Towards modelling social habits: an organismically inspired evolutionary robotics approach

Manuel G. Bedia [1], Manuel Heras-Escribano[2], Diego Cajal [1], Miguel Aguilera[1,2], and Xabier E. Barandiaran [2,3]

[1] ISAAC Lab, Aragón Institute of Engineering Research, University of Zaragoza, Zaragoza, Spain.
[2] IAS-Research Center for Life, Mind and, Society, UPV/EHU, University of the Basque Country, Spain.
[3] Department of Philosophy, Faculty of Labor Relations and Social Work, UPV/EHU, University of the Basque Country, Spain.
mgbedia@unizar.es

Abstract

There has been a revival of the notion of habit in the embodied and situated cognitive sciences. A habit can be understood as 'a self-sustaining pattern of sensorimotor coordination that is formed when the stability of a particular mode of sensorimotor engagement is dynamically coupled with the stability of the mechanisms generating it' (Barandiaran, 2008, p. 281). This view has inspired models of biologically-inspired homeostatic agents capable of establishing their own habits (Di Paolo and Iizuka, 2008). Despite recent achievements in this field, there is little written about how social habits can be established from this modelling perspective. We hypothesize that, when the stability of internal behavioural mechanisms is coupled to the stability of a behaviour and other agents are present during this behaviour, a social interdependence of behaviour takes place: a social habit is established. We provide evidence for our hypothesis with an evolutionary robotics simulation model of homeostatic plasticity in a phototactic behaviour. Agents evolved to couple internal homeostasis to behavioural fitness display social interdependencies in their behaviour. The social habit of these agents was not interrupted when blindness to phototactic stimuli was introduced as long as social perception remained active. This did not happen when internal homeostasis was not coupled to the fitness of the agent. The results allow us to propose a possible conjecture about the character of social habits and to offer a potential theoretical framework to understand how habits develop from neurodynamics to the level of social interaction.

From individual to social habits: philosophy, psychology, neural and social sciences

Habits in the history of philosophy

The concept of habit has been key for making sense of our cognitive abilities. Philosophers, psychologists, and cognitive scientists from the most diverse traditions have used this concept to make sense of behavior and cognition from Ancient Greece to the 20th Century (Sparrow and Hutchinson, 2013). The advent of computationalism and information processing prompted a decline of the notion of habit in the last Century (Barandiaran and Di Paolo, 2014), although the notion has gained momentum recently with the rise of the embodied and situated cognitive sciences (Di Paolo et al., 2017; Egbert and Barandiaran, 2014; Barandiaran, 2017).

Despite its popularity, there has not been a unified notion of habit through the ages. However, we can differentiate two main branches: (1) The associacionist view, that defines habits as automatic responses, and (2) the organicist view, that defines habits as self-organizing structures linked to environmental aspects through which agents establish tendencies with positive or negative balance (Di Paolo et al., 2014). This second branch, which is the one we will focus on, was received by pragmatism, and it was combined with the idea of habit as mentioned in the works of Charles Darwin (James 1890: Chapter 4). On the other side, the first branch became relevant with the scientific revolution, and developed a view that was partially mechanistic and focused on subpersonal processes of automatization. The associacionist conceptualization of habits was inherited by the contemporary neuroscientific approach (Wood and Rünger, 2016).

The organicist view goes back to Aristotle, and it was based on three main ideas: habits are dispositions or tendencies (active capacities of the agent), they are agential (they belong to the agent as a whole) and they are relational (those tendencies of the agent are directed towards an environmental element). The term 'habit' comes from the Greek words 'ethos' and 'hexis' and Aristotle understood hexis as the capacitiy of the agent for being in possession of something in an active way, as 'a kind of activity of the haver and of what he has something like an action or movement' (Aristotle, 2007, 5.20, 1022b1214). This is, a kind of disposition towards through which someone established a tendency, which can be positive or negative (Faucher and Roques, 2018, 2). In this sense, a habit is an active capacity possessed by the agent, who establishes a structure that gives rise to a particular behavior. This is tightly related to the meaning of 'ethos', which means 'custom' or 'habit': that particular behavior is established to be repeated. Habits are also relative to something external, as Aristotle himself defines them (Aristotle, 2004, 8, 11a2032), because these behavioral patterns or tendencies are always related to an object or being of the outer world. This Aristotelian view of habits inaugurates an understanding of habits as self-structured patterns of action that depend on environmental elements, something

that was inherited by traditions as diverse as pragmatism or phenomenology.

