena. Video showing this trial is available at <https://www.youtube.com/watch?v=2v1TyvKz9qw>.

4.2. Auditory sensors

We performed a similar analysis on the auditory robot. Figure 7A shows the complete path taken by the auditory robot in its environment, broken down by habit (Figures 7B–H). Figure 8 shows the weight of each node which also provides

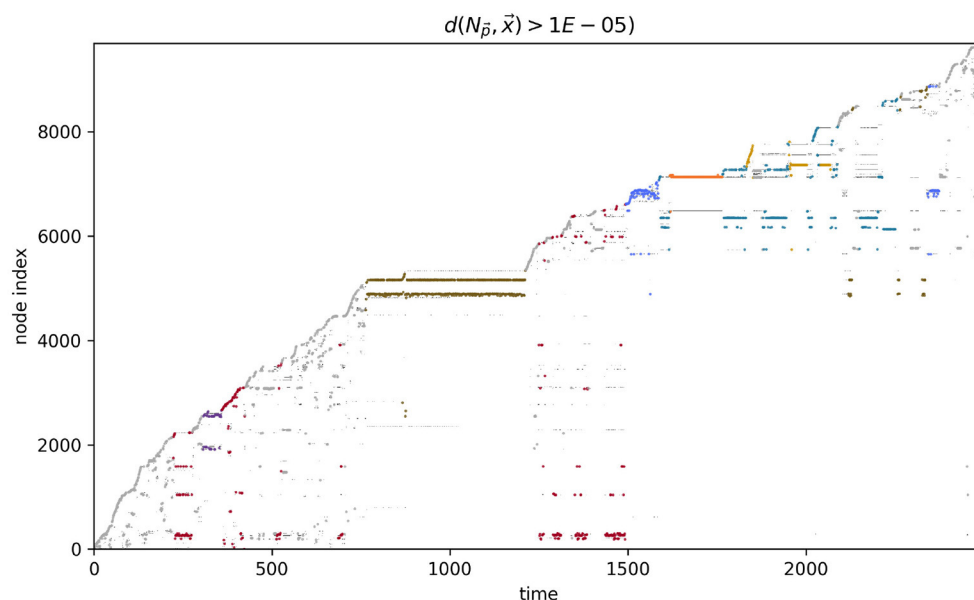
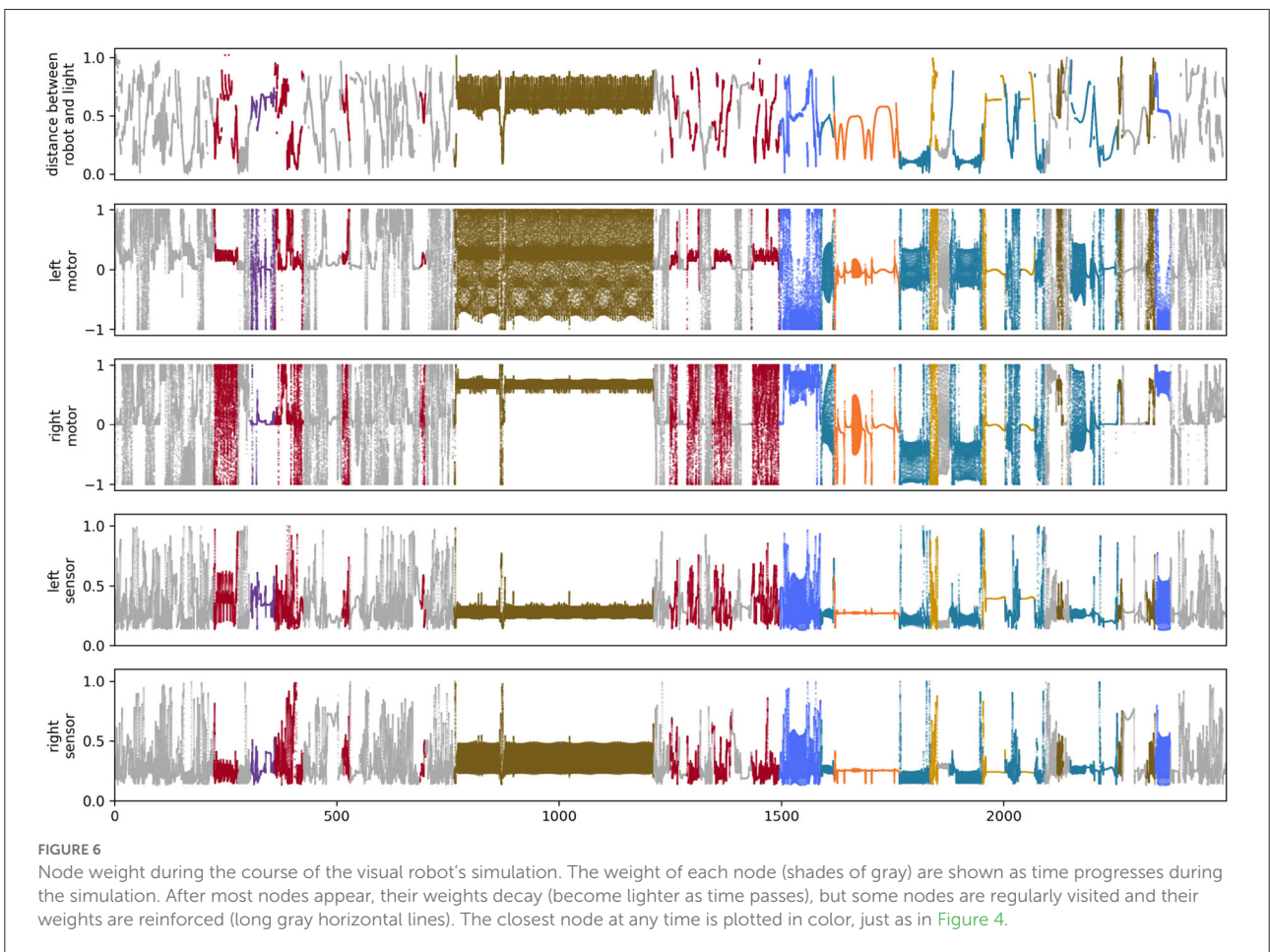
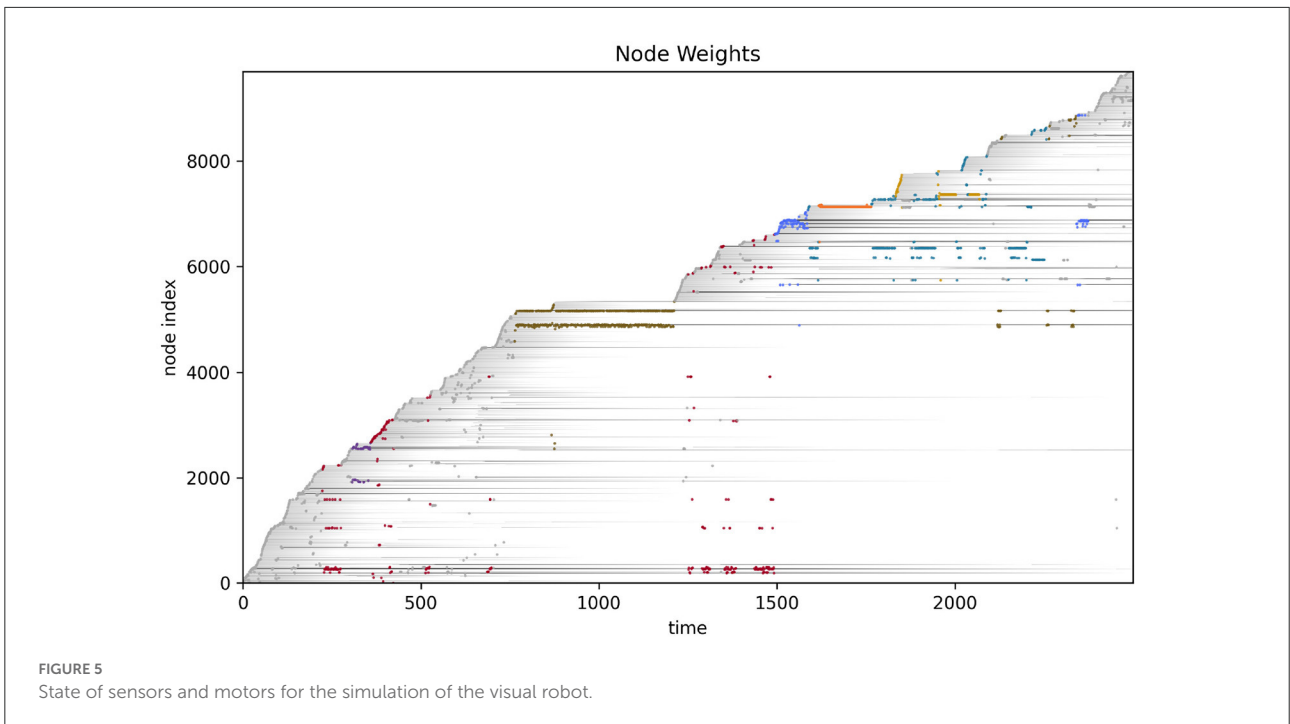
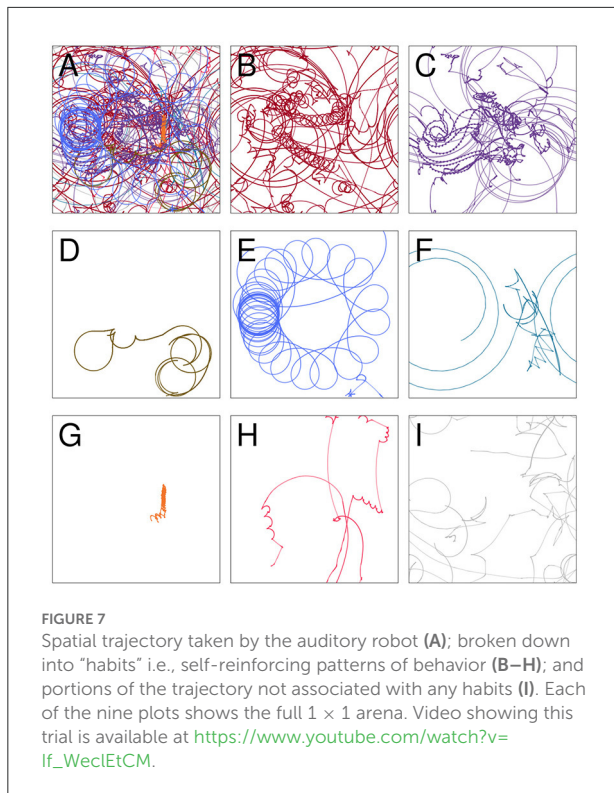


