cognitive and systemic postulates. Systems' self-referent operations driving the learning process are stimulated by irritations (internal states) (Luhmann 2006: 627).

« 3 » The second level is perhaps even more important from a sociological perspective. It corresponds to a social system that is drawn upon the interaction of agents' behavior. To talk about communicative systems, at least in the way I understand Luhmann's model, it is not enough that agents are able to broadcast signals (§62). Contingency is the link between the two levels. Contingency arises when an agent faces another agent whose behavior cannot be predicted but only expected. Alterity (otherness) is the heart of a sociological perspective. Luhmann depicted double contingency as two black boxes that deal with one another under the constraints of their own capacity to observe and influence (Luhmann 1991: 118f). The authors suggest that double contingency raised by this interaction could motivate an agent to select a particular aspect of its environment as a complexity reduction strategy (§61). The agents then could specialize into a role or particular behavior (§57) by triggering a modification in their neuronal structures. The agents are not individually preprogrammed to adopt one particular role. They are just programmed to adapt their behaviors under the conditions of their environment (§64), which is only possible because they have been first programmed to observe the environment through distinctions (e.g., food sources). The emergent order (social system) is conditioned by the complexity of the individual systems, without depending on their capacity to coordinate or control it (Luhmann 1991: 119).

« 4 » Does this process illustrate the emergence of social subsystems in accordance with Luhmann's theory? To answer this question, I think two concepts should be analyzed: "differentiation" and "binary code." The authors clearly take this direction in their conclusion. And I agree with their remarks, as pointed out in what follows.

« 5 » According to Luhmann (2006), the process of system differentiation can be triggered spontaneously as a result of evolution and can induce structural transformations. Differentiation takes place when a system/environment distinction is drawn within an existing system; the latter is seen as global in the eyes of the just-differentiated system (Luhmann 1991: 42). The authors illustrate the simplest and most trivial form of this differentiation, namely, one that carries no reference to the society as a whole and produces no subsequent formations within the system.

« 6 » A binary code can be established when the system is able to distinguish-and-select without losing the reference that created the distinction in the first place. Then there is always the possibility to turn to a (preexisting) negative value. This type of selection differs from basal operations that can only refer to identical elements and are blind to what is different (Esposito 1996). The distinction driven by binary code does not correspond to difference between system and environment but it refers to internal, accessible and contingent states of the system that duplicate the reality. A communication can accept or deny a certain value (e.g., a paper suggesting that an entire theory is false) but it cannot deny the importance of this distinction (Luhmann 1996: 222). In my opinion, this is the most substantial part of Luhmann's sociological work: it enables the treatment of a multi-contextural significant construction of reality within modern society.

« 7 » Porr & Di Prodi's model for simulation looks to be far from the conditions for generating a binary code. The systems they describe rely more on actions than on communications, and have no place for understanding. Without communication and language, is not that clear how acceptance/denial could play a part in this model (Luhmann 2006: 170). However, having a binary code is not a *sine qua non* condition for illustrating how social subsystems emerge. It is important to keep this in mind to avoid raising false expectations while evaluating the authors' contribution and its relevance to sociology. They are not simulating the emergence of functional systems, but showing that double contingency can lead to subsystem differentiation at the basal operations level.

Gastón Becerra is an assistant professor at the Universidad de Buenos Aires. He is pursuing a Ph.D. on philosophy of science and epistemology.

Received: 20 February 2014
Accepted: 22 February 2014

Authors' Response

What to Do Next: Applying Flexible Learning Algorithms to Develop Constructivist Communication

Bernd Porr
University of Glasgow, UK
bernd.porr/at/glasgow.ac.uk

Paolo Di Prodi
University of Glasgow, UK
robomotic/at/gmail.com

> Upshot • We acknowledge that our model can be implemented with different reinforcement learning algorithms. Subsystem formation has been successfully demonstrated on the basal level, and in order to show full subsystem formation in the communication system at least both intentional utterances and acceptance/rejection need to be implemented. The comments about intrinsic vs extrinsic rewards made clear that this distinction is not helpful in the context of the constructivist paradigm but rather needs to be replaced by a critical reflection on whether one has truly created autopoietic agents or just an engineering system.