The relevance of the organicist view of habits declined with the raise of mechanicism and natural philosophy after the achievements of the scientific revolution. Natural philosophers often described the body in mechanical terms, and this picture has been present thorough history until nowadays. This is the basis of the associacionist view of habits: habits are unconscious, rigid and automatic behaviors that are formed thanks to the association of different stimuli to certain behavioral outputs. The gap between stimuli and response is offered in subpersonal mechanistic terms. This mechanistic view of the body governed the main developments in psychology and neuroscience, explaining the connection between stimuli and response as a bunch of discrete subpersonal steps that are mechanically linked to each other from perception to action. This associacionist view was at the basis of a wide variety of theories, from behaviorism to cognitivism (Reed, 1996) and also inspired the main current views of neuroscience (Bennett and Hacker, 2003, 2008; Kandel et al., 2000)

In the 20th Century, pragmatism and phenomenology inherited the organicist view of habits. First of all, Jamesian psychology included the idea of habit as a cognitive capacity inspired by the work of Darwin. James claimed that there are habits or tendencies that are innate (instincts) or educated (acts of reason) (James, 1890). In any case, these tendencies belong to the agent as a whole and, given the anti-structuralist and anti-elementarist approach of James, habits cannot be reduced to a series of mechanical and discrete subpersonal steps. James claimed that habits are an example of the plasticity of the behavior of organic life that helps us compensating the perturbations of the environment. At the same time, habits help us to automatize our behavior, relieving us from excessive conscious attention (Blanco, 2014). In this sense, James' notion of habit combines the agential, active and dispositional advantages of organicism with the emphasis on automatization and unreflectiveness of the associacionist view.

Habits are understood in Merleau-Ponty's phenomenology as the bunch of ways in which our body establishes a meaningful dialogue with our surroundings by means of our action, allowing us to experience the world as a network of attractions and repulsions (Moya, 2014). Merleau-Ponty claimed that our understanding (this is, our capacity for making sense of the environment) is not based on abstract rationality but on our bodily and unreflective understanding with the world (Kaufer and Chemero, 2015). In this sense, Merleau-Ponty served as an inspiration for those cognitive scientists that endorsed an embodied and situated approach (Moya, 2014; Kaufer and Chemero, 2015; Gallagher, 2017; Heras-Escribano, 2019). Taking this, the embodied and situated approach to cognitive sciences is a direct heir of the organismic view of habits, although passed through the lens

of the pragmatist and phenomenological views offered here; this is, habits are self-structured dispositions or tendencies that allow us to related to the world, but in an unreflective and embodied way. The tendencies of our body towards environmental elements allow us to navigate it and compensate its perturbations in an skillful but totally unreflective way.

From individual to social habits

The term 'ethos', which is at the basis of the idea of habit in the Aristotelian tradition, is also understood in moral terms, such as the ethos of a person or a community (this is why 'ethics' and 'morals' derive from this word). Thus, individual habits also have a moral dimension and, more importantly, a social dimension. In fact, the work of some pragmatist philosophers, such as Dewey, relies on the constitutive aspect of the social environment for establishing both individual and shared habits.

Dewey proposed an ontology of mind based on habits in order to outcompete the subpersonal, mechanistic view of the reflex arc concept in psychology (Dewey, 1896). This reflex arc concept endorsed the idea that perception was passive and separated from action. According to the reflex arc view, the senses received the impingements from the environment and, thanks to a series of discrete and mechanic steps, it was delivered a behavioral output that was automatic and dis-embedded. Dewey claimed that this is not a reliable picture of perception and action: in his view, perception was continuous with action and context-dependent. First of all, perception starts with the exploration of the environment, so stimuli are informed by the previous activity of the organism. Also, the response is based on the particular situation in a given spatiotemporal context that includes the particularities of the natural and social environment that the organism is exploring. In this sense, stimulus and response are two abstractions from the continuous and cyclic process of exploration of the environment. This is why Dewey discards the mechanistic and subpersonal approach of behaviorism and develops an ontology of mind based, among other things, on the organicist idea of habit.

Habits in the Deweyan view are understood as the constitutive dispositional patterns that organize human behavior (Dewey, 1958, pp. 20-6, 40-1), but the key point is that they are socially acquired or established. This is because, according to Dewey, our individual habits are built upon previous collective habits in the sense that there is a priority of the society over the individual (Dewey, 2002, p. 58). This priority is not based on metaphysical assumptions, but on the natural origins of humans: every individual is born in a particular society, which means that individuals develop their own action patterns in accordance with the social ones and thanks to the feedback of the community (Dewey, 2002, p. 59). So, following Dewey, habits are always evaluated or assessed by the individual's social environment, and this assessment comes in the form of a reinforcements and sanc-

342

tions (Dewey, 2002, pp. 16-17). For this reason, Dewey claimed that habits are 'ways of incorporating the environment' (Dewey, 2002, p. 15) and this includes, of course, the social environment. The social constitution of habits is so crucial for Dewey that he claimed that it is practically impossible for humans to form habits that are not socially-established, so our social environment plays a constitutive role in the shaping and establishment of individual habits (Dewey, 2002, p. 16).

