

Enabling robot autonomy through biomimetic self-regulatory dynamics

1st Oscar Guerrero Rosado

Donders Institute for Brain, Cognition and Behaviour
Radboud University
Nijmegen, Netherlands
oscar.guerrerorosado@donders.ru.nl

2nd Ismael T. Freire

Donders Institute for Brain, Cognition and Behaviour
Radboud University
Nijmegen, Netherlands

3rd Adrian F. Amil

Donders Institute for Brain, Cognition and Behaviour
Radboud University
Nijmegen, Netherlands

4th Paul F. M. J. Verschure

Donders Institute for Brain, Cognition and Behaviour
Radboud University
Nijmegen, Netherlands
paul.verschure@donders.ru.nl

Abstract—From weaving spiders to hibernating mammals and migratory birds, nature presents numerous examples of organisms exhibiting extraordinary autonomous behaviors that ensure their self-maintenance. However, physiological needs often interact and compete with each other, which demands super-autonomous organisms to orchestrate them as a complex set of internal needs rather than as isolated subsystems. This paper presents a synthetic agent equipped with a brain-based neural mass model replicating fundamental self-regulatory behaviors observed in desert lizards. Our results show this agent not only autonomously regulates its internal temperature by navigating to areas with optimal environmental conditions but also harmonizes this process with other internal needs, such as energy, hydration, security, and mating. Using game theory metrics, we observe that such a biomimetic agent outperforms an interoceptive-agnostic agent in efficiency, fairness, and stability. Together, our results suggest that grounding robot behavior in biological processes of self-regulation provides an excellent approach for addressing trade-off situations in the physical deployments of autonomous robots.

Index Terms—Autonomy, Biomimetics, Embodied Cognition, Self-maintenance, Self-Regulation, Allostasis.

I. INTRODUCTION

When attempting to build super autonomous robots capable of making decisions without any human intervention, we open the door to the question of what grounds such decisions. In this paper, we argue that drawing inspiration from biological systems not only helps us better understand the cognitive mechanisms behind the decisions of these systems but also enables us to foster robot autonomy in challenging environments.

Indeed, nature provides numerous examples of organisms exhibiting extraordinary autonomous behaviors. Ranging from weaving spiders to hibernating mammals and migratory birds, these biologically autonomous individuals ground their behavioral plans in a specific purpose: self-maintenance [1]. Thus, in order to survive and thrive, autopoietic systems produce

autonomous behavior to self-regulate a set of physiological needs critical for their subsistence [2], [3].

Autonomous behavior and its fundamental self-regulatory mechanisms can be found across different organizational scales, ranging from single cells to networks, organs, whole organisms, and even groups of organisms [4]. One of these self-regulatory mechanisms was first defined by Walter Cannon [8]. He coined the process of maintaining an internal variable within a steady desired range “homeostasis” and described it as a negative feedback loop. In this loop, any deviation from the desired range would be sensed as a homeostatic error, triggering a proportional error-correcting response delivered by the system.

The concept of homeostasis has not only revolutionized biology but has also influenced psychology. One major theory that integrates such self-regulatory mechanisms in the explanation of goal-directed autonomous behavior is the Drive Reduction theory [9]. According to this theory, organisms pursue goals that fulfill internal needs, which are understood in homeostatic terms as deviations from a desired state. The correction of these needs results rewarding to the individuals, encouraging them to reproduce the same behavior in future similar situations.

However, homeostatic mechanisms fall short when trying to explain self-regulation in its entirety. First, it is challenging to identify steady states in physiological measures. Variables such as heart rate [5], body temperature [6], or body mass index [7] exhibit variations over minutes, hours, or days. Second, for autopoietic systems to self-maintain and reproduce, the balance of isolated homeostatic systems is insufficient. Organisms embody a rather complex set of internal needs that often interact and compete with each other, which makes their optimal orchestration critical for survival. Indeed, learning environmental contingencies and developing anticipatory behavior is key to avoiding catastrophic deviations from homeostasis. In an attempt to address these issues, Sterling and Eyer [10] proposed the term “allostasis” as the

This study was supported by Counterfactual Assessment and Valuation for Awareness Architecture—CAVAA (European Commission, EIC 101071178).

adaptive stability of the self in response to the demands and opportunities that the environment offers. Thus, allostasis considers stability as a dynamic process that concerns the self rather than its components and occurs in consonance with environmental events.