« 1 » **Olivier Georgeon**'s commentary shows that there is a deep mistrust between two communities: on the one hand the reinforcement community, which has emerged from both engineering (e.g., optimal control) and economics, and on the other hand the community that has emerged from behaviour-based robotics, now branding itself as "enactive cognition." As **Patrick Pilarski** (§6) rightfully emphasized, the trouble comes from introducing the notion of reward. While in reinforcement learning the reward plays a central role, in enactive cognition it is very fashionable to claim that rewards are not part of the equation (**Georgeon** §9). However, rewards are then introduced through the back door in the form of "enjoyment," "dislike" or "intrinsic satisfaction." Even "sensor loops" (**Fabien Hervouet** §15), which learn to detect novelty, essentially reward themselves to be able to predict sensor states. As pointed out by Pilarski throughout his commentary, the boundary between intrinsic and extrinsic reward is blurred (§4) and rewards are and always will play a role in an experimental design in a more or less explicit manner. For us as authors writing a paper, using a reinforcement learning algorithm is treading on thin ice, being essentially halfway in the one camp and halfway in the other. However, bringing these two opposing views back together is probably easier than it appears to be (§13): the enactive cognition community needs to accept that rewards are part of the equation, for example as average rewards, and that it is politically correct to call them rewards (§7).

« 2 » At the same time, the reinforcement community needs to reflect on why the constructivist community and, more generally, the cognitive robotics community has such reservations against any form of rewards. This brings us back to its roots, namely engineering and economics. It comes as no surprise that engineering in particular is focused on results, namely *outputs*. However, from a constructivist perspective, agents control their *inputs,* which is the opposite of what an engineer wants to achieve. In Porr & Wörgötter (2005), we called this "the second chicken/egg problem": while the farmer wants to have the egg, the chicken wants to keep it. In other words, it is not so much an issue that we call a reward a "reward" or "satisfaction" but rather a problem of perspective. For example, a car should drive a passenger from A to B and not suddenly decide to make a detour to the cinema. However, for an autonomous agent that would be alright as long as it maintains its autopoiesis, which translates into the requirement that rewards need to be beneficial for the agent and not for the designer. This means that the use of the "R"-word in enactive cognition should be perfectly reasonable as long as the rewards help to maintain the agents' autopoiesis, usually through learning, as explained next.

« 3 » Based on these general remarks about the concept of reward, we can re-visit ICO learning, which has been used in this article. It is a biologically-inspired machine learning algorithm and belongs to the class of reinforcement learners where the actor and critic have been merged into one unit (Wörgötter & Porr 2005). We used it because of its very fast convergence, which shortened the execution of experiments considerably – at the expense of flexibility. Clearly, other learning algorithms of the class of reinforcement learning (Pilarski §10; Georgeon §11; Hervouet §4) would allow more flexible loops, the creation of new loops and also the generation of long-range predictions (Pilarski §8, §12; Hervouet, §§6f). This also addresses the criticism by Hervouet (§§6f, 13-15), who pointed out that ICO learning does not start off from random loops. This is certainly possible in any actor/critic scenario where the actor is allowed the freedom to create/remove any sensor/motor loop so that emergence is possible, in particular when employing the architectures suggested by Pilarski. We are in complete agreement that the learning suggested in his thought experiment (§§11f) would certainly lead to subsystem formation and would allow much more sophisticated behaviours. We also agree that an active exploration mechanism would greatly improve the agent's learning (Hervouet §9); however, this has been already discussed in our target article. In any case, all these alternative algorithms would have drastically increased the complexity of the experiment, which we wanted to keep as simple as possible to make our point about both the role of disturbances and subsystem formation.

« 4 » Hervouet (§16) pointed out that loops should become proactive, which is achieved with ICO learning and also with other reinforcement learning algorithms. However, it is debatable that an agent needs to become proactive at all to survive. One could argue that reactive behaviour is sufficient as long as it has the requisite variety and that this could be improved (Nakanishi & Schaal 2004). For example, a rabbit will always just run away from a fox but will never try to poison the fox in a proactive way to prevent further attacks. So being proactive is a huge evolutionary advantage and for that reason is used in our paper. However, one could argue that even humans could get by with being largely reactive.