The notion of habit has also been vindicated as foundational for social sciences. One such line of development departs from Norbert Elias and Marcel Mauss and finds it most prominent exponent with Pierre Bourdieu (1977). They all share in common the view that habits are not only social constructions shaping individual behaviour, but the most important building block or structuring process of society (in contrast with intellectualist views on how ideology, rational thinking or individual goal-oriented strategic planning shapes society).

Bourdieu defines habitus as:

> systems of durable, transposable *dispositions*, structured structures predisposed to function as structuring structures, that is, as principles of the generation and structuring of practices and representations which can be objectively 'regulated' and 'regular' without in any way being the product of obedience to rules, objectively adapted to their goals without presupposing a conscious aiming at ends or an express mastery of the operations necessary to attain them and, being all this, collectively orchestrated without being the product of the orchestrating action of a conductor. (Bourdieu, 1977, p. 72)

We aim to enrich recent contributions in which habits have been modelled (Di Paolo and Iizuka, 2008) by introducing a social dimension of habits inspired by the main insights of social habits as defined by Dewey and social theorists. These social habits, then, satisfy both the organismically-inspired homeostatic model of habit-formation (because we rely on processes of homeostasis as a key for the establishment of social habits, offering a naturalistic continuum between individual and social habits) as well as the Deweyan demands of understanding habits as 'ways of incorporating the [social] environment' (Dewey, 2002, p. 15). At the same time, Bourdieu's notion of structuring structures resembles a classical theme of Artificial Life, that of self-organizing structures, where a process produces and re-produces its conditions for self-perpetuation or self-structuring and it affords the opportunity to model social life using artificial life modelling techniques. Artificial Life should thus be ready to face both challenges: the social dimension of habit and the habitual dimension of society.

In the next subsection we will explain how a recent modelling version of individual habits have been proposed in the embodied and situated cognitive sciences. This version includes homeostasis as a key aspect for offering a naturalization of habits as understood within the organicist tradition.

Modelling habits for organismically-inspired homeostatic agents

An organismically inspired approach to habit formation in evolutionary robotics and artificial life is not new. Ezequiel Di Paolo pioneered the modelling of homeostatic behavior generating mechanisms to simulate re-adaptation to visual inversion(Di Paolo, 2000), a paradigmatic experimental setup aimed and discovering the deep entanglement of sensorimotor habits and processes of re-habituation (Kohler, 1963). Variations of these homeostatic mechanisms have been many (Aguilera et al., 2016, 2015; Iizuka and Di Paolo, 2008; Di Paolo, 2003; Williams and Noble, 2007; Iizuka and Di Paolo, 2007), but perhaps the most relevant for our purpose here is that of Barandiaran and Di Paolo (2010), where the authors develop a model of habit development, maintenance and switching for an operant conditioning task and the spontaneous emergence of new associative habits not present during artificial evolution. The internal mechanism capable of supporting this behaviour was homeostatic plasticity in a continuous time recurrent neural network and a sensorimotor embodied task. A behavioural sensorimotor pattern is linked to an emergent region of stable synaptic plasticity. Being plasticity activity dependent, variations on sensorimotor history are capable to generate, reshape those regions or to switch between them. Long terms disruption of sensorimotor correlations can also destroy the underlying neurosynaptic structures supporting them. However, the high dimensionality of the system makes it hard to analyse. In an attempt to simplify the model Egbert and Barandiaran (2014) chose a mesoscopic level of modelling, avoiding neural mechanisms, and proposing instead an iterant deformable sensorimotor medium (IDSM) a mode of plastic sensorimotor mapping that became structured through repetition. The spontaneous emergence of an ecology of habits was displayed on that model (see Egbert, 2018, for further developments). Furthermore, other works have used minimal models to analyze the relation between agency and agency and social interaction Di Paolo et al. (2008), although they focus on interactive dynamics in a social context without considering the capacity of agents of developing plastic habits.

From an organismic perspective a habit can be defined as 'a self-sustaining pattern of sensorimotor coordination that is formed when the stability of a particular mode of sensorimotor engagement is dynamically coupled with the stability of the mechanisms generating it' (Barandiaran, 2008, p. 281). As we can see, this model offers a naturalization of habits that satisfies the organicist view shown here, as self-organizing structures that govern the action patterns of an agent. In particular, it satisfies the two main requirements: they are active and relational, because the modelled agent

Figure 1: Types of sociality of a habit

displays a behavior based on dispositions and tendencies actively and in relation to the source of stimulation in the environment. The introduction of a social dimension to this definition of habits has to do with the social or interactive dependence of sensorimotor coordinations, their generation and stability.