Previous works have implemented allostatic mechanisms as an efficient approach to designing autonomous robots [11]–[15]. In this paper, we update such an approach by biologically grounding the design of a novel allostatic model on hypothalamic dynamics. This evolutionary ancient brain structure encompasses a set of nuclei responsive to physiological changes and nominated as the neural substrates of homeostatic drives (i.e. the motivation to correct a homeostatic error) [19]. Nonetheless, the competing dynamic between multiple drives imposed by their mutual inhibitory projections [16]–[18] has not been explored as a source of allostatic orchestration.

To test the autonomy of a synthetic agent endowed with this hypothalamic model, we use the Nabim desert lizard [20] as a benchmark. Like any other ectothermic lizard, this species needs to be exposed to an external source of heat (usually sunlight) to regulate its body temperature. However, the transient high temperatures of its natural environment force it to seek colder areas from time to time, which it accomplishes by diving beneath the sand. While this self-regulatory behavior allows the individual to achieve thermoregulation, it also hampers its ability to encounter other relevant resources, thus preventing it from fulfilling other physiological needs such as energy, hydration, security, and mating.

II. MATERIALS

To replicate the desert lizard agent and its natural environment, we designed a virtual experimental setup using Webots [21]. The setup consisted of three e-puck mobile robots representing the allostatic agent, a mate, and a predator (Fig. 1, Video). In addition, the simulated environment contained two objects representing available sources of energy and hydration. Importantly, the location of each resource-providing object (including the optimal temperature area) was simultaneously modeled as gradients. The motivation to implement such gradients is dual: On the one side, nature offers physical gradients, such as those based on chemical concentration [22] or environmental temperature [19], to organisms (from micro-organisms to highly complex cognitive species) so they can use them to guide their behavior. On the other side, it is known that cognitive systems are capable of learning and embodying such gradients as cognitive representations when familiarizing with a new environment and its reward locations [23], [24]. Moreover, gradients are popularly modeled in reinforcement learning as policies, where states close to rewards augment their state value function [25]. And the implementation of such gradients has been shown effective to guide navigation in similar scenarios [6], [12], [26].

III. METHODS

A set of internal needs are incorporated into the agent as homeostatic systems. That is, every internal need (i.e. home-

ostatic error) is defined as the Euclidean distance between the current internal state and the setpoint (i.e. the desired steady state) of such homeostatic system. In order to force competition and maximization of the internal states, all desired values were set at 1 (max. internal state).

Internal states are slowly but continually reduced unless the e-puck agent is in the vicinity (peak area of the gradient) of a resource. In that case, the internal state increases. These internal state dynamics hold for all needs with the exception of security, which obtains the actual value of its internal state from the distance between the robot and its predator. Importantly, temperature, security, and mating gradients change over time. Specifically, the temperature gradient follows a process of shrinking and widening the optimal temperature area to simulate desert temperature across day-night cycles. Five of these cycles were reproduced. In contrast, the security and mating gradient were also continually updated but in this case, considering the location of the respective e-puck (either the predator or the mate).

The agent was endowed with a neural mass model that, using the homeostatic errors as inputs, is capable of efficiently orchestrating the robot’s internal needs. Derived from modified Wilson-Cowan equations [27], the neural mass model is grounded in hypothalamic dynamics [16]–[18] by incorporating a relationship of competition between drives. Mutual and shared feedback inhibition regulates such a competitive relationship as follows:

$$\tau_i \frac{dD_i(t)}{dt} = -D_i + f(w_+ D_i + I - Q w_- \sum_{j \in N, j \neq i} D_j - (1 - Q) w_- f(\sum_{j \in N} D_j)) + \sigma \xi(t) \quad (1)$$

where $f(x)$ is the logistic $f - I$ function,

$$f(x) = \frac{F_{max}}{1 + e^{\frac{-(x-\theta)}{k}}} \quad (2)$$

In these equations, τ_i is the time constant determining the timescale of population dynamics. D is the set of drive magnitudes and D_i is the specific drive being updated. I is the input each excitatory population receives, the homeostatic error in our case. w_+ , w_- , and w_- are the weights for recurrent connections within the excitatory population, mutual inhibition, and feedback inhibition, respectively. Q represents the mutual/feedback inhibition ratio, a variable that allows inducing a controlled level of competition. σ and ξ are the variance and magnitude of Gaussian noise provided to the excitatory populations. Finally, F_{max} , k , and θ are the maximum firing rate, gain, and threshold parameters of the $f - I$ logistic curve respectively (equation 2).