« 5 » Hervouet (§§16f) also comments on the actual physical instantiation of the system. Do we need a physical system such as the Watt governor because its digital implementation in silico would not be appropriate? Any standard textbook on digital signal processing can easily provide an answer to that. The Watt governor is essentially a linear control system and performs its computations mechanically. A computer processes numbers that, in the case of navigating a real agent, arise from quantising analogue values, performing computations and transforming them back into analogue motor outputs. However, as long as the quantisation is smaller than the noise in the sensors (and the actuators), the digital processing will not change the dynamics of the system. It was shown *analytically* that one can perfectly replace parts of the Watt Governor with a digital controller without altering the dynamics of the system at all. Again, the main problem is not the implementation. This brings us back to the thorny issue of who defines the "reward" functions. Looking into the mechanical computations the Watt governor performs, it can easily be seen that it can also be described in terms of rewards because it has a setpoint (i.e., target value). For virtually all organisms, including humans, rewards have been provided by evolution (see the next paragraph). For artificial agents, we either need to simulate evolution or we can be inspired by the neural networks of animals and observe/measure how they have managed to stay alive in terms of rewards and punishments.

« 6 » Looking at real brains, primary rewards are usually coded, no matter if mouse or human brains, in the lateral hypothalamus (Nakamura & Ono 1986). These reward signals are then processed in the limbic system, where the most famous signal is the reward prediction error encoded by dopaminergic activity (Schultz 1998). More recently, it has been discovered that dopamine has two modes of operations: a phasic prediction error and slow tonic changes, where the latter seems to be encoding average rewards (Niv 2007). However, while the average reward plays an important role in the limbic (i.e., emotional) system, higher-level (conscious) decision making, especially in the basal ganglia, might operate reward-free and is probably driven by habits and/or novelty (Redgrave, Gurney & Reynolds 2008; Hervouet §9). This shows again that one needs to be careful when dealing with rewards: primary rewards certainly play an important role but at the same time one needs to acknowledge that they are not the sole drivers of action selection.

« 7 » Regardless of whether it is implemented in the traditional formulation as TD learning or in the advanced versions using average rewards, reinforcement learning employs closed loop processing. In his OPC, Georgeon stresses the fact that learning needs to be closed loop (§§1f) and that this kind of learning creates new loops, which he calls "schemas" in the spirit of Piaget. Our target article acknowledges the fact that ICO learning, though limited, is able to create new loops. However, one needs to be clear about why agents employ closed loop learning. Certainly, it has not been implemented to create "interesting" behaviours (Georgeon §§4-6). Such a requirement would create two problems.

1 | The first problem arises from the fact that the experimenter essentially performs output control by evaluating the agent's behaviour and then concludes that closed loop control should be preferred over open loop because it creates more "interesting" behaviour. However, while the agent might provide an output that is "interesting" for the observer, it might not be very beneficial for its own survival. For example, in the worst case, an external observer finds it highly "interesting" to observe an animal starving to death.

2 | To make things worse, Georgeon's writing suggests that one can freely choose between open and closed loop control when modelling autonomous agents. However, as outlined in the previous paragraphs, closed loop processing is not just a nice feature for creating "interesting" behaviours but it is fundamental for the constructivist paradigm because it performs input control, which guarantees that an agent can maintain its autopoiesis.

« 8 » Having now discussed input and output control extensively and concluded that agents perform input control, one needs to be careful when stating that the agent cannot access the environment's state (Georgeon §3). Again, it is a matter of perspective: experimenter vs agent. If one could ask the agent: "Are you aware of all of the environment's states, agent?," it would say: "Yes, totally." From the agent's perspective, its loops that act against the corresponding disturbances represent its *entire* environment, and the loops and disturbances represent the agent's "reality" of the world (Georgeon §7). Only when we observe the agent from the outside do we realise that its environment is infinite and thus the number of both accessible and hidden states is infinite as well. To have all states of the environment finite and observable to the agent can only be created artificially in very simple environments, which is not the case for actual organisms. For artificial agents this can be overcome – for example, by using embodied agents that have to deal with a complex environment.