FIGURE 4
Node proximity during the course of the visual robot's simulation. The darker the point in this plot, the closer that node is to the robot's sensorimotor state at that point in time. The larger colored points indicate the index of the node that was closest (in sensorimotor state) to the robot's sensorimotor state at that time. The color of the closest-node points indicates the habit that those nodes are associated with. These colors match those in Figure 3.





an indication for how close each node is to the current sensorimotor state of the robot and the timeseries plots for the auditory robot are shown in Figure 9. It is worth noting that though these patterns of behavior may seem random, they all (both visual and auditory) relate to the moving stimulus source in regular, non-random ways. The regularities in the interaction with the stimulus are much more easily seen in the animations linked to in the captions of Figures 3, 7).

4.3. Comparison of sensorimotor structures

Even when two behavior's functions are the same (e.g., they both accomplish taxis) the forms of the underlying sensorimotor habit can be different. As a case in point, we can compare auditory habit E with visual habit F. In both cases, the robot circles around the stimulus, maintaining a similar approximate distance from the stimulus as the stimulus moves in spurts around the environment. When we look at the two patterns of sensorimotor activity (Figure 10), we see two different pictures. Each row of this visualization shows the spatial trajectory of the robot (left) followed by four projections of the 4D sensorimotor trajectory of the robot and it is obvious that the two sensorimotor trajectories are qualitatively different. The auditory pattern involves less diversity in the range of states visited; the mean and other statistical properties of the

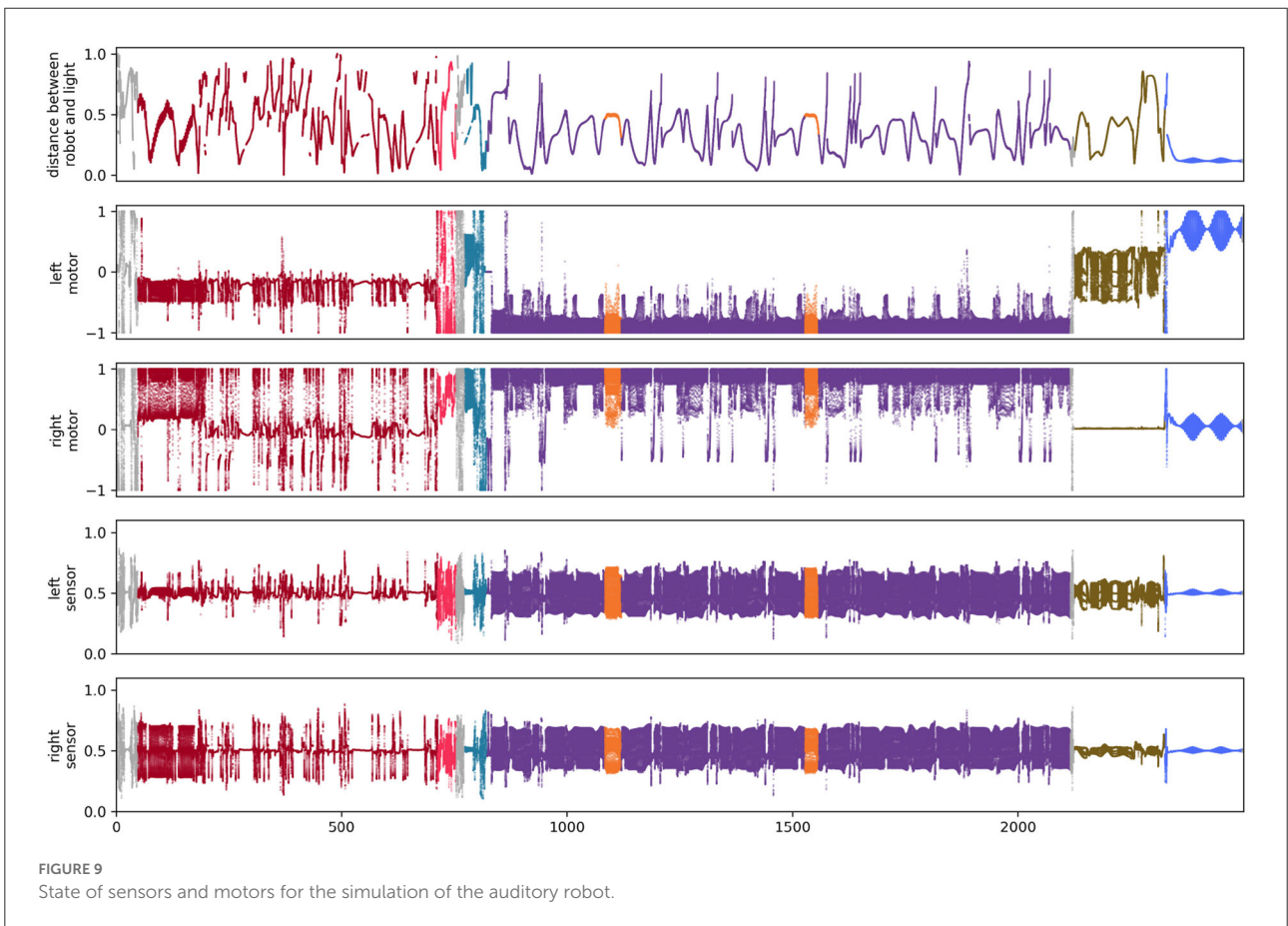
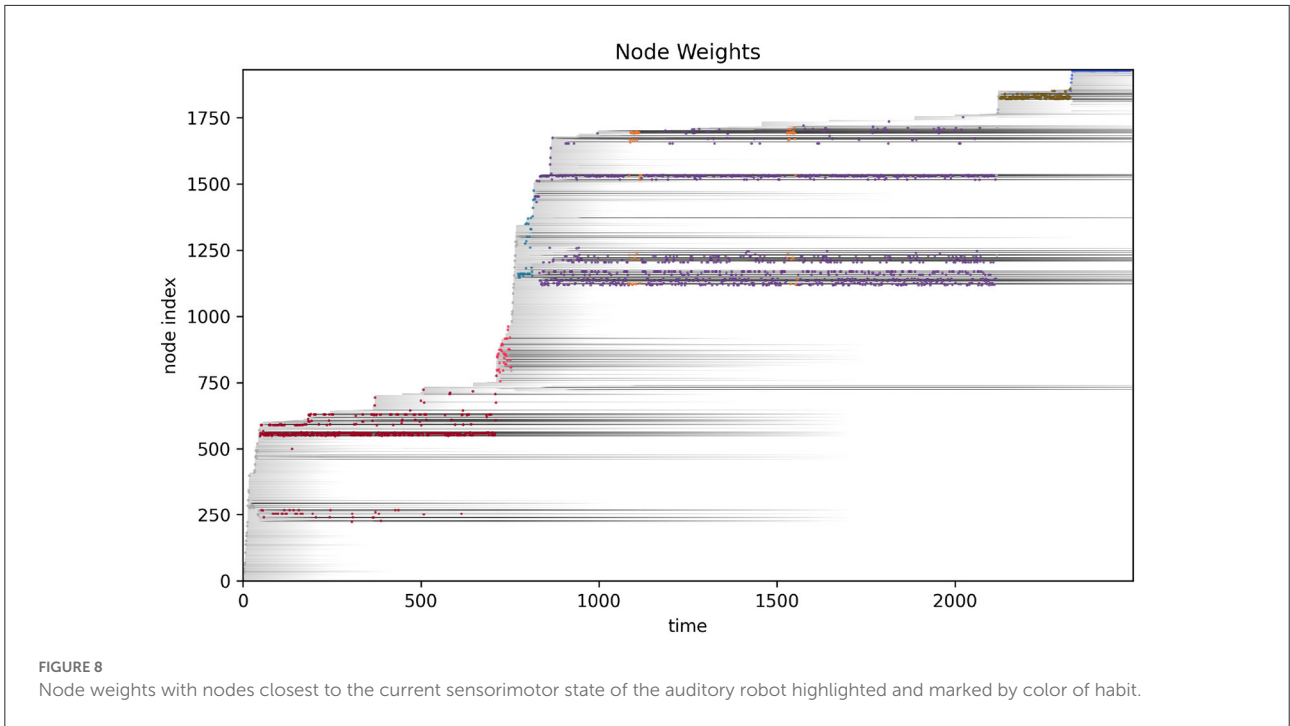
sensorimotor trajectories are also clearly different, etc. But every habit in this model is unique—how much of the difference between these habits is simply due to the fact that they are different habits, and how much is due to the difference in the robot's sensorimotor modalities?