The model output results in attractor dynamics that ensure the consideration of the stronger drive while neglecting irrelevant ones until a change in the internal state occurs. Such dominating drive informs the agent about the gradient that must be followed in order to satisfy the internal need.

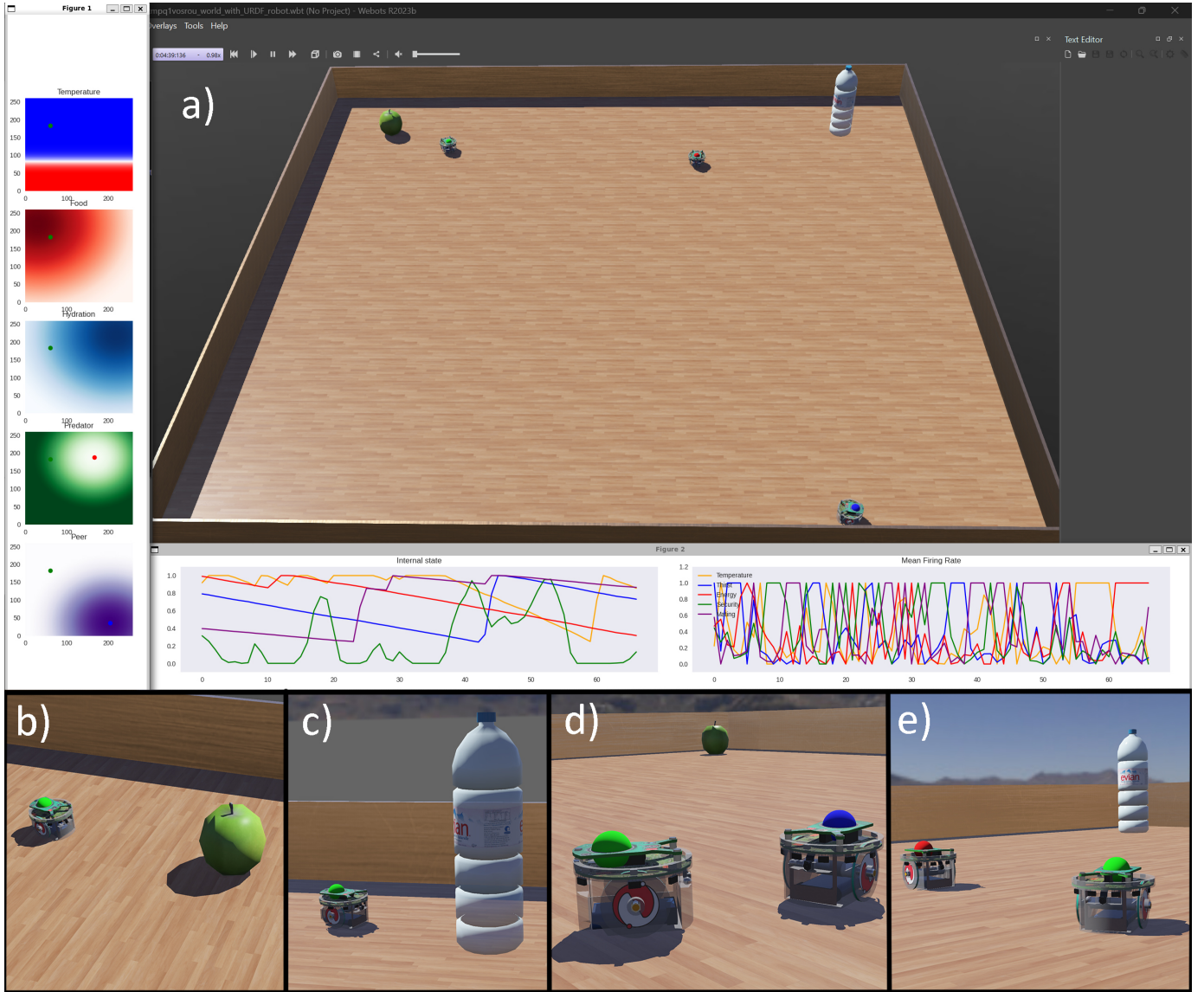


Fig. 1. Webots simulated experimental setup. a) General view of the experimental setup where the e-puck agent (marked in green) is approaching the energy source. To the left of the simulation view, we have the gradients that inform the agent about the location of the need-fulfilling resources. Below the simulation view, we can observe to the left the internal states of the agent (temperature in yellow, hydration in blue, energy in red, insecurity in green, and mating in purple), and to the right the resulting drives from the neural mass model (similar color code). At this precise moment, we observe that the agent has a low internal state for energy, and this results in a higher firing rate of the neural population encoding the energy drive. Thus, the robot navigates to the resource that will fulfill such a need. We can also observe in the temperature gradient that the agent (green dot) is located in the low-temperature area of the gradient and this makes the internal state of temperature decrease. However, if we observe the history of this signal, we can see that it was recently attended and that the previous dominating neural population of the neural mass model was the one encoding the temperature drive. b) Epuck agent approaching the energy source. c) Epuck agent approaching the hydration source. d) Allostatic agent approaching its mate (marked in blue) e) e-puck agent escaping from a predator (marked in red).

Orientation and internally-driven navigation are supported by a local observation of such gradient. For a thorough description of these methods, please consult [15].

IV. RESULTS

To analyze the performance of our allostatic agent in self-regulating its internal states we gathered data during the five day-night cycles that the simulation lasted and compare it with an interoceptive-agnostic agent that followed random navigation.