What is a social habit? We can establish the sociality of a habit by degrees along three axes or *dimensions* of sociality: the behavioural or coordinative dimension (whether the enactment itself in action demands coordination with other agents), the structural or stabilizing dimension (whether the stability of the action demands social interaction) and the generative dimension (whether the habit can be acquired without social interaction). Building upon De Jaegher et al. (2010) we can distinguish three *degrees* of sociality or interactivity: contextual, enabling and constitutive. Contextual sociality means that other agents' behavior S simply affect or produce variations on habit H. Enabling sociality means that habit H cannot happen without S and constitutive sociality means that S is inherently part of H. We want to add a further distinction within the enabling degree, which can be *strong* when S has to happen, *ceteris-paribus*, for H to occur, and *weak* when S has to happen for H to occur only in cases in which conditions change (e.g. an impairment has taken place). So, for example, dressing up might be something that is done individually, without coordination with others. But it is certainly generatively social (wild infants have not been seen to dress up) and is probably a structurally social habit (meaning that out of a social context, the care and disposition to dress up fades away). However dressing up *fashionably* is constitutively social at the coordinative level. Similarly if your partner has often a say on how you should dress up, whether you follow the advice or directly and systematically oppose it, the habit will be contextually social at the coordinative dimension.

We hypothesize that, when the stability of internal behavioural mechanisms is coupled to the stability of a behaviour, and other agents are present during this behaviour, a social interdependence of behaviour takes place: a social habit is established, even if the task is not coordinatively social. It follows that no specific 'social mechanism' is required for a social habit to emerge. We provide evidence for our hypothesis with an evolutionary robotics sim-

ulation model of homeostatic plasticity in a phototactic behaviour. Agents evolved to couple internal homeostasis to behavioural fitness display social interdependencies in their behaviour so that blindness to the very object of taxis does not disrupt the habit, provided that social perception is still active. This constitutes a mode of *weak social enabling* for habit enactment (see Figure 1).

In order to explore this hypothesis we defend a phototactic task is a sufficiently complex task so as to discover some relevant sociality phenomena. Pierre Bourdieu often conceives social phenomena as occurring in a social field where agents (individuals) move trying to maximize certain values (e.g. social status). Interestingly for Bourdieu, both the perception of the field is the result of habit and, at the same time, moving through the field creates the field itself. In this paper we are not going to develop such a deep conception of social *habitus*, but it suffices to note that moving toward a light source can potentially mimick the structure and phenomenology of more complex forms of social behaviour.

Model

In Ashby's seminal work (Ashby, 1952), the author states that the adaptation of organisms can be understood as the maintenance of the internal stability through the homeostatic regulation of some essential variables that ensure survival. By doing so, Ashby proposes a conceptual framework for the development of artificial agents in which, rather than assuming that organisms have behaviors that seek intentional goals, they are able to generate self-induced adaptive behaviors exclusively through the maintenance of their internal stability. Based on Ashby's ideas, Di Paolo (2000) applied notions of homeostatic regulation to the synaptic activity of agents involved in phototaxis, demonstrating that they could adapt to sensory inversion despite not having been developed specifically for this task. Di Paolo explains how that adaptation arises in this model: although agents do not evolve to adapt to inversion of the visual field, it is suggested that evolution has created a link between structural stability and desired behavior. This link appears because the process shapes a stable attractor when a certain pattern of sensorimotor activity is present.

We use the same kind of minimal agent guided by homeostasis to search stable configurations where the plasticity effects generate the behaviour and behaviour affects plasticity, leading to the creation of invariants and habits in a social scenario. This allows us to investigate how ultrastability can lead to adaptation in a social dimension.

Model: We simulated a pair of minimal agents evolved to exhibit phototaxis on a series of 1000 light sources. Agents are evaluated in a two-dimensional environment with only a single light source at a time. During each trial, agents are encouraged to remain close to the light sources. Light sources are placed randomly at a distance between 10 and 25 times the agent's radius from the agent position, for a random pe-
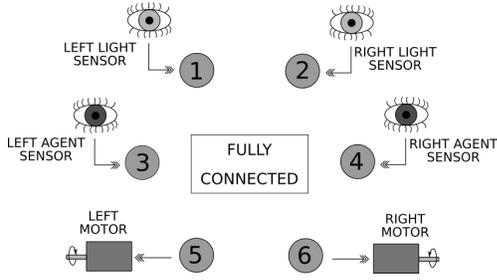
Figure 2: Agent's controller

riod of time T, chosen from the interval $[2000, 4000]$ time steps and with a random intensity.