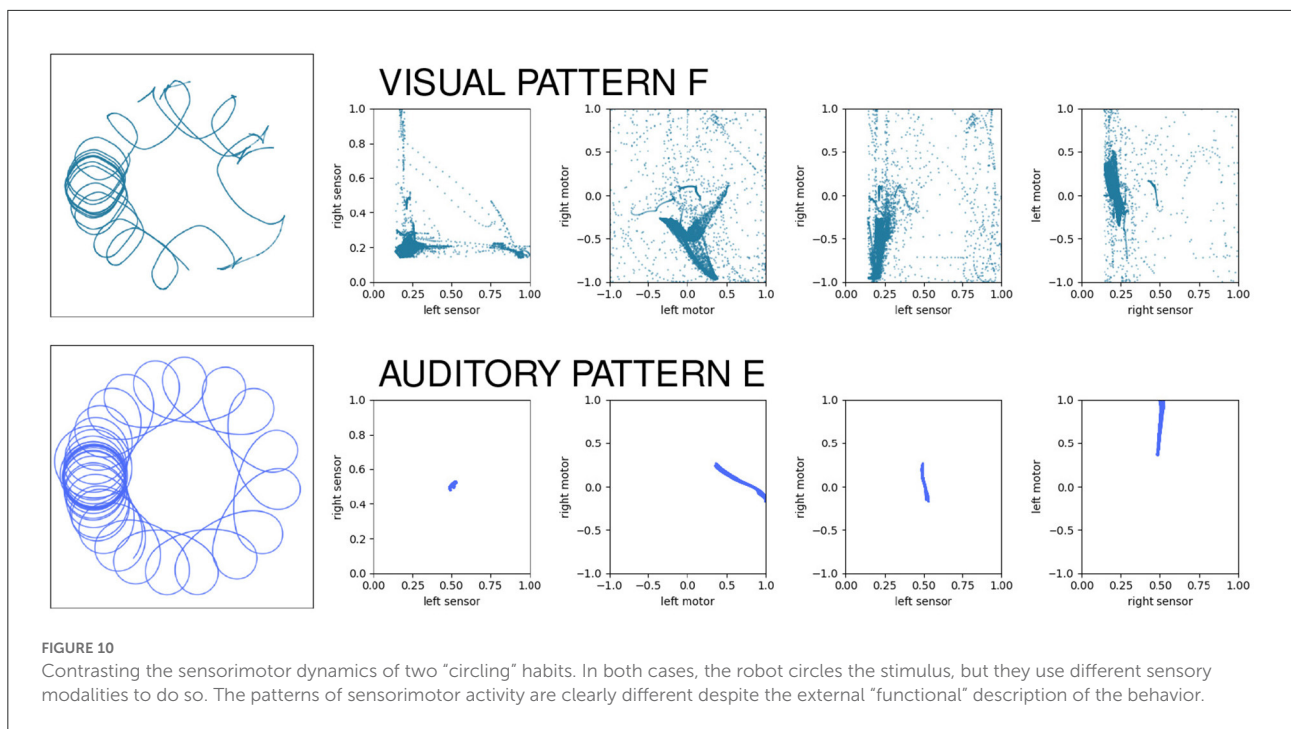
To address this question, we can look at projections of the sensorimotor trajectories of all of the habits of the visual robot (Figure 13) and observe that there are regularities in the forms of the habits of this robot. This is most readily seen in the left-sensor vs. right-sensor projection (upper right) which reveals elements of this robot's sensorimotor contingencies, including aspects of both its sensorimotor habitat and environment (Buhrmann et al., 2013) of the robot. Specifically, the bilateral symmetry of the robot's embodiment result in a (statistical) mirror symmetry across the left sensor = right sensor diagonal. The directionality of the visual sensors and their different orientations mean that it is impossible (in the environment with a single stimulus) to maximally stimulate both left and right sensors concurrently, leading to the L-shape apparent in upper-right plot of Figure 13, with arcs that pass between the stems of the L as the robot turns such that the stimulus moves from being in front of one sensor to in front of the other. These features are also found in the other nine visual robot trials.

Similar plots for the auditory robots (Figure 14) also reveal regularities. Like the visual robot, the auditory robot's embodiment is bilaterally symmetric and so it has the same symmetry across the diagonal in left-sensor/right-sensor plot, but the non-directionality of the auditory sensors and the fact that their excitation does not fall off with distance make other aspects of the sensorimotor activity different. The auditory trajectories tend to repeat patterns of diverging from a center point in short arcs in a pattern of returning and diverging again, in a different direction. These dynamics can be seen in greater detail in Appendix A, which shows the projection in sensorimotor space of each habit individually.

Sensors and motors are involved in a recursive relationship of influence, where sensors influence motors (*via* the controller) and motors influence sensors (*via* the effect of actions upon the environment and the agent's relationship to the environment). It therefore would make sense that the constraints imposed upon sensory dynamics by different sensory modalities would constrain the motor dynamics. In other words, under one sensory modality, certain motor trajectories will be more readily repeated (and thus stabilized) than others, and which motor trajectories are more readily repeated would depend upon the sensory modality.

We do, in fact, see differences in the distributions of motor states between the two sensory conditions (Figure 11). Visually, we can see that the distributions of motor (and abstract motor) states in the visual robots seem different than that of the auditory robots (Figure 11); and the distributions of motor states within these groups in independent runs of the visual trial seem more similar to each other (Figure 12) than they are





to the other group. We confirmed the statistical significance of these differences by using Kullback–Leibler divergence to assess the distributional conformity of 50 visual trials and 50 auditory trials. The within-group (intra-modal) distributions were significantly more similar than the between-group (inter-modal) distributions for m ($t = -4.67, p = 0.0114$) and for μ ($t = -8.507, p = 0.0215$).

In the absence of the self-reinforcing dynamic of the IDSM, habits, motor activity would be random, in which case we would expect the same distribution of motor states in both sensory modalities. The histograms in (new) (Figures 11, 12) show that this is not the case, and thus provide support for the claim that the sensorimotor contingencies implicit in the different sensory modalities constrain not only the sensory dynamics, but also motor dynamics—and in doing so, they play an important role of constraining the form of emergent autonomous sensorimotor dynamics.