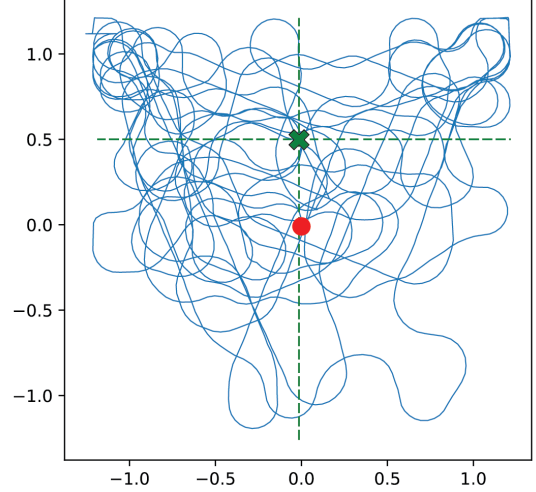
When analyzing agents' behavior we identify a spatial preference of the allostatic agent toward the upper part of the arena (where both energy and hydration sources are allocated) (Fig. 2a). We can observe this in both, the pattern of its trajectory and in the Y-axis mean position of the allostatic agent. In contrast, since the random navigation of the interoceptive-agnostic agent was produced in a closed environment and was enhanced with obstacle avoidance mechanisms, the navigation inevitably explores the complete arena but without any spatial bias (Fig. 2b).

When analyzing the internal states of both agents, we observe the allostatic agent performed better in maintaining higher internal states for hydration, energy, security, and mating. The lower internal state of temperature in the allostatic agent with respect to the interoceptive-agnostic agent is an expected effect imposed by the trade-off of orchestrating different internal needs. Moreover, we can see that not only thermoregulation is facilitated by higher environmental temperature, but it also facilitates hydration and energy (Fig. 3b-d). This effect is also expected since hydration and energy sources are allocated on the opposite side of the optimal temperature area. When the optimal temperature area is large, the distance between the temperature source and the hydration and energy sources is shorter, therefore facilitating their collective self-regulation. It is trickier to interpret the negative correlations of both the security and mating internal states of the agents (more pronounced in the allostatic one) since the location of the sources varied through the simulation (Fig. 3e, f). A possible explanation is that when the optimal temperature area shrinks and the allostatic agent needs to thermoregulate, its navigation is reduced to a smaller area where is less probable to accidentally encounter the predator or the mate.