« 9 » We are grateful that Hervouet comments on the environment (§§11f) and that altering it will impact on the subsystem formation. Indeed, a changing environment over time should cause a re-organisation of the subsystems and thus the predictive value readings of the individual agents. It would be an interesting follow-up experiment to change the number of food sources in a dynamic way. In a more general context, this also means that the makeup of the environment has a strong impact on the agent because (a) the disturbances force the agent to create appropriate loops and (b) these loops need to be operational with sufficient requisite variety.

« 10 » Finally, Gastón Becerra's excellent commentary is more an outlook than a criticism. It points out exactly how to develop our model further. Double contingency in our model is only established on the level of behaviour, where agents aim to predict each other. However, as already outlined in the target article, the level of communication has only been partially implemented. So far, our agents broadcast their states into the world continuously (Becerra §3), which leads to complete predictability in terms of the amount of food an agent carries. In a more advanced version of our model, the agents should be able to decide whether they want to release their states to the world in the form of utterances and to which agent. This may sound simple, but it requires loops and learning algorithms on the level of utterances and not just on the level of actions. In other words, we need to have both "motor" loops that generate actions (as done in our model) and loops that control the release of utterances. By creating separate loops for actions and utterances, we effectively create two subsystems. Consequently, the question arises of how they interpenetrate each other (to use Luhmann's words). The communication system would have an impact on the action system and vice versa. To enable a rich selection of utterances, one needs to allow the agent to release not only internal states but also sensor information and how sensor states lead to behaviour (by using weights, etc.).

« 11 » So far, we have just dealt with the release of utterances, which means that we have covered ego but not alter. In particular, alter observes an utterance by ego and can integrate it into its loops or just ignore it. However, this requires ultimately second-order cybernetics in the communication system so that agents can accept or reject certain communications (Becerra §§6f). This would then lead to a binary code that could accept or reject an utterance. As such, it is part of our on-going research.

Received: 3 March 2014
Accepted: 5 March 2014

Combined References

Ashby W. R. (1956) An introduction to cybernetics. Methnen, London.

Barber M., Blanchard P., Buchinger E., Cessac B. & Streit L. (2006) Expectation-driven interaction: A model based on Luhmann's contingency approach. Journal of Artificial Societies and Social Simulation 9(4). Available at http://jasss.soc.surrey.ac.uk/9/4/5.html

Beekman M., Nicolis S., Meyer B. & Dussutour A. (2009) Noise improves collective decision-making by ants in dynamic environments. Proceedings of the Royal Society London B 276: 4353–4361.

Braitenberg V. (1984) Vehicles: Experiments in synthetic psychology. MIT Press, Cambridge MA.

Camazine S., Franks N. R., Sneyd J., Bonabeau E., Deneubourg J.-L. & Theraula G. (2001) Self-organization in biological systems. Princeton University Press, Princeton NJ.

Cangelosi A., Tikhanoff V., Fontanari J. & Hourdakis E. (2007) Integrating language and cognition: A cognitive robotics approach. IEEE Computational Intelligence Magazine 2(3): 65–70.

Degris T., Pilarski P. M. & Sutton R. S. (2012) Model-free reinforcement learning with continuous action in practice. In: Proceedings of the 2012 American Control Conference, 27–29 June 2012, Montreal, Canada. IEEE Press, Piscataway NJ: 2177–2182.

Di Paolo E. A. (2005) Autopoiesis, adaptivity, teleology, agency. Phenomenology and the Cognitive Sciences 4(4): 429–452.

Dittrich P., Kron T. & Banzhaf W. (2003) On the scalability of social order. Modeling the problem of double and multi contingency following Luhmann. Journal of Artificial Societies and Social Simulation 6(1). Available at http://jasss.soc.surrey.ac.uk/6/1/3.html

Drescher G. L. (1991) Made-up minds: A constructivist approach to artificial intelligence. MIT Press, Cambridge MA.

Emmeche C., Koppe S. & Stjernfelt F. (2000) Levels, emergence, and three versions of downward causation. In: Andersen P. B., Emmeche C., Finnemann N. O. & Christiansen P. V. (eds.) Downward causation. University of Aarhus Press, Århus DK: 13–34.

Esposito E. (1996) From self-reference to autology: How to operationalize a circular approach. Social Science Information 35(2): 269–281.