The primary point that we wish to communicate using this model is that sensorimotor contingencies influence which patterns of autonomous sensorimotor behavior can emerge and persist. It is worth emphasizing that the plots in Figures 10, 13, 14 are not plots of the sensorimotor contingencies of the robot—they are plots of the robots’ self-maintaining habits. Nevertheless, we see in these plots the *influence* of sensorimotor contingencies—the way that the contingencies have constrained the set of habits that can form and self-stabilize, producing the regular patterns described in the paragraphs above.

The two different sensory modalities entail two different sets of sensorimotor contingencies. The different sensorimotor

contingencies each imply a different set of possible self-stabilizing habits, and so the habits that form with one sensory modality are more similar to each other than they are to the habits that form with the other sensory modality. Two *visual* habits will be different from each other, but they will generally be more similar to each other than auditory habits. This is what we mean by the claim that the form of sensorimotor habits are constrained by their sensory modality and it is apparent in the regularities described above and seen in Figures 13, 14. In this account, sensorimotor contingencies play an important role. They influence the form of the emergent autonomous sensorimotor structures, but the rules themselves are not internalized or learned in any other way.

5. Discussion

5.1. Solving somebody else’s problems vs. having your own problems

The robots we have just presented display a number of interesting and challenging features for traditional problem-solving centered artificial intelligence (biologically inspired or otherwise). We see in this model how sensorimotor dynamics can have problems of their own. As we have seen AI and robots more specifically are generally built to solve problems. But whose problems? Perhaps the engineers’ problems or those of their clients, perhaps animal problems that require robotic models to be better understood, maybe societies problems but certainly

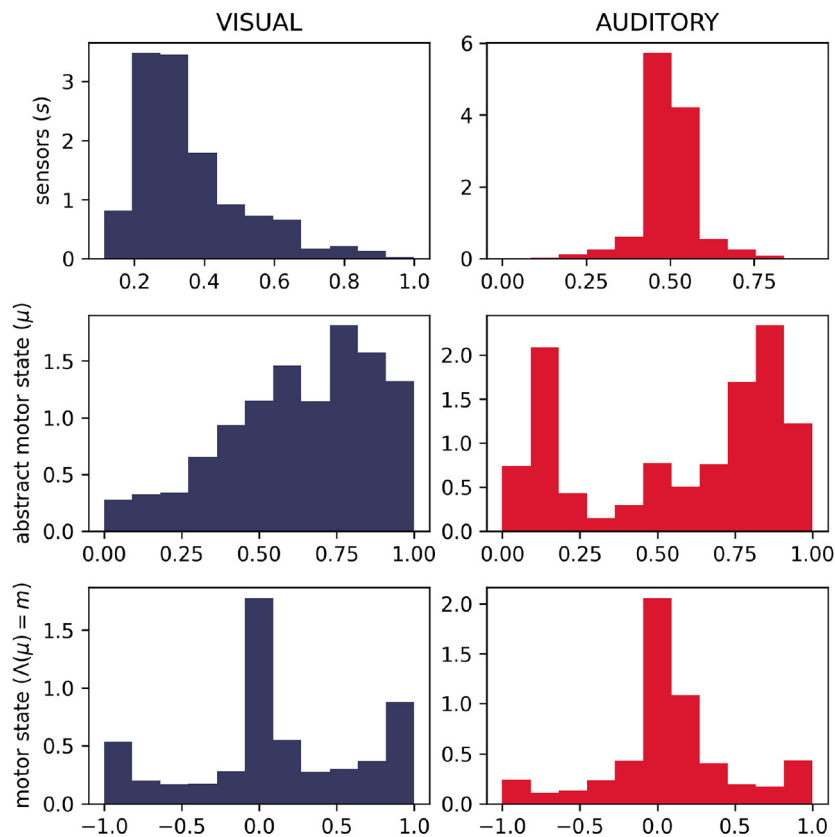


FIGURE 11

Histograms showing difference in distributions of sensorimotor states for 50 visual and 50 auditory trials. As expected, the different sensory modalities produce different distributions of the sensor state. It can also be seen that in visual trials, the abstract motor-state tended to be higher (corresponding to faster rates of motor state change—see Figure 2) and the motor state tended to be more extreme.

not those of the robot itself. It is thus important to distinguish between having one's problem vs. being able to solve a problem. A bacterium that swims up a gradient toward the resources it needs to survive is doing both. It has a problem (it needs food to persist) and it solves that problem by navigating through its environment. AI, on the other hand, is capable only of the former. A "self-driving" car that takes me to a restaurant dinner is solving a problem (my need for food), but that problem is mine, and not the car's.

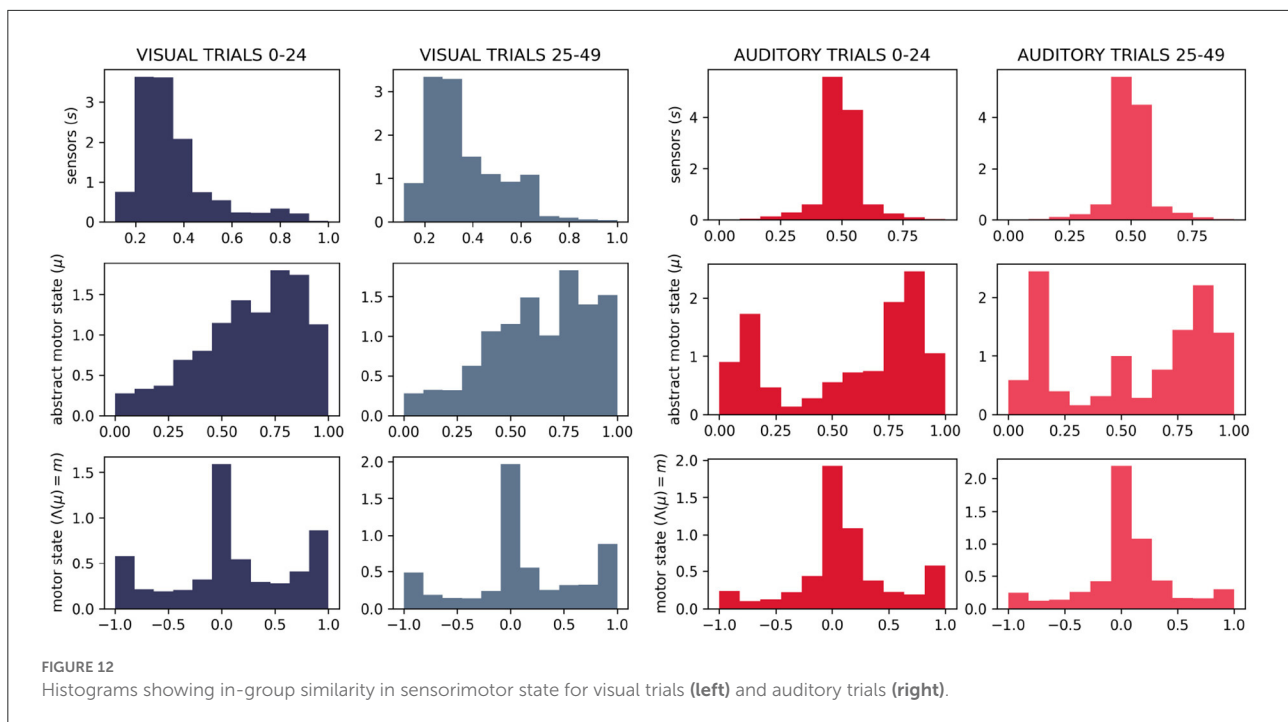
The enactive robots presented in this paper were not designed to do anything, to solve any problem. None of our analysis was described in terms of success or failure at the performance of a task. The architecture makes habits emerge and with them a basic sense of having a problem: the habit *needs* to enact itself in order to persist. The norms of behavior here are thus only in terms of the persistence of the behavior itself (and concurrently the mechanism that produces it). Surely, life is full of problems, we certainly have enough with our own and don't need to be concerned with creating new ones ... unless we want to understand what it is to have a problem and how life is, itself, a source of problems. Enactivism embraces a deep conception of

life as the self maintenance of a network of precarious processes. In this sense life is inherently a source of problems. By making it possible for habits to emerge and sustain their own existence through recurrent sensorimotor interaction we have shown how this conception of life can be practically transferred to a "way of life" in sensorimotor robotics.