To better understand the self-regulatory performance of the allostatic agent from a more holistic approach we computed the game-theoretic metrics of Efficiency, Fairness, and Stability [28], [29], and compare them with the interoceptive-agnostic agent. Efficiency informs about the ability of the agent in maintaining its internal states close to the setpoints. Fairness informs about the proportional consideration of the internal states when self-regulating. Finally, computed as the mean square homeostatic error, Stability represents a general measure of self-regulation performance. In the following equations, S is the internal state and P the setpoint (i.e. the desired value).

$$Efficiency = \frac{\sum_{i=0}^N \bar{S}_i(t)}{N} \quad (3)$$

a) Allostatic agent's Trajectory



b) Random agent's Trajectory

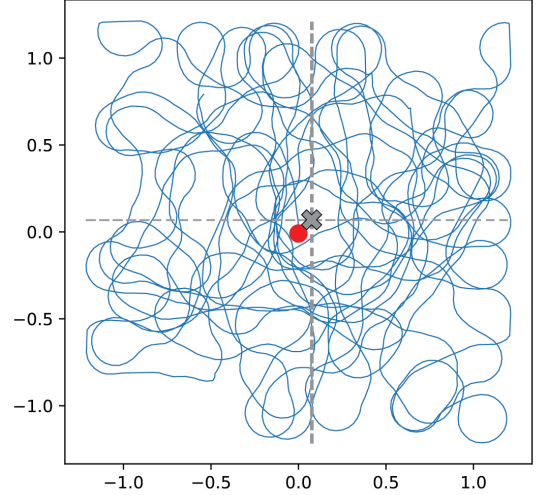


Fig. 2. **Agent navigation.** a) Allostatic agent's trajectory. b) Interoceptive-agnostic agent's trajectory. Red dots represent the agent's starting position. Dashed lines represent the X and Y-axis mean position for the five day-night periods that the simulation lasted. Cross marks represent the mean position of the agent.

$$Fairness = 1 - \max_{i \in N}(\bar{S}_i(t)) - \min_{j \in N}(\bar{S}_j(t)) \quad (4)$$

$$Stability = 1 - \frac{\sum_{i \in N} \|S_i(t) - P_i(t)\|_2^2}{N \cdot T} \quad (5)$$

The allostatic agent performed better in each of these measures. It showed a higher Efficiency (Fig. 4a), and therefore, it was able to maintain its internal states higher than the interoceptive-agnostic agent. It achieves it while increasing the Fairness between internal states (Fig. 4b). That is, although we observed that the allostatic agent maintained its internal temperature at a lower level than the interoceptive-agnostic agent, this internal need was not completely neglected and was