Foerster H. von (2003) Understanding understanding: Essays on cybernetics and cognition. Spinger, New York.

Franks N. R., Mallon E. B., Bray H. E., Hamilton M. J. & Mischler T. C. (2003) Strategies for choosing between alternatives with different attributes: Exemplified by house-hunting ants. Animal Behaviour 65(1): 215–223.

Froese T. & Ziemke T. (2009) Enactive artificial intelligence: Investigating the systemic organization of life and mind. Artificial Intelligence 173(3–4): 466–500.

Gadenne V. (2010) Why radical constructivism has not become a paradigm? Constructivist Foundations 6(1): 77–83. Available at http://www.univie.ac.at/constructivism/journal/6/1/077.gadenne

Georgeon O. & Hassas S. (2013) Single agents can be constructivist too. Constructivist Foundations 9(1): 40–42. Available at http://www.univie.ac.at/constructivism/journal/9/1/040.georgeon

Georgeon O. & Marshall J. (2013) Demonstrating sensemaking emergence in artificial agents: A method and an example. International Journal of Machine Consciousness 5(2): 131–144.

Georgeon O., Marshall J. & Manzotti R. (2013) ECA: An enactivist cognitive architecture based on sensorimotor modeling. Biologically Inspired Cognitive Architectures 6: 46–57.

Georgeon O. & Sakellariou I. (2012) Designing environment-agnostic agents. In: Howley E., Vrancx P. & Knudson M. (eds.) Proceedings of the Adaptive and Learning Agents Workshop, 4–5 June 2012, Valencia, Spain: 25–32. Available at http://ai.vub.ac.be/ALA2012/

Georgeon O., Wolf C. & Gay S. (2013) An enactive approach to autonomous agent and robot learning. Proceedings of the IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (EPIROB2013) Osaka, Japan. 1–6.

Glasersfeld E. von (1984) An introduction to radical constructivism. In: Watzlawick P. (ed.) The invented reality: How do we know what we believe we know? W. W. Norton, New York: 17–40. Available at http://www.vonglasersfeld.com/070.1

Grant C. B. (2002) Complexities of self and social communication. In: Grant C. B., (ed.) Radical communication: Rethinking interaction and dialogue. John Benjamins, Amsterdam: 101–125.

Grondman I., Busoniu L., Lopes G. A. D. & Babuska R. (2012) A survey of actor-critic reinforcement learning: Standard and natural policy gradients. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 42(6): 1291–1307.

Heath B., Hill R. & Ciarallo F. (2009) A survey of agent-based modeling practices (January 1998 to July 2008) Journal of Artificial Societies and Social Simulation 12(4). Available at http://jasss.soc.surrey.ac.uk/12/4/9.html.

Hülse M. & Pasemann F. (2002) Dynamical neural Schmitt trigger for robot control. In: Dorronsoro J. R. (ed.) Artificial Neural Networks ICANN 2002: 783–788.

Jost J., Bertschinger N., Olbrich E., Ay N. & Fraenkel S. (2007) An information theoretic approach to system differentiation on the basis of statistical dependencies between subsystems. Physica A 378: 1–10.

Kaelbling L., Littman M. & Cassandra A. (1998) Planning and acting in partially observable stochastic domains. Artificial Intelligence 101: 99–134.

Kenny V. (2009) "There's nothing like the real thing." Revisiting the need for a third-order cybernetics. Constructivist Foundations 4(2): 100–111. Available at http://www.univie.ac.at/constructivism/journal/4/2/100.kenny

Kernbach S., Thenius R., Kornienko O. & Schmickl T. (2009) Re-embodiment of honeybee aggregation behavior in an artificial micro-robotic swarm. Adaptive Behavior 17: 237–259.

Klopf A. H. (1982) The hedonistic neuron: A theory of memory, learning, and intelligence. Hemisphere, Washington DC.

Leydesdorff L. (2005) Anticipatory systems and the processing of meaning: A simulation inspired by Luhmann's theory of social systems. Journal of Artificial Societies and Social Simulation 8(2). Available at http://jasss.soc.surrey.ac.uk/8/2/7.html.

Linden D. J. (2003) From molecules to memory in the cerebellum. Science 301: 1682–1685.