A research program on enactive robotics can and must address how sensorimotor life is a source of problems. We can move beyond the single habit and envision how the network of habits that constitutes the agents can display emergent problems of its own: that of keeping the whole network alive, coping with variations on the way the environment affords or precludes to enact them. This line of enquiry requires further work but the present model requires little development to start addressing it.

5.2. Constraints and form in biological and behavioral explanation

Biological inspiration in robotics can also move beyond the analogy between metabolic or physiological life and its



sensorimotor counterpart. In particular the models presented and analyzed in this paper show how sensorimotor life and the problems that emerge within are constrained by sensory embodiment and how such constraints constitute an important part of what it means to explain life.

Whether it is understood as more fundamental-than, complementary-to or directly at odds-with (Newman, 2018) natural selection, it is the repeated, regular, robust and often phylogenetically independent appearance of forms or structures in biology what becomes in itself an object of study that cannot be reduced-to or deduced-from life as a response to environmental problems. In this sense biologically inspired enactive robotics can learn from biological development and its organization and import into psychology some explanatory strategies followed by biologists. Darcy Thompson, René Thom, Pere Albert, etc. all conceived biological forms to be both *explananda* and *explanans* on themselves. Homologies, analogies, inherency, are all concepts directed to capture the re-emergence of certain forms (and functions!) in biological organisms. The notion of constraint plays here a fundamental role. The way in which different layers of materiality and self-organization limit and channel the emergence of viable forms is essential to biological explanation.

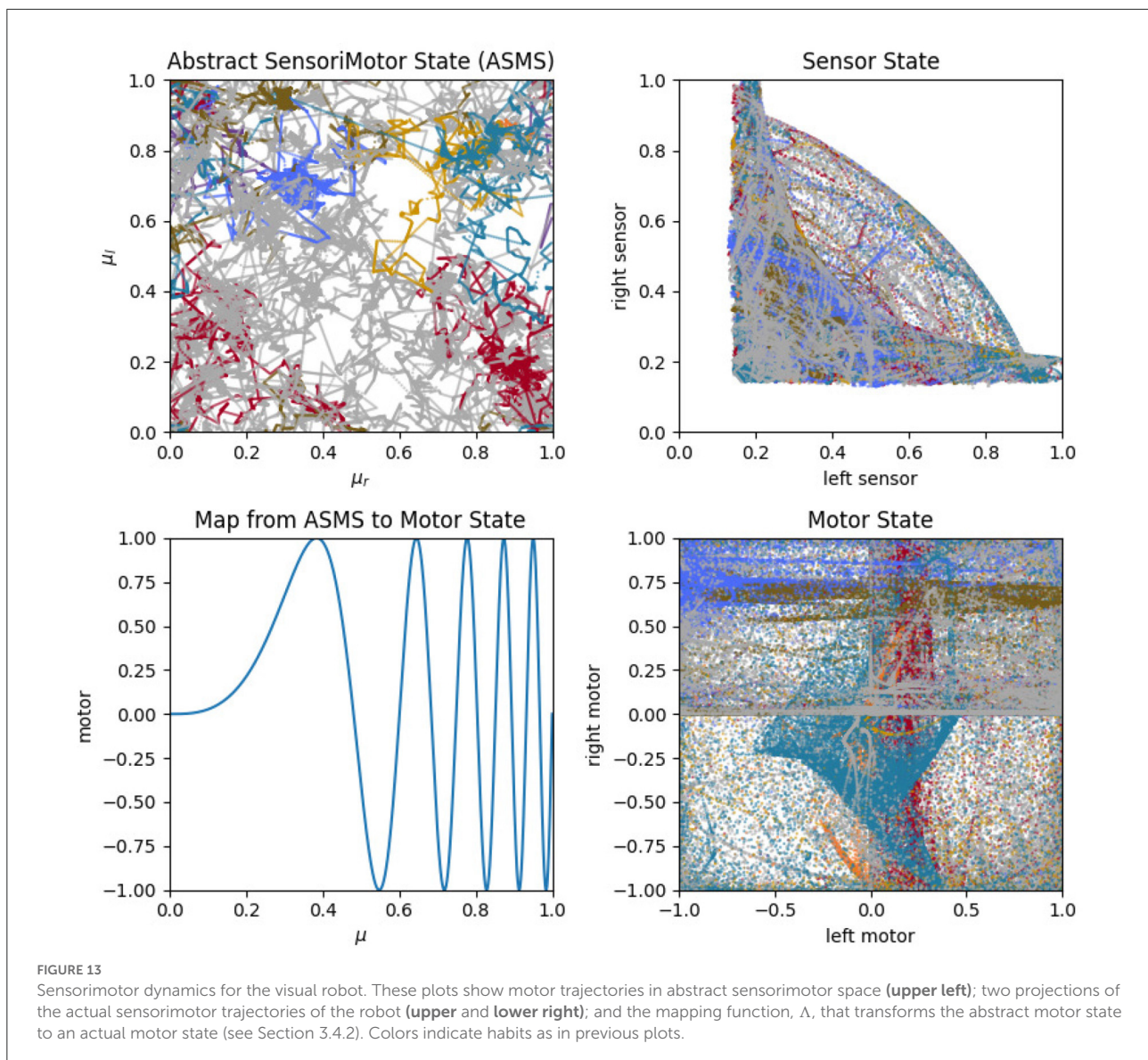
If biological inspiration within the problem-solving paradigm concluded that the material embodiment of robots permits to offload and transform the problems to be solved, the biological inspiration of the enactive paradigm can conclude that the embodiment loads and informs the problems that constitute the sensorimotor agent. As we

have seen, sensory modalities constraint motor trajectories, which in turn, shape how habits get stabilized. Thus, in a way that parallels the explanatory role of constraints in evolution, we can hypothesize that behavioral variability is not free (to explore potential solutions to cognitive problems) but is constrained and channeled by the embodiment of the agent.

Now, before the advent and widespread influence of evolutionary theory in biology, the latter synthesis with molecular biology and subsequent expansion to psychology, sociology and even epistemology, or philosophy more generally, the concept that was key for the continuity between life and mind was that of habit.

5.3. Enactive robotics revisited

There are different takes on how enactivism translates into robotics. As mentioned in the introduction it is possible to simply reject representational functionalism, or to claim for the importance of embodiment, or to demand that robot be endowed with living bodies, or to introduce some feedback mechanisms that parallel those provided by emotion bearing bodies. We have taken a different approach—that envisioned by Smithers (1997) and latter developed by Di Paolo (2003), Barandiaran and Moreno (2006), and Barandiaran (2008). What a research programme in enactive robotics entails is the study of the organization of sensorimotor life: the form of viability of different habits, the topology of the network of habits that