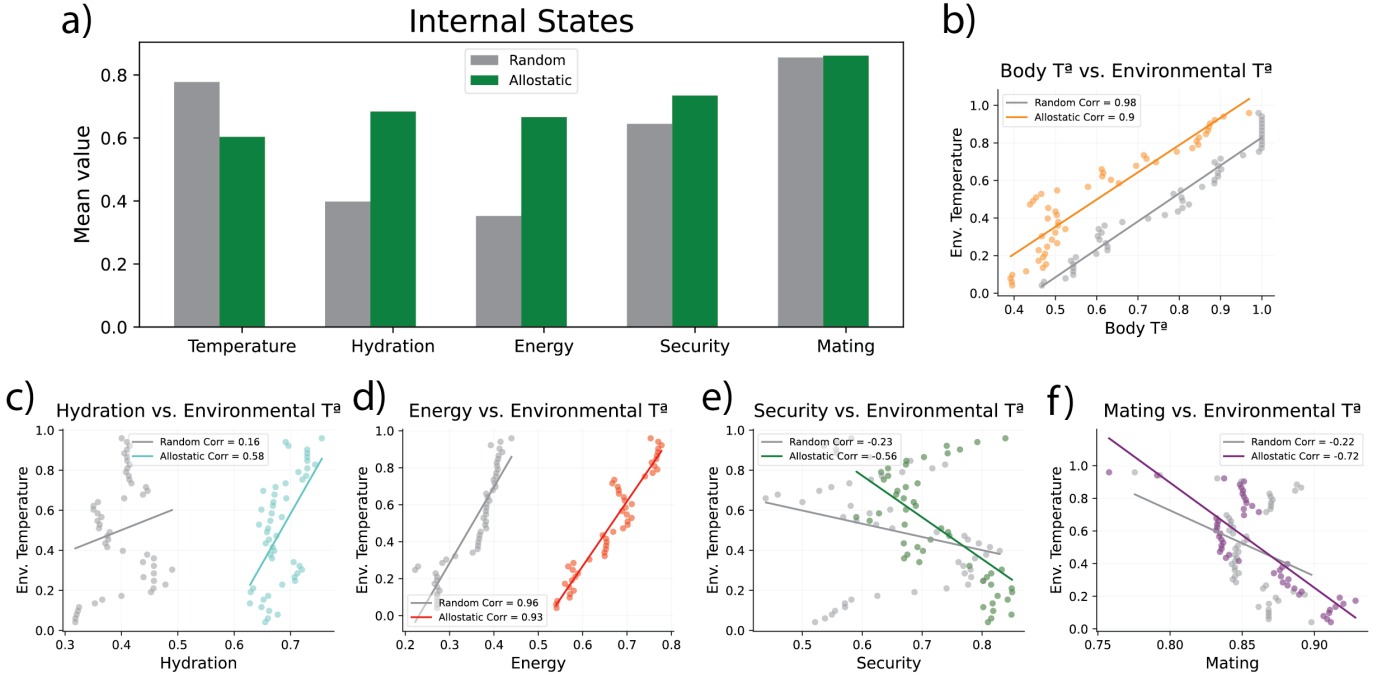


Fig. 3. **Agent internal states.** a) Mean internal states of the interoceptive-agnostic agent (gray) and the allostatic agent (green). b) Internal temperature measures correlated with the environmental temperature. c) Hydration states correlated with the environmental temperature. d) Energy states correlated with the environmental temperature. e) Security states correlated with the environmental temperature. f) Mating states correlated with the environmental temperature.

in favor of considering the other needs. Finally, the allostatic agent achieved higher Stability than the interoceptive agnostic agent (Fig. 4c), not only proving our methods effective but setting a record to be improved by future works.

V. DISCUSSION

In this study, we propose that robot autonomy can be guided by the principles of biological autonomy. To support this claim, we conducted a robot experiment replicating the core self-regulatory behaviors observed in the Namib desert lizard. The self-regulatory behaviors of the robot were driven by a biologically-grounded neural mass model of the hypothalamus, an ancient brain structure preserved across several species known for its role in drive regulation. This allostatic agent demonstrated remarkable autonomy, as shown by its capacity to attend to a set of several internal needs by navigating a dynamic environment and making decisions grounded in interoceptive measures such as temperature, hydration, energy, security, and mating. When compared to an interoceptive-agnostic agent, the allostatic agent outperformed it in three key game theory metrics: Efficiency, Fairness, and Stability. While the internal temperature was maintained at a lower level, it was done in favor of orchestrating other needs and ensuring self-stability.

Based on our findings, we propose that the adoption of a biomimetic approach toward self-regulation can significantly enhance autonomy in artificial systems. To illustrate the practical implementation of this approach, we consider the use case of an autonomous robot builder sent to the

moon for constructing base camps. Given the extreme lunar temperatures that could damage vital components like batteries or sensors, thermoregulation becomes a crucial internal need to self-regulate. Additionally, assuming our fictional robot relies on solar panels for energy, a trade-off arises between maintaining sufficient energy levels and managing temperature. Furthermore, foraging resources to build necessitates addressing the internal needs of material availability, which can be represented by the need to access moon ground and surface water. Lastly, the agent must also be equipped with a security drive to detect and avoid collisions with meteorites, which are relatively frequent on the moon. By employing a neural mass allostatic model, as presented in this study, the need for human teleoperation or rigid decision trees is unburdened.

Nonetheless, it is important to acknowledge the limitations of our study. While our synthetic agent demonstrated successful self-regulation of five internal needs, further research is required to explore the scalability of our model to a larger set of internal needs. This should also include considering cognitive needs in a motivational hierarchy, thus enabling more comprehensive and complex autonomous behaviors. Additionally, integrating our model into a cognitive architecture remains an avenue for future investigation.