Littman M. L., Sutton R. S. & Singh S. (2002) Predictive representations of state. In: Dietterich T. G., Becker S. & Ghahramani Z. (eds.) Advances in Neural Information Processing Systems 14. Proceedings of the 2001 Conference. MIT Press, Cambridge MA: 1555–1561.

Luhmann N. (1984) Soziale Systeme. Suhrkamp, Frankfurt am Main.

Luhmann N. (1991) Sistemas sociales. Lineamientos para una teoría general. Anthropos, Barcelona. German original published in 1985.

Luhmann N. (1996) Die Realität der Massenmedien. Westdeutscher Verlag, Opladen.

Luhmann N. (1996) La ciencia de la sociedad. Anthropos, México City. German original published in 1990.

Luhmann N. (2006) La sociedad de la sociedad. Herder, México City. German original published in 1997.

Matarić M. J. (1997) Reinforcement learning in the multi-robot domain. Autonomous Robots 4(1): 73–83.

Maturana H. R. & Varela F. J. (1980) Autopoiesis and cognition: The realization of the living. Reidel, Dordrecht.

McCallum A. (1996) Learning to use selective attention and short-term memory in sequential tasks. In: Maes P., Mataric M. J., Meyer J.-A., Pollack J. & Wilson S. W. (eds.) From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior. MIT Press, Cambridge MA: 315–324.

Meyer B., Beekman M. & Dussutour A. (2008) Noise-induced adaptive decision-making in ant-foraging. Lecture Notes in Computer Science 5040: 415–425.

Modayil J., White A. & Sutton R. S. (2014) Multi-timescale nexting in a reinforcement learning robot. Adaptive Behavior, published online 7 February 2014.

Nakamura K. & Ono T. (1986) Lateral hypothalamus neuron involvement in integration of natural and artificial rewards and cue signals. Journal of Neurophysiology 55(1): 163–181.

Nakanishi J. & Schaal S. (2004) Feedback error learning and nonlinear adaptive control. Neural Networks 17(10): 1453–1465.

Niv Y. (2007) Cost, benefit, tonic, phasic: What do response rates tell us about dopamine and motivation? Annals of the New York Academy of Sciences 1104: 357–376.

O'Regan J. K., Myin E. & Noë A. (2005) Phenomenal consciousness explained (better) in terms of bodiliness and grabbiness. Phenomenology and the Cognitive Sciences 4(4): 369–387.

O'Regan J. K. & Noë A. (2001) A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences 24(5): 939–1031.

Oudeyer P.-Y., Kaplan F. & Hafner V. (2007) Intrinsic motivation systems for autonomous mental development. IEEE Transactions on Evolutionary Computation 11(2): 265–286.

Parsons T. (1968) Social interaction. In: Sills D. L. (ed.) International encyclopedia of the social sciences. Volume 12. Macmillan/Free Press, New York: 429–440.

Pfeifer R. & Scheier C. (1994) From perception to action: The right direction? In: Gaussier P. & Nicoud J.-D. (eds.) From perception to action. IEEE Computer Society Press, Los Alamitos CA: 1–11.

Pilarski P. M., Dick T. B. & Sutton R. S. (2013) Real-time prediction learning for the simultaneous actuation of multiple prosthetic joints. In: Proceedings of the 2013 IEEE International Conference on Rehabilitation Robotics, Seattle, USA, 24–26 June 2013. IEEE Press, Piscataway NJ: 1–8.

Porr B., Egerton A. & Wörgötter F. (2006) Towards closed loop information: Predictive information. Constructivist Foundations 1(2): 83–90. Available at http://www.univie.ac.at/constructivism/journal/1/2/083.porr

Porr B., Ferber C. von & Wörgötter F. (2003) ISO-learning approximates a solution to the inverse-controller problem in an unsupervised behavioural paradigm. Neural Computation 15: 865–884.

Porr B. & Wörgötter F. (2002) Isotropic sequence order learning using a novel linear algorithm in a closed loop behavioural system. Biosystems 67(1–3):195–202.

Porr B. & Wörgötter F. (2003) Isotropic sequence order learning. Neural Computation 15: 831–864.

Porr B. & Wörgötter F. (2005) What means embodiment for radical constructivists? Kybernetes 34(1/2): 105–117.