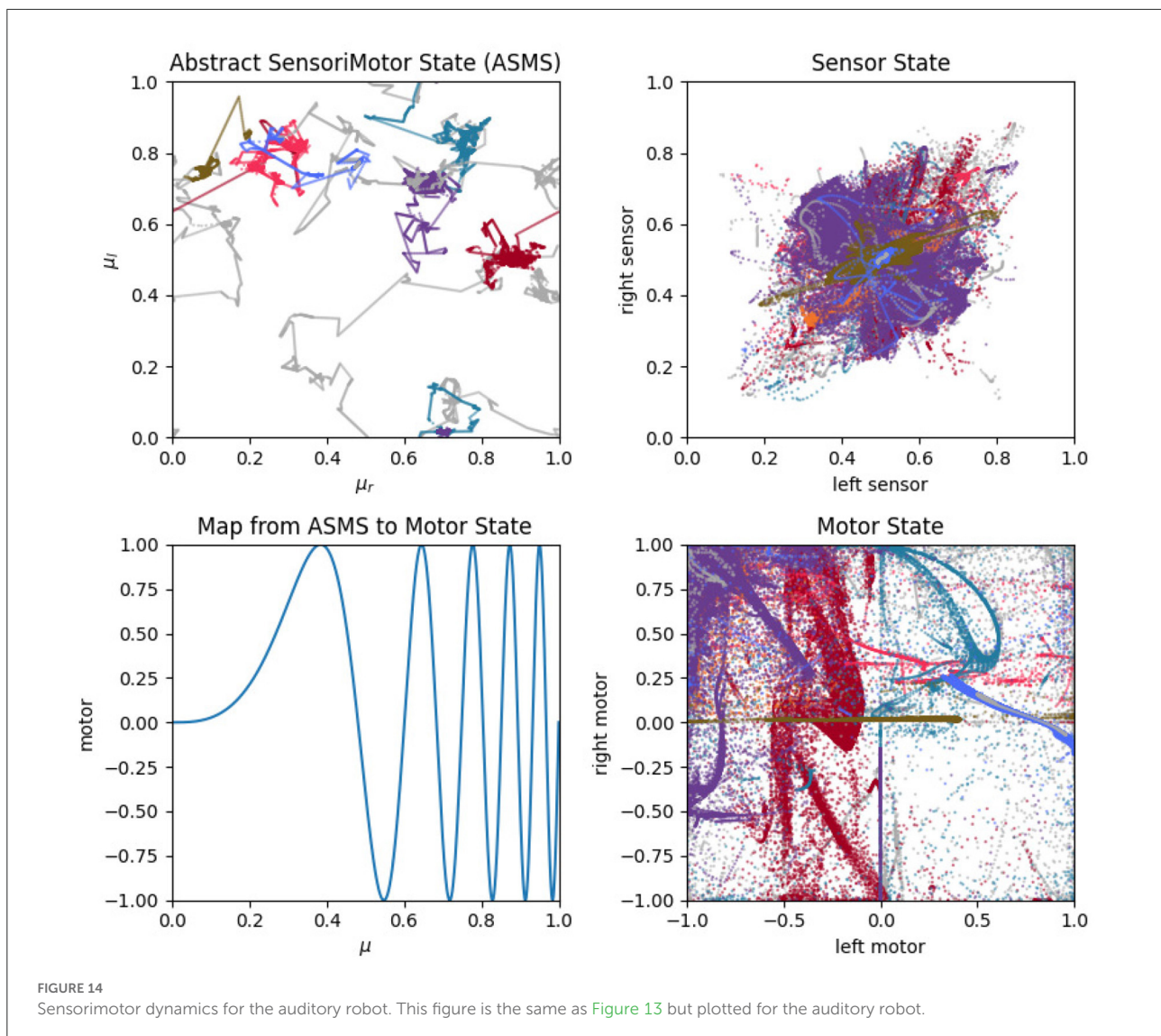


unfolds over development, the shape of the habitat that is thus constructed, the structure of the world that is experienced. Habit-based robots as designed here are capable of individuating habits and of creating an ecology of habits that can easily be understood as a form of sensorimotor self.

Not only has enactivism informed robotics (Ziemke and Lowe, 2009; Vernon, 2010) but robotics has often served enactivism (Beer, 1995; Di Paolo, 2003; Aguilera et al., 2016) by clarifying its claims, pushing theoretical development, operationalizing its concepts or penetrating diverse problems. The model we developed here can be aligned with the latter. It can be used to clarify and make explicit the often obscure original formulation of enactivism that “cognition in its most encompassing sense consists in the enactment or bringing forth of a world by a viable history of structural coupling” (Varela et al., 1991, p. 205). We have shown how a robot endowed with an

IDSMS can bring into being a number of habits, the world that it brings forth is the habitat, more specifically, the structured set of sensorimotor contingencies that the agent inhabits or enacts. This habitat must be viable in the sense that habits must be sustainable and results from a (developmental) history of sensorimotor (structural) coupling. And in so doing, we have demonstrated how sensorimotor contingencies can *directly* constrain or “sculpt” the form taken by sensorimotor habits without requiring any virtualization, i.e., without any internal model or representation of the sensorimotor contingency.

Perhaps it is worth clarifying why the nodes do not constitute an internal model or representation. The nodes do not stand-for something else other than the sensorimotor dynamics they partake in. It is not possible to operate upon the nodes in a decoupled offline mode that is not itself the enactment of a behavior, and there is not additional module or subsystem



“consuming” such nodes to carry out any further operation. In these ways, the nodes do not represent a behavior or habit, but they embody it—they constitute the habit *together* with elements of the robot’s bodily and environmental dynamics.

6. Conclusion

A central focus across the cognitive sciences is upon problem-solving ability and tremendous progress has been made in understanding how to mechanize problem solving. Much of AI and robotics research is validated by how well some artifact (neural network, human being, robot, etc.) performs at a problem-solving task (chess, maze navigation, bipedal walking, etc.).

However, the conflation of “problem-solving ability” with all of the phenomena associated with “being a mindful body leaves out a number of features that demand be put at

the center of (enactive) theorizing: historicity, embodiment, habitat, precariousness, identity, norm-establishing, etc. All these dimensions of mental life are worthy of study and remain outside of the problem-solving frame that scaffolded the development of Artificial Intelligence and Robotics.

We have here presented a set of robots with different sensory modalities that spontaneously develops a complex ecology of sensorimotor habits. These are constrained by the sensory modality of the robot and give rise to sensorimotor habitats of specific forms. The very nature of habits thus developed, understood as self-sustaining forms of sensorimotor activity, has shown how robots must first have their own problems instead of solving those posited by external observers; and, that, in doing so, they must assert a form of life whose structure and topology must be taken as the object of study. In particular we have seen how the form of sensory embodiment shapes potential sensorimotor contingencies and these constraint the shape and type of sensorimotor habits that emerge during development.

To be alive is not a computable function but the way in which materiality (implementation), behavior and function are deeply intertwined. What was once claimed as the triumph of functionalism as the clear conceptual separation between matter, behavior and machine state transitions is now its deepest weakness. The enactive approach brings all three together again. To be fair it is not the materiality of the robot's "body" that is a stake here (neither does the simulation itself possess any body beyond the computer in which the simulation was carried out, not the physical body, e.g. wheels and sensors, would be at stake was the robot to be implemented in real life), but the materiality of the sensorimotor mapping, its precarious existence, its fading structural stability.

Perhaps the "artificial sciences" (AI, artificial life, robotics, exploratory computational modeling, etc.) would benefit from similarly investing more time in targets other than problem-solving ability. Biologists are sometimes accused of suffering "physics envy"—i.e., wishing that the objects of their study were more easily and completely summarized by simple, provable equations. Perhaps we enactivist and embodied researchers can be accused of "problem-solving envy" a desire for our artifacts and theories to be equally or more capable of solving problems as the expert systems or disembodied neural networks of other problem-solving focused approaches. And perhaps this envy is a distraction, and impeding our progress toward understanding minds. In fact, robotics cannot only reveal itself as an engineering practice directed at solving problems but also as a philosophical practice aimed at posing the right problems. This is a contribution that enactive robotics is ready to do.

7. Permission to reuse and copyright

Figures, tables, and images will be published under a Creative Commons CC-BY license and permission must be obtained for use of copyrighted material from other sources (including re-published/adapted/modified/partial figures and images from the internet). It is the responsibility of the authors to acquire the licenses, to follow any citation instructions requested by third-party rights holders, and cover any supplementary charges.

References

- Aguilera, M., Bedia, M. G., and Barandiaran, X. E. (2016). Extended neural metastability in an embodied model of sensorimotor coupling. *Front. Syst. Neurosci.* 10, 76. doi: 10.3389/fnsys.2016.00076
- Ashby, W. R. (1952). *Design for a BRAIN: The Origin of Adaptive Behaviour, 2nd Edn.* London: J. Wiley.
- Barandiaran, X., and Moreno, A. (2006). On what makes certain dynamical systems cognitive: a minimally cognitive organization program. *Adapt. Behav.* 14, 171–185. doi: 10.1177/105971230601400208

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

ME conceived, implemented, ran simulations and experiments, and authored the manuscript. XB contextualized research, interpreted results, and authored the manuscript. Both authors contributed to the article and approved the submitted version.

Funding

This work was supported in part *via* funding from the Digital Life Institute, University of Auckland. XB acknowledges funding from the Spanish Ministry of Science and Innovation for the research project Outonomy PID2019-104576GB-I00 and IAS-Research group funding IT1668-22 from Basque Government.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Barandiaran, X. (2007). "Mental life: conceptual models and sythetic methodologies for a post-cognitivist psychology," in *The World, the Mind and the Body: Psychology after Cognitivism*, eds B. Wallace, A. Ross, T. Anderson, and J. Davies (Exeter: Imprint Academic), 49–90.

- Barandiaran, X. E. (2008). *Mental Life. A Naturalized Approach to the Autonomy of Cognitive Agents* (Ph.D. thesis). University of the Basque Country, Spain.