In conclusion, our research highlights the promising link between biological autonomy and robot autonomy, presenting new insights into the achievement of autonomous behavior in biological agents and upgrading state-of-the-art robot autonomy with it. By utilizing the principles of biological self-

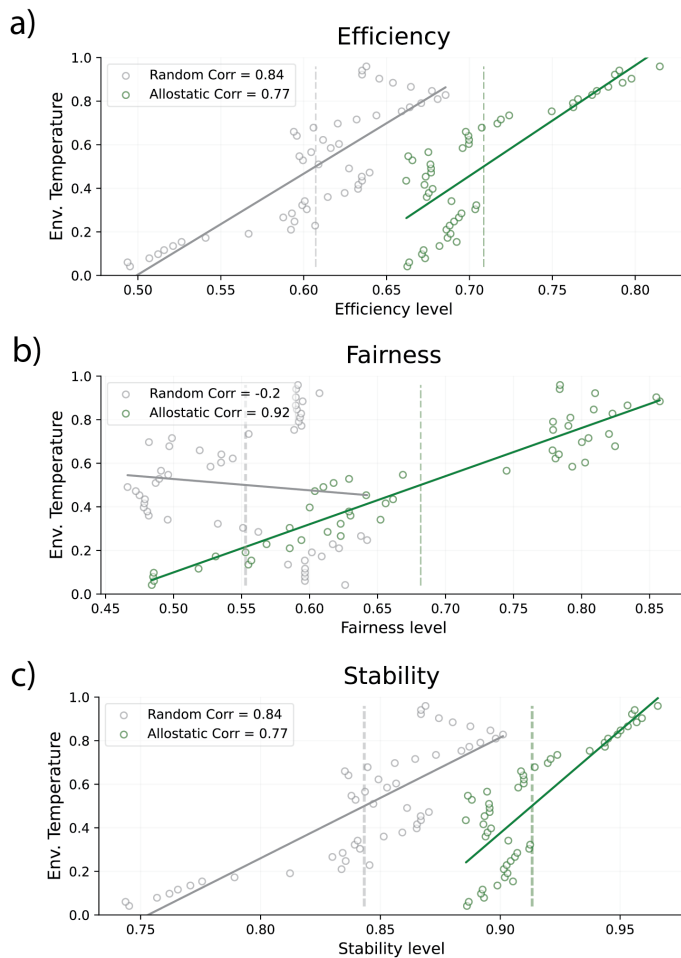


Fig. 4. **Efficiency, Fairness, and Stability measures.** a) Mean Efficiency measures correlated with the environmental temperature. b) Mean Fairness measures correlated with the environmental temperature. c) Mean Stability measures correlated with the environmental temperature. Dashed lines showed the mean value for each of these measures for the interoceptive-agnostic agent (gray) and the allostatic agent (green).

regulation, autonomous systems can become more adaptive, efficient, and robust, paving the way for groundbreaking applications in diverse fields.