Porr B. & Wörgötter F. (2006) Strongly improved stability and faster convergence of temporal sequence learning by utilising input correlations only. Neural Computation 18(6): 1380–1412.

Reading N. C. & Sperandio V. (2006) Quorum sensing: The many languages of bacteria. FEMS Microbiology Letters 254(1): 1–11.

Redgrave P., Gurney K. & Reynolds J. (2008) What is reinforced by phasic dopamine signals? Brain research reviews 58(2): 322–339.

Redish A. D. (2013) The mind within the brain: How we make decisions and how those decisions go wrong. Oxford University Press, New York.

Riegler A. (2007) The radical constructivist dynamics of cognition. In: Wallace B. (ed.) The mind, the body and the world: Psychology after cognitivism? Imprint, London: 91–115. Available at http://www.univie.ac.at/constructivism/riegler/44

Roesch E., Spencer M., Nasuto S., Tanay T. & Bishop J.-M. (2013) Exploration of the functional properties of interaction: Computer models and pointers for theory. Constructivist Foundations 9(1): 26–32. Available at http://www.univie.ac.at/constructivism/journal/9/1/026.roesch

Salgado M. & Gilbert N. (2008) Emergence and communication: Overcoming some epistemological drawbacks in computational sociology. In: Proceedings of the Third Edition of Epistemological Perspectives on Simulation. Lisbon: 105–124.

Schmidhuber J. (1991) Curious model-building control systems. In: Proceedings of the International Joint Conference on Neural Networks, Singapore, Volume 2. IEEE Press, Piscataway NJ: 1458–1463.

Schultz W. (1998) Predictive reward signal of dopamine neurons. Journal of Neurophysiology 80: 1–27.

Singh S., Barto A. & Chentanez N. (2005) Intrinsically motivated reinforcement learning. In: Saul L. K., Weiss Y. & Bottou L. (eds.) Advances in Neural Information Processing Systems. MIT Press, Cambridge MA: 1281–1288.

Steels L. (2004) The autotelic principle. In: Fumiya I., Pfeifer R., Steels L. & Kunyoshi K. (eds.) Embodied artificial intelligence. Lecture Notes in AI 3139. Springer, Berlin: 231–242.

Sutton R. S. & Barto A. G. (1998) Reinforcement learning: An introduction. MIT Press, Cambridge MA.

Sutton R. S., Modayil J., Delp M., Degris T., Pilarski P. M., White A. & Precup D. (2011) Horde: A scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction. In: Tumer K., Yolum P., Sonenberg L. & Stone P. (eds.) Proceedings of the Tenth International Conference on Autonomous Agents and Multiagent Systems, 2–6 May 2011, Taipei, Taiwan. International Foundation for Autonomous Agents and Multiagent Systems, Richland SC: 761–768.

Tani J. & Nolfi S. (1999) Learning to perceive the world as articulated: An approach for hierarchical learning in sensory-motor systems. Neural Networks 12: 1131–1141.

Touchette H. & Lloyd S. (2004) Information-theoretic approach to the study of control systems. Physica A Statistical Mechanics and its Applications, 331: 140–172.

van Gelder T. (1997) Dynamics and cognition. In: Haugland J. (ed.) Mind design II: Philosophy, psychology, artificial intelligence. MIT Press, Cambridge MA: 421–450.

Varela F. J. (1989) Autonomie et connaissance: Essai sur le vivant. Edition Seuil, Paris.

Varela F. J., Thompson E. & Rosch E. (1991) The embodied mind: Cognitive science and human experience. MIT Press, Cambridge MA.

Weisbuch G. & Stauffer D. (2000) Hits and flops dynamics. Working Papers 00–07–036. Santa Fe Institute, Santa Fe NM.

Whitehead S. D. & Ballard D. H. (1991) Learning to perceive and act by trial and error. Machine Learning 7(1): 45–83.

Wischman S., Pasemann F. & Hülse M. (2004) Structure and function of evolved neurocontrollers for autonomous robots. Connections Science 16(4): 249–266.

Wörgötter F. & Porr B. (2005) Temporal sequence learning, prediction and control. A review of different models and their relation to biological mechanisms. Neural Computation 17(2): 245–319.