- Barandiaran, X. E. (2017). Autonomy and enactivism: towards a theory of sensorimotor autonomous agency. *Topoi* 36, 409–430. doi: 10.1007/s11245-016-9365-4

- Barandiaran, X. E., and Egbert, M. D. (2013). Norm-establishing and norm-following in autonomous agency. *Artif. Life* 20, 5–28. doi: 10.1162/ARTL_a_00094
- Beer, R. D. (1995). A dynamical systems perspective on agent-environment interaction. *Artif. Intell.* 72, 173–215. doi: 10.1016/0004-3702(94)00005-L
- Buhrmann, T., Di Paolo, E. A., and Barandiaran, X. (2013). A dynamical systems account of sensorimotor contingencies. *Front. Psychol.* 4, 285. doi: 10.3389/fpsyg.2013.00285
- Buss, D. M. (1998). *Evolutionary Psychology: The New Science of the Mind*. London: Allyn and Bacon.
- Canaan, R., Salge, C., Togelius, J., and Nealen, A. (2019). “Leveling the playing field: fairness in AI versus human game benchmarks,” in *Proceedings of the 14th International Conference on the Foundations of Digital Games, FDG '19* (New York, NY: Association for Computing Machinery), 1–8.
- Cliff, D., Husbands, P., and Harvey, I. (1993). Explorations in evolutionary robotics. *Adapt. Behav.* 2, 73–110. doi: 10.1177/105971239300200104
- Damiano, L., and Stano, P. (2021). A wetware embodied AI? Towards an autopoietic organizational approach grounded in synthetic biology. *Front. Bioeng. Biotechnol.* 9, 724023. doi: 10.3389/fbioe.2021.724023
- Di Paolo, E. A. (2003). “Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop,” in *Dynamical Systems Approach to Embodiment and Sociality*, eds K. Murase and T. Asakura (Adelaide, SA: Advanced Knowledge International), 19–42.
- Di Paolo, E. A., Buhrmann, T., and Barandiaran, X. (2017). *Sensorimotor Life: An Enactive Proposal, 1st Edn*. Oxford: Oxford University Press.
- Egbert, M. D., and Barandiaran, X. E. (2014). Modeling habits as self-sustaining patterns of sensorimotor behavior. *Front. Hum. Neurosci.* 8, 590. doi: 10.3389/fnhum.2014.00590
- Egbert, M. D., and Barandiaran, X. E. (2015). Corrigendum: Modeling habits as self-sustaining patterns of sensorimotor behavior. *Front. Hum. Neurosci.* 9, 209. doi: 10.3389/fnhum.2015.00209
- Egbert, M. D., and Cañamero, L. (2014). “Habit-based regulation of essential variables,” in *Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, eds H. Sayama, J. Rieffel, S. Risi, R. Doursat, and H. Lipson (New York, NY: MIT Press), 168–175.
- Egbert, M. D. (2018). “Investigations of an adaptive and autonomous sensorimotor individual,” in *The 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE)*, eds T. Ikegami, N. Virgo, O. Witkowski, M. Oka, R. Suzuki, and H. Iizuka (New York, NY: MIT Press), 343–350.
- Fodor, J. A. (1968). *Psychological Explanation: An Introduction to the Philosophy of Psychology*. New York, NY: Random House Inc.
- Gánti, T. (1975). Organization of chemical reactions into dividing and metabolizing units: the chemotons. *Biosystems* 7, 15–21. doi: 10.1016/0303-2647(75)90038-6
- Graves, A., and Schmidhuber, J. (2009). “Offline handwriting recognition with multidimensional recurrent neural networks,” in *Advances in Neural Information Processing Systems, Vol. 21*, eds D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou (New York, NY: Curran Associates, Inc.)
- Grey Walter, W. (1950). An imitation of life. *Sci. Am.* 182, 42–45. doi: 10.1038/scientificamerican0550-42
- Hambrick, D. Z., Burgoyne, A. P., and Altmann, E. M. (2020). “Problem-solving and intelligence,” in *The Cambridge Handbook of Intelligence, Cambridge Handbooks in Psychology, 2nd Edn*, ed R. J. Sternberg (Cambridge: Cambridge University Press), 553–579.
- Harvey, I., Husbands, P., Cliff, D., Thompson, A., and Jakobi, N. (1997). Evolutionary robotics: the sussex approach. *Rob. Auton. Syst.* 20, 205–224. doi: 10.1016/S0921-8890(96)00067-X
- Jacquey, L., Baldassarre, G., Santucci, V. G., and O'Regan, J. K. (2019). Sensorimotor contingencies as a key drive of development: from babies to robots. *Front. Neurobot.* 13, 98. doi: 10.3389/fnbot.2019.00098
- Luisi, P. L. (2006). “The emergence of life,” in *From Chemical Origins to Synthetic Biology* (Cambridge: Cambridge University Press).
- Maes, P. (1990). *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*. Cambridge, MA: MIT Press.
- Man, K., and Damasio, A. (2019). Homeostasis and soft robotics in the design of feeling machines. *Nat. Mach. Intell.* 1, 446–452. doi: 10.1038/s42256-019-0103-7
- Maturana, H. R., and Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: Springer.
- Maye, A., and Engel, A. K. (2013). Extending sensorimotor contingency theory: prediction, planning, and action generation. *Adapt. Behav.* 21, 423–436. doi: 10.1177/1059712313497975
- MLPerf (2021). *MLCommons*. Available online at: <https://mlcommons.org/>
- Newell, A., Shaw, J. C., and Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychol. Rev.* 65, 151–166. doi: 10.1037/h0048495
- Newell, A., and Simon, H. A. (1976). Computer science as empirical inquiry: symbols and search. *Commun. ACM* 19, 113–126. doi: 10.1145/360018.360022
- Newman, S. A. (2018). “Inherency,” in *Evolutionary Developmental Biology*, eds L. Nuno de la Rosa and G. Müller (Cham: Springer International Publishing), 1–12.
- Noë, A. (2004). *Action in Perception*. Cambridge, MA: The MIT Press.
- Nolfi, S., Bongard, J., Husbands, P., and Floreano, D. (2016). “Evolutionary robotics,” in *Springer Handbook of Robotics* (Berlin: Springer), 2035–2068.
- Norvig, P., and Russell, S. (2021). *Artificial Intelligence: A Modern Approach, Global Edition*. Harlow: Pearson.
- O'Regan, J. K., and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–973; discussion 973–1031. doi: 10.1017/S0140525X01000115
- Pfeifer, R., and Scheier, C. (2001). *Understanding Intelligence*. Cambridge, MA: MIT Press.
- Pinker, S. (1997). *How the Mind Works*. New York, NY: W. W. Norton & Company.
- Putnam, H. (1965). “The mental life of some machines,” in *Philosophical Papers: Volume 2, Mind, Language and Reality* (Cambridge: Cambridge University Press), 408–428.
- Rosen, R. (1991). *Life Itself, 1st Edn*. New York, NY: Columbia University Press.
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* 65, 386. doi: 10.1037/h0042519
- Ruiz-Mirazo, K., Pereto, J., and Moreno, A. (2004). A universal definition of life: autonomy and open-ended evolution. *Origins Life Evolut. Biospheres* 34, 323–346. doi: 10.1023/B:ORIG.0000016440.53346.d
- Rumelhart, D. E., McClelland, J. L., and Group, P. R. (1988). *Parallel Distributed Processing, Vol. 1*. Cambridge, MA: MIT Press.
- Simon, J. R., and Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus-response correspondence and age. *Ergonomics* 6, 99–105. doi: 10.1080/00140136308930679
- Smithers, T. (1997). Autonomy in robots and other agents. *Brain Cogn.* 34, 88–106. doi: 10.1006/brcg.1997.0908
- Steels, L. (1990). “Towards a theory of emergent functionality,” in *Proceedings of the First International Conference on Simulation of Adaptive Behavior on From Animals to Animats* (Cambridge, MA: MIT Press), 451–461.
- Varela, F. J. (1992). “Whence perceptual meaning? A cartography of current ideas,” in *Understanding Origins: Contemporary Views on the Origin of Life, Mind and Society, Boston Studies in the Philosophy and History of Science 130, 1st Edn*, eds J.-P. Dupuy and F. J. Varela (Dordrecht: Springer Netherlands), 235–263.
- Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Varela, F., Maturana, H., and Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems* (Cambridge, MA). 5, 187–196. doi: 10.1016/0303-2647(74)90031-8
- Vernon, D. (2010). Enaction as a conceptual framework for developmental cognitive robotics. *Paladyn J. Behav. Rob.* 1, 89–98. doi: 10.2478/s13230-010-0016-y
- Webb, B. (1995). Using robots to model animals: a cricket test. *Rob. Auton. Syst.* 16, 117–134. doi: 10.1016/0921-8890(95)00044-5
- Winograd, T., and Flores, F. (1987). *Understanding Computers and Cognition*. Boston, MA: Addison-Wesley Professional.
- Zarco, M., and Egbert, M. (2019). “Different forms of random motor activity scaffold the formation of different habits in a simulated robot,” in *Artificial Life Conference Proceedings* (New York, NY: MIT Press).
- Ziemke, T., and Lowe, R. (2009). On the role of emotion in embodied cognitive architectures: from organisms to robots. *Cogn. Comput.* 1, 104–117. doi: 10.1007/s12559-009-9012-0