REFERENCES

- [1] Maturana, H. R., & Varela, F. J. (1991). *Autopoiesis and cognition: The realization of the living* (Vol. 42). Springer Science & Business Media.
- [2] Varela, F. G., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems*, 5(4), 187-196.
- [3] Montévil, M., & Mossio, M. (2015). Biological organisation as closure of constraints. *Journal of theoretical biology*, 372, 179-191.
- [4] Levin, M. (2021). Life, death, and self: Fundamental questions of primitive cognition viewed through the lens of body plasticity and synthetic organisms. *Biochemical and Biophysical Research Communications*, 564, 114-133.
- [5] Rajendra Acharya, U., Paul Joseph, K., Kannathal, N., Lim, C. M., Suri, J. S. (2006). Heart rate variability: a review. *Medical and biological engineering and computing*, 44, 1031-1051.
- [6] Szymusiak, R. (2018). Body temperature and sleep. *Handbook of clinical neurology*, 156, 341-351.
- [7] Buck, C. L., Barnes, B. M. (1999). Annual cycle of body composition and hibernation in free-living arctic ground squirrels. *Journal of Mammalogy*, 80(2), 430-442.
- [8] Cannon, W. B. (1939). *The wisdom of the body*
- [9] Hull, C. L. (1952). *A behavior system; an introduction to behavior theory concerning the individual organism*.
- [10] Sterling, P., & Eyer, J. (1988). *Handbook of Life Stress, Cognition and Health* Edited by S. Fisher and J. Reason Copyright 1988 John Wiley & Sons. *Handbook of Life Stress, Cognition and Health*, January 1988.
- [11] Verschure, P. F. (2012). Distributed adaptive control: a theory of the mind, brain, body nexus. *Biologically Inspired Cognitive Architectures*, 1, 55-72.
- [12] Sanchez-Fibla, M., Bernardet, U., Wasserman, E., Pelc, T., Mintz, M., Jackson, J. C., ... & Verschure, P. F. (2010). Allostatic control for robot behavior regulation: a comparative rodent-robot study. *Advances in Complex Systems*, 13(03), 377-403.
- [13] Vouloutsi, V., Lallée, S., & Verschure, P. F. (2013). Modulating behaviors using allostatic control. In *Biomimetic and Biohybrid Systems: Second International Conference, Living Machines 2013, London, UK, July 29–August 2, 2013. Proceedings 2* (pp. 287-298). Springer Berlin Heidelberg.
- [14] Maffei, G., Santos-Pata, D., Marcos, E., Sánchez-Fibla, M., & Verschure, P. F. (2015). An embodied biologically constrained model of foraging: from classical and operant conditioning to adaptive real-world behavior in DAC-X. *Neural Networks*, 72, 88-108.
- [15] Rosado, O. G., Amil, A. F., Freire, I. T., & Verschure, P. F. (2022). Drive competition underlies effective allostatic orchestration. *Frontiers in Robotics and AI*, 9, 1052998.
- [16] Burnett, C. J., Li, C., Webber, E., Tsaousidou, E., Xue, S. Y., Bruning, J. C., et al. (2016). Hunger-driven motivational state competition. *Neuron* 92, 187-201
- [17] Osterhout, J. A., Kapoor, V., Eichhorn, S. W., Vaughn, E., Moore, J. D., Liu, D., ... & Dulac, C. (2022). A preoptic neuronal population controls fever and appetite during sickness. *Nature*, 606(7916), 937-944.
- [18] Petzold, A., Van den Munkhof, H. E., Figge-Schlensock, R., & Korotkova, T. (2023). Complementary lateral hypothalamic populations resist hunger pressure to balance nutritional and social needs. *Cell Metabolism*, 35(3), 456-471.
- [19] Stellar, E. (1954). The physiology of motivation. *Psychological review*, 61(1), 5.
- [20] Houston, A., & McFarland, D. (1976). On the measurement of motivational variables. *Animal Behaviour*, 24(2), 459-475.
- [21] Webots. (n.d.). Open-source Mobile Robot Simulation Software. Retrieved from <http://www.cyberbotics.com>.
- [22] Stock, J. B., & Baker, M. D. (2009). Chemotaxis. In *Encyclopedia of Microbiology*, Third Edition (pp. 71-78). Elsevier.
- [23] Basu, R., Gebauer, R., Herfurth, T., Kolb, S., Golipour, Z., Tchumatchenko, T., et al. (2021). The orbitofrontal cortex maps future navigational goals. *Nature* 599, 449-452.
- [24] Hollup, S. A., Molden, S., Donnett, J. G., Moser, M. B., & Moser, E. I. (2001). Accumulation of hippocampal place fields at the goal location in an annular watermaze task. *Journal of Neuroscience*, 21(5), 1635-1644.
- [25] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- [26] Sanchez-Fibla, M., Bernardet, U., & Verschure, P. F. (2010, October). Allostatic control for robot behaviour regulation: An extension to path planning. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1935-1942). IEEE.
- [27] Amil, A. F. & Verschure, P. F. (2021). Supercritical dynamics at the edge-of-chaos underlies optimal decision-making. *Journal of Physics: Complexity* 2, 045017
- [28] Binmore, K. (2005). *Natural justice*. Oxford: Oxford University Press.
- [29] Freire, I. T., Moulin-Frier, C., Sanchez-Fibla, M., Arsiwalla, X. D., and Verschure, P. F. (2020). Modeling the formation of social conventions from embodied real-time interactions. *PloS one* 15, e0234434.