Agent: Each agent is modelled as a simple circular agent with two light sensors and two motors. If we name V_l and V_r as the motors' velocities and r as the agent's radius (for the experiments, the body's radius is set at 4), we calculate the translational movement as $(V_r + V_l)/2$ and the rotational movement by calculating the angular speed as $(V_r - V_l)/2r$. Collisions are not considered.

Sensors: Each agent has two light sensors, separated by 120°, symmetrically placed at the front of each agent and each sensor has a viewing cone of 80°. At each time step, the angle from the agent to the light source is first calculated and then adjusted by the agent's rotation. This angle is then used to determine if the agent's sensors are active. Light intensity on each sensor can be calculated as inversely proportional to the square of the distance to the light source. In addition, agents have two agent sensors, placed on the same positions and with the same cone of vision that light sensors. Each agent can be seen as a mobile source with constant intensity, so another agent can see it as it see lights. Intensity values on sensors are calculated in the same way as light case.

Controller: A 6-neuron fully connected Continuous Time Recurrent Neural Network is used as the agent's controller. Each neuron is a CTRNN node described by

$$\dot{y_i} = \frac{1}{\tau_i} \left(-y_i + \sum_{j=1}^{N} w_{ji} \sigma \left(y_j + \theta_j \right) + I_i \right) \quad (1)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

where $i = 1, 2, ..., 6$. Each neuron's state is controlled by y_i, representing the cell's potential, τ_i is the decay constant, b_i the bias, z_j the firing rate of the j_{th} neuron, w_{ij} the strength of synaptic connection from node i to node j, and I_i the input from sensors. Inputs and outputs are connected as depicted in Figure 2.

Plasticity: Controllers are able to regulate synaptic activity. Plastic changes in the recurrent network occur locally on each connection governed by both the synaptic activity and

a plasticity rule encoded genetically. The plasticity rule is given by one of these hebbian learning rules

$$\Delta w_{ij} = \delta_i n_{ij} p_j z_i z_j \quad (3)$$

$$\Delta w_{ij} = \delta_i n_{ij} p_j (z_i - z_{ij}^0) z_j \quad (4)$$

$$\Delta w_{ij} = \delta_i n_{ij} p_j z_i (z_j - z_{ij}^0) \quad (5)$$

$$\Delta w_{ij} = 0 \quad (6)$$

where Δw_{ij} is the change per unit of time to a synaptic weight w_{ij}, z_i and z_j are the firing rates of the presynaptic and postsynaptic neurons, n_{ij} is the learning rate factor, and p_j is the degree of local plasticity. δ_i models a damping factor, and z_{ij}^0 a weakening of a synapse.

Genetic Algorithm: Agent's genotypes are divided into a real component and an integer component. The real component has 87 genes: 3 for motor and sensor gains, one per each type; 6 bias and 6 decay constants, one per neuron; 36 weights and 36 learning rates, one per synapse. Finally, 36 integer values are used to select a hebbian rule from (3) to (6) for each synapse. A genetic algorithm is used to test agents against a sequence of five lights. Two agents with the same genotype are selected at random to test their aggregate phototaxis in every step. Fitness scores of each pair of agents are assessed as the averaged individual fitness. Then, selected agents, using crossover and mutation, form a new offspring, which replaces the lowest scoring pairs. A mutation probability of 0.5 is used for all real components and a probability of 0.1 is used for integer components. The GA is repeated until no perceivable fitness increase is observed.

Homeostatic fitness function: The agent's behaviour is assessed using a fitness function with three terms where the first, F_d, is a measure of how near to the source an agent is (measuring the reduction in final and initial starting positions for the agent), F_p is a measure of time spent near the source (proportion of time over the evaluation period that the agent is within 4 radius units of the light source) and F_h is a measure of the homeostatic behaviour of an agent's neurons (the time-averaged proportion of neurons that act homeostatically, without inducing plasticity in the controller). The results of all three sub-fitness functions are ranged between $[0, 1]$. $F_H = 0.2F_d + 0.64F_p + 0.16F_h$. Each component has got a weight to adjust the contribution of each term of the fitness function in order to have the sum equal to 1.