Appendix A

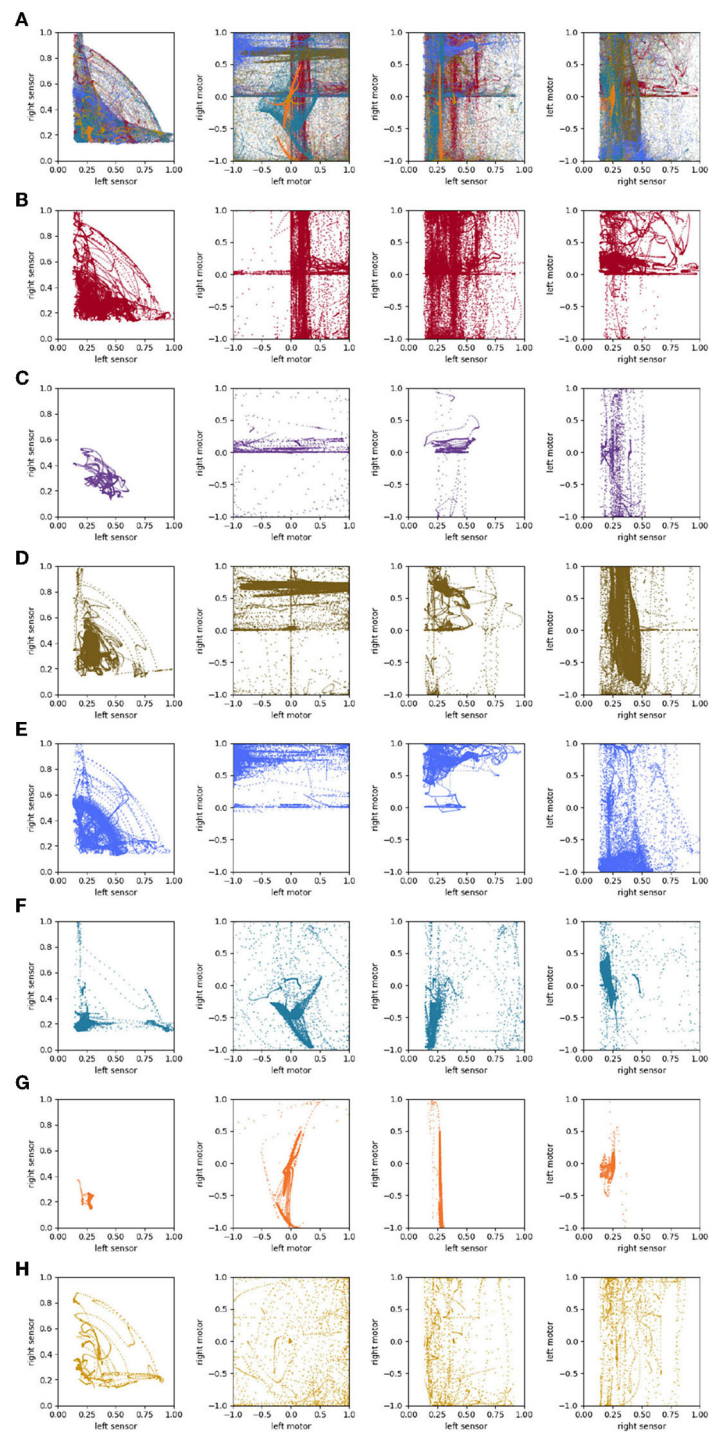


FIGURE A1

Projections of the sensorimotor trajectories of the visual robot. Colors and letters (B–H) are included to facilitate comparison with Figure 3. (A) Shows the superposition of trials (B–H).

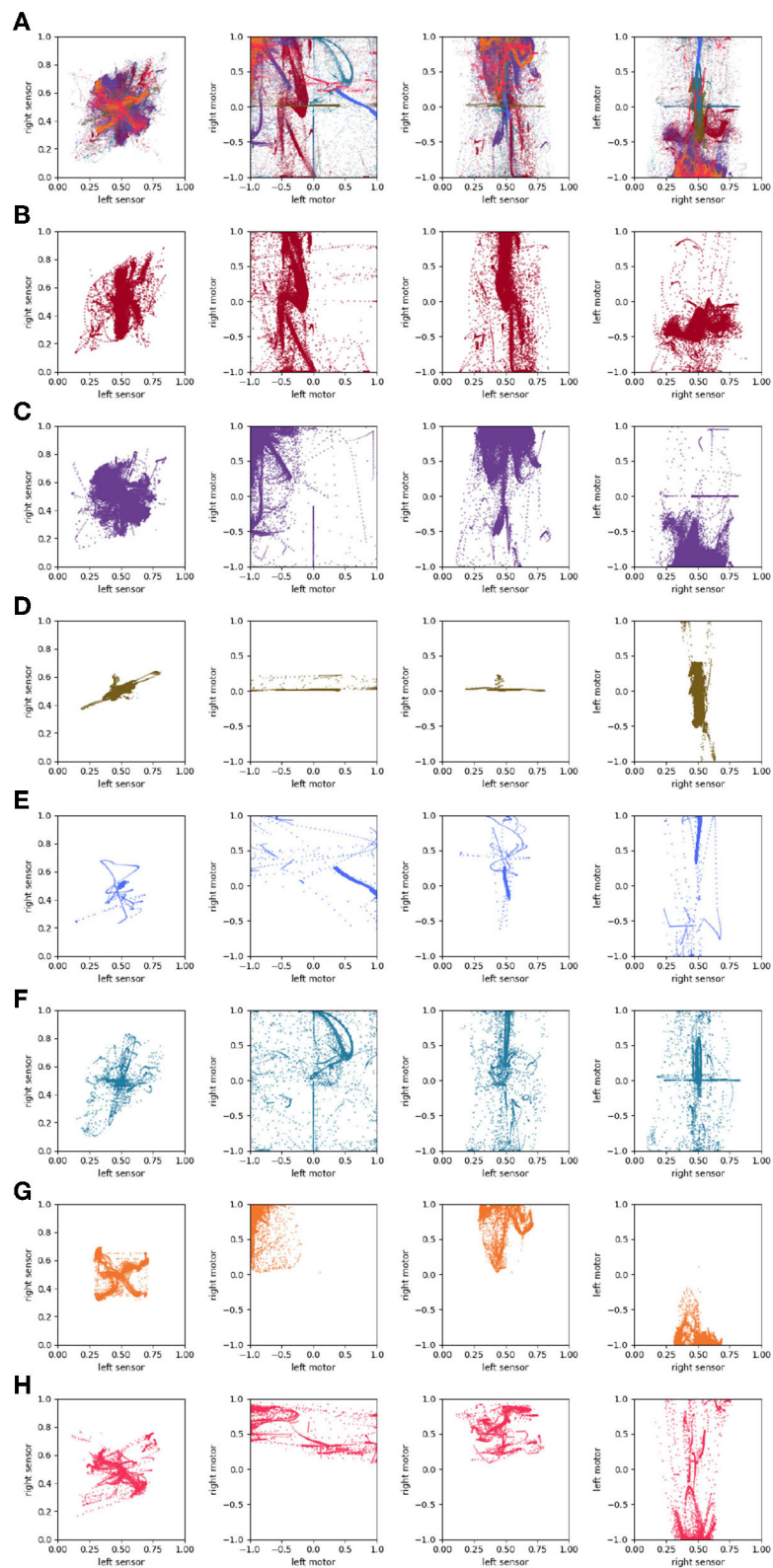


FIGURE A2

Projections of the sensorimotor trajectories of the visual robot. Colors and letters (B–H) are included to facilitate comparison with Figure 7. (A) Shows the superposition of trials (B–H).