Non-homeostatic fitness function: In order to evaluate how homeostasis affects the behaviour of agents, two different types of fitness are used: one taking into account the term for homeostatic stability (F_H), and the other $F = 0.2F_d + 0.8F_p$ not taking it into account. Evolving populations of pairs of robots with these two possible fitness functions, we obtain two types of agents: homeostatic and non-homeostatic. From each population, we select the best pair of agents with best performance for the subsequent analysis.
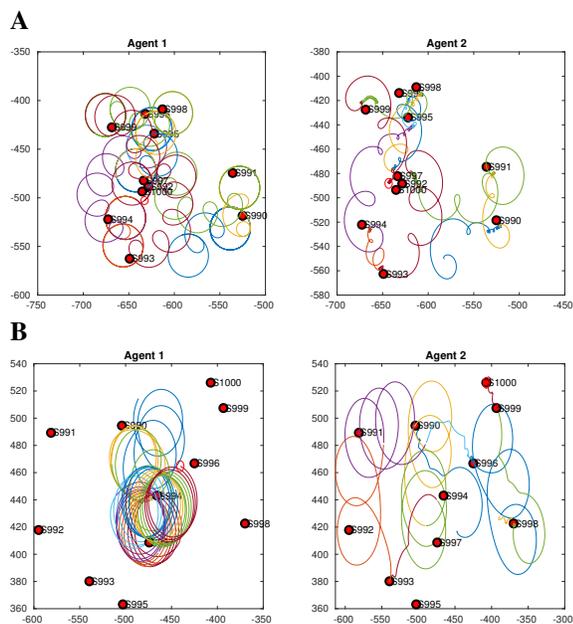
Figure 3: Behaviour of two agents performing the task in homeostatic (A) and non-homeostatic (B) conditions. Red dots are the light sources, numbered as they appears. Agent 1 is blinded, so it can perceive the other agent but not the light. We observe that, in the homeostatic case, both agents are able to behave in coordination and reaching the lights, suggesting that a collective habit is displayed.

Figure 4: Fitness values of Agent 1 (gray) and Agent 2 (white) when homeostatic plasticity is **(A)** included **(B)** not included in evolution. Fitness is compared between normal conditions and a situation in which Agent 1 is blind to the light. Fitness of Agent 1 only maintained to some degree in the case of homeostatic stability.

Results

We want to explore the relation between individual capacities and social interaction in the maintenance of habits. For doing so, we artificially 'blind' one of the agents and compare the case of the robots with synaptic plasticity with agents not evolved for homeostatic stability. We only blind light sensors. Thus, blinded agents can keep the sight of other agents, which are the only clue to reach light sources. We then observe how the behaviour of the agents is reconfigured after synaptic plasticity stabilizes the activity of agents.

In Figure 3, we show the behaviour of the two agents in the case when Agent 1 is blinded, for the case with homeostatic stability (A) and without (B). In the first case, we observe that, even when Agent 1 is not able to perceive the light anymore, the agents are able to collectively achieve the task after homeostatic reconfiguration. In the case without homeostatic stability, this is no longer the case, and the behaviour of the blind agent is disrupted. This suggests that synaptic plasticity maintains a collective pattern of behaviour.

In order to test this change in a more systematic way, we simulate the behaviour of a (normal or blinded) Agent 1 and a (normal) Agent 2 reaching a series of 20 lights. We repeat each experiment 40 times and measure the fitness F (without considering homeostatic stabilization terms F_h). In Figure 4
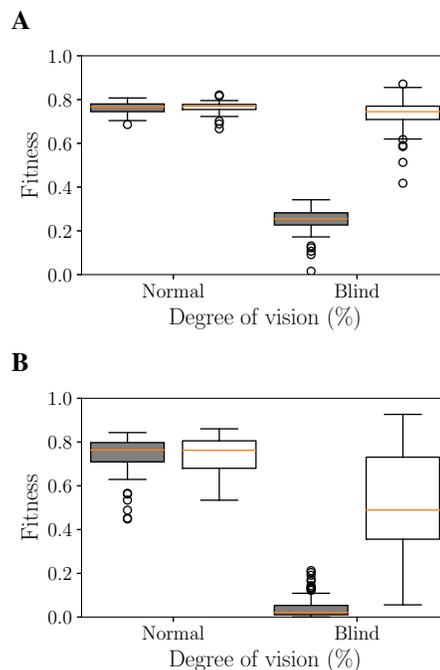
we observe how fitness F change when Agent 1 is blinded to the light. In the case with homeostatic stabilization, the fitness of agents in the normal case is of 0.77, and when Agent 1 is blinded its fitness is reduced to 0.26, while the fitness of Agent 2 is reduced to 0.74. This means that, while Agent 2 is practically unaffected, the fitness of Agent 1 is reduced to one third, although it is still able to reach a number of lights thanks to its coordination to Agent 2. Thus, even when Agent 1 has lost its individual abilities for phototaxis, it is able to maintain its phototactic habit to some extent thanks to social interaction.

In the case without homeostatic stability, the situation is different. Now the mean fitness is reduced from around 0.76 to 0.02 for the blinded agent and 0.49 for Agent 2. In this case, loosing its individual vision means that Agent 1 is no longer able to maintain its phototactic habit. Furthermore, the phototactic habit of Agent 2 is disrupted when Agent 1 can no longer reach the light, and its performance is reduced. In further tests, we tested the effect of 'blinding' the agent to social interaction instead of the light, without practically no effect in the fitness of the agents, so results are not reported.

We tried to increase the number of agents working together on a step-wise manner. Good performances are achieved in populations of 3 ($F = 0.71$) and 4 ($F = 0.67$)

agents with only a blind agent per group. However, the GA did not provide any successful case of more than four agents. An element to consider to explain this is that agents can't see the position of each other agent but a sum of their contributions to an agent sensor.

Discussion, conclusions and future work

In this article, we have explored the social dimension of habits understood under the lens of homeostatic neurodynamic models of habit generation. These models have been mostly focused in the emergence of individual sensorimotor habits and we have shown how a richer picture should incorporate the social dimension of habits.

But... what is the nature of social habits? Is it a merely aggregation of individual habits? Or does the social dimension play a constitutive role for generating emergent collective habits? How can we attempt a gradual path towards social habits? A social account of habits should clarify these aspects and offer a sound theoretical framework of how habits are developed from neural and sensorimotor dynamics to the level of social interaction and back.

We have provided a theoretical framework that classifies the types and degrees of sociality of habits, and we have illustrated a weak enabling sociality of the coordinative dimension of habits with a model of evolutionary robotics: a minimal model of agents controlled by a dynamical neural network incorporating mechanisms of homeostatic synaptic plasticity evolved to solve a task in a social scenario. A pair of agents have to perform a phototactic task. Each agent has sensors that perceive the presence of light sources, and another set of sensors that perceive the presence of the other agent. These agents are compared with a similar population adapted to solve the task without a pressure for stabilizing its synaptic plasticity, therefore reducing drastically the capacity for self-regulation of emerging patterns or habits.

In order to investigate the degree of sociality of the phototactic habits in our robotic model we perform an experiment over the resulting agents in which we artificially 'blind' one of the agents of the group. Thus, the capacity to express an individual habit is removed, yet its capacity for social interaction remains present. In this case, we can see that agents with homeostatic stabilization are able to maintain to some extent the performance of the task, therefore illustrating that the social dimension has a role in the maintenance of a phototactic habit. In contrast, agents that are not selected for homeostatic stabilization lose their capacity for phototactic behaviour entirely when blinded, suggesting that in this case the phototactic habit cannot be maintained by social interaction when direct photosensitivity is precluded. Still, performance of non-blinded agents is reduced when an agent is blinded, suggesting that the phototactic habit also has some social component. This suggests that the relation between habits and its individual and social components can be intricate, and further studies should clarify these relations. The

modeling results presented here are tentative and just seek to illustrate the relation between some ideas presented in this article. Further work could explore in more depth different possibilities to study different ways in which a habits can be socially constituted at different levels and degrees.

As we have seen, the dependence between the maintenance of habits, the presence of individual abilities and social coordination is complex. A possibility to unveil the complexities of these relations could be to introduce different scenarios in which noise or perturbations selectively disturb the individual or the social behaviour of the agents. This should force the agents to rely more on their individual capacities or in one another, depending on the case. This could allow the exploration of different kinds of the habits in our classification (Figure 1).

A natural next step on the evolution of social habits is to apply this framework to a constitutively social coordination task. We have focused on a simple phototactic behaviour that can be carried out individually. On the other hand, it is typical of collective behavior to include a necessary or constitutive coordination between agents: e.g. dancing, moving as a whole, etc. Long term studies of the structural stability of the habits out of the social dimension would also be interesting as future work.

Another research line that could be explored is the social constitution of the behavioural field. For instance, habits could be formed not (or not only) by exploiting the synaptic plasticity of an agent's neural controller, but by plastically modifying its own environment. For example, the activation of lights could be the result of the behaviour of the agents (e.g. light intensity being reinforced or consumed by the proximity of an agent). This will help approaching the concept of social field by Bourdieu.

References

Aguilera, M., Barandiaran, X. E., Bedia, M. G., and Seron, F. (2015). Self-Organized Criticality, Plasticity and Sensorimotor Coupling. Explorations with a Neurorobotic Model in a Behavioural Preference Task. *PLoS ONE*, 10(2):e0117465.

Aguilera, M., Bedia, M. G., and Barandiaran, X. E. (2016). Extended Neural Metastability in an Embodied Model of Sensorimotor Coupling. *Frontiers in Systems Neuroscience*, 10.

[1]The Foundation accepts no responsibility for the opinions, statements and contents included in the project and/or the results thereof, which are entirely the responsibility of the authors

Aristotle (2004). *The Categories*. Kessinger Publishing.

Aristotle (2007). *Metaphysics*. Mineola, NY: Dover.

Ashby, W. R. (1952). *Design for a Brain*. J. Wiley, 2 edition.

Barandiaran, X. E. (2008). *Mental Life: a naturalized approach to the autonomy of cognitive agents*. PhD Thesis, University of the Basque Country (UPV-EHU), Donostia - San Sebastin, Gipuzkoa, Spain. 17th June 2008.

Barandiaran, X. E. (2017). Autonomy and Enactivism: Towards a Theory of Sensorimotor Autonomous Agency. *Topoi*, 36(3):409–430.

Barandiaran, X. E. and Di Paolo, E. A. (2010). Homeostatic Plasticity in Robots: from development to operant conditioning to habit formation. In *Proceedings of CogSys2010*, Zurich.

Barandiaran, X. E. and Di Paolo, E. A. (2014). A genealogical map of the concept of habit. *Frontiers in Human Neuroscience*, 8:522.

Bennett, M. R. and Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. Wiley-Blackwell.

Bennett, M. R. and Hacker, P. M. S. (2008). *History of cognitive neuroscience*. John Wiley & Sons.

Blanco, C. A. (2014). The principal sources of William James' idea of habit. *Frontiers in Human Neuroscience*, 8.

Bourdieu, P. (1977). *Outline of a Theory of Practice*, volume 16. Cambridge university press.

De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 14(10):441–447.

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological review*, 3(4):357.

Dewey, J. (1958). *Experience and nature*, volume 471. Courier Corporation.

Dewey, J. (2002). *Human nature and conduct*. Courier Corporation.

Di Paolo, E., Buhrmann, T., and Barandiaran, X. E. (2017). *Sensorimotor Life: An enactive proposal*. Oxford University Press.

Di Paolo, E. A. (2000). Homeostatic adaptation to inversion of the visual field and other sensorimotor disruptions. In *From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior*, Cambridge, MA. MIT Press.

Di Paolo, E. A. (2003). Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop. In Murase, K. and Asakura, editors, *Dynamical systems approaches to embodiment and sociality*, pages 19–42. Advanced Knowledge International, Adelaide.

Di Paolo, E. A., Barandiaran, X. E., Beaton, M., and Buhrmann, T. (2014). Learning to perceive in the sensorimotor approach: Piagets theory of equilibration interpreted dynamically. *Frontiers in Human Neuroscience*, 8.

Di Paolo, E. A. and Iizuka, H. (2008). How (not) to model autonomous behaviour. *Biosystems*, 91(2):409–423.

Di Paolo, E. A., Rohde, M., and Iizuka, H. (2008). Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing. *New Ideas in Psychology*, 26(2):278–294.

Egbert, M. (2018). Investigations of an Adaptive and Autonomous Sensorimotor Individual. *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)*, pages 343–350.

Egbert, M. D. and Barandiaran, X. E. (2014). Modeling habits as self-sustaining patterns of sensorimotor behavior. *Frontiers in Human Neuroscience*, 8:590.

Faucher, N. and Roques, M., editors (2018). *The Ontology, Psychology and Axiology of Habits (Habitus) in Medieval Philosophy*. Historical-Analytical Studies on Nature, Mind and Action. Springer International Publishing.

Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind*. Oxford University Press.

Heras-Escribano, M. (2019). *The Philosophy of Affordances*. Palgrave Macmillan, New York, NY, 1st ed. 2019 edition edition.

Iizuka, H. and Di Paolo, E. A. (2007). Toward Spinozist Robotics: Exploring the Minimal Dynamics of Behavioral Preference. *Adaptive Behavior*, 15(4):359–376.

Iizuka, H. and Di Paolo, E. A. (2008). Extended Homeostatic Adaptation: Improving the Link between Internal and Behavioural Stability. In *From Animals to Animats 10. 10th International Conference on Simulation of Adaptive Behavior*, pages 1–11.

James, W. (1890). *The Principles of Psychology*. Henry Holt and Company.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Department of Biochemistry and Molecular Biophysics Thomas Jessell, Siegelbaum, S., and Hudspeth, A. (2000). *Principles of neural science*, volume 4. McGraw-hill New York.

Kaufer, S. and Chemero, A. (2015). *Phenomenology: an introduction*. John Wiley & Sons.

Kohler, I. (1963). The formation and transformation of the perceptual world. *Psychological issues*.

Moya, P. (2014). Habit and embodiment in Merleau-Ponty. *Frontiers in Human Neuroscience*, 8.

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. Oxford University Press.

Sparrow, T. and Hutchinson, A. (2013). *A History of Habit: From Aristotle to Bourdieu*. Lexington Books.

Williams, H. and Noble, J. (2007). Homeostatic plasticity improves signal propagation in continuous-time recurrent neural networks. *Biosystems*, 87(23):252–259.

Wood, W. and Rünger, D. (2016). Psychology of Habit. *Annual Review of Psychology*, 67:289–314.