### **EDITORIAL**



# Extrinsic rewards, intrinsic rewards, and non-optimal behavior

Mousa Karayanni<sup>1</sup> · Israel Nelken<sup>1,2</sup>

Received: 5 October 2021 / Revised: 20 January 2022 / Accepted: 25 January 2022 / Published online: 5 February 2022 @ The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022, corrected publication 2022

#### Abstract

The optimality of behavior in experimental settings is usually determined with respect to an extrinsic reward defined by the experimenters. However, actions that do not lead to reward are ubiquitous in many species and in many experimental paradigms. Modern research on decision processes commonly treat non-optimal behaviors as noise, often excluding from analysis animals that do not reach behavioral performance criteria. However, non-optimal behaviors can be a window on important brain processes. Here we explore the evidence that non-optimal behaviors are the consequence of intrinsically motivated actions, related to drives that are different from that of obtaining extrinsic reward. One way of operationally characterizing these drives is by postulating intrinsic rewards associated with them. Behaviors that are apparently non-optimal can be interpreted as the consequence of optimal decisions whose goal is to optimize a combination of intrinsic and extrinsic rewards. We review intrinsic rewards that have been discussed in the literature, and suggest ways of testing their existence and role in shaping animal behavior.

Keywords Decision making · Optimal behaviour · Curiosity · Intrinsic motivation · Exploration · Reward processing

Modern research on behavior usually defines optimality using an objective designed by the experimenters. This objective is often based on the extrinsic reward (e.g. food pellets) obtained by the experimental subject, providing a clear and straightforward procedure to quantify behavior succinctly. The use of experimenter-defined behavioral goals makes it possible to compare behaviors under considerably different circumstances, manipulations and even across species.

However, animal behavior may be driven by more than just experimenter-defined objectives. For example, if the subject is driven by the extrinsic reward together with another one (e.g. the wish of the subject to maximize the frustration of the experimenter), optimal behavior would differ from that required to solely maximize the extrinsic-defined reward. Consequently, the observed behavior cannot be derived solely from the objectives set by the experimenter. Indeed, animal behavior often show evidence for deviations from optimality.

Mousa Karayanni mousa.karayanni@mail.huji.ac.il Take for example the paper by Trujano and Orduna (2015). By their definition rats are "optimal" while pigeons are "non-optimal". This classification is based on a choice task, in which subjects could choose one of two regimes for each trial. In one regime, two stimuli predicted distinct outcomes deterministically but with a low probability of reward. In the other regime stimuli poorly predicted the outcomes but the probability of reward was overall higher. Pigeons were considered to be "non-optimal" because they preferred the first regime, while rats were "optimal" for preferring the other. The underlying assumption of the researchers was that their externally defined reward is the sole driver of behavior. In this paper, such behavior will be termed *extrinsically directed behavior*.

It has become common to design and implement behavioral paradigms that minimize non-optimal actions, and to ignore them when interpreting behavior. For example, in experiments addressing decision making processes, researchers employ training phases aimed at reducing nonextrinsically directed behavior (e.g. Miller et al., 2017), omit non-extrinsically directed decisions from the analysis (e.g. Constantinople et al., 2019; Garenne et al., 2011) and exclude animals who fail to learn the task (where success and failure are determined by extrinsic measures) from the analyzed data (e.g. Brunton et al., 2013; Miller et al., 2017).

<sup>&</sup>lt;sup>1</sup> Edmond and Lily Safra Center for Brain Sciences, Hebrew University, Jerusalem, Israel

<sup>&</sup>lt;sup>2</sup> Department of Neurobiology, Silberman Institute of Life Sciences, Hebrew University, Jerusalem, Israel

Importantly, the behavior of the pigeons in Trujano and Orduna (2015) can be interpreted as optimal if the pigeons maximize a weighted sum of the extrinsic reward with the predictive information supplied by the stimulus. One way of conceptualizing such an optimization is through the introduction of the notion of an *internal reward*. Extrinsic reward induces internal reward that drives behavior, but internal reward may be acquired in other ways too. If we postulate that for the pigeons, high predictive information results in a substantial amount of internal reward, their behavior indeed optimizes the total internal reward.

Support for the notion that predictive information is a kind of intrinsic reward is provided by experiments employing the observing task (see Bromberg-Martin & Monosov, 2020; Cervera et al., 2020). In the observing task, the decision maker is faced with two risky options (or gambles), each delivering a certain reward quantity with a given probability, after a fixed delay period. Importantly, one option offers predictive information during the delay period on the success of the gamble (informative gamble). Results from multiple species (see Cervera et al., 2020) show strong preference for the informative gamble, even if its expected extrinsic reward is lower than for the non-informative gamble.

In general, these results can be viewed as part of the study of motivated decisions that are not directly driven by extrinsic reinforcements. Evidence for these motivated but unrewarded processes were found across multiple species in behaviors directed towards exploration, manipulation or probing of the environment. The motivations that drive these behaviors are postulated to be intrinsic (Baldassarre & Mirolli, 2013; Ryan & Deci, 2000). Notable examples of intrinsic motivations include novel object investigation in birds (Berlyne, 1966), explorative decision making in humans (Daw et al., 2006), and explorative behaviors by infants (Nöe, 2004; Rochat, 1989). Intrinsic motivations may shape behavior even in the presence of extrinsic reward. For example, hungry rats preferred to explore a novel feature before settling to eat (Berlyne, 1966; see also Thompson, 1953; Woods & Bolles, 1965; Zimbardo & Miller, 1958). Moreover, subjects may actively choose to trade extrinsic for intrinsic rewards. In (Deaner et al., 2005), macaque monkeys consistently traded juice for social information. In other experiments similar effects were observed in humans for different types of information (Reggev et al., 2021; Tamir & Mitchell, 2012). Furthermore, decisions directed towards novel options at the expense of other rewarding options were observed in humans and rats (Miller et al., 2017; Wittmann et al., 2008; see also Kakade & Dayan, 2002). Evidence also exists for intrinsic costs (or negative intrinsic rewards). These values represent behavioral outcomes and characteristics that animals and humans prefer to avoid or reduce. Examples include the complexity of behaviors and information collection/integration (Amir et al., 2020; Sims, 2010; Tishby & Polani, 2011).

Additionally, other reported intrinsic motivations have yet to be thoroughly researched, such as empowerment and competence driven behaviors (Klyubin et al., 2005; Baldassarre & Mirolli, 2013; Oudeyer & Kaplan, 2007). Thus, intrinsically motivated behaviors can be observed even in the presence of extrinsic reward, and can be found across species, settings, and experiments.

Some theoretical accounts of the intrinsically motivated behaviors mentioned above make use of the concepts of curiosity and play (Berlyne, 1966; Cervera et al., 2020; L. Fox et al., 2020; Loewenstein, 1994; van Lieshout et al., 2020), whose objectives are information gathering/integration and environment manipulation, respectively. Interestingly, two of the prominent psychologists to first address curiosity described it as an intrinsic cost (Berlyne, 1966; Loewenstein, 1994).

Some of the brain mechanisms underlying intrinsic rewards are known. For example, behaviorally-observed intrinsic rewards have been linked to brain regions critical for decision making processes. Notably, novel events and stimuli evoked responses in midbrain dopaminergic (DA) neurons known to encode reward prediction errors (RPE's) in mice and primates (Horvitz, 2000; Kakade & Dayan, 2002). Similarly, Wittman et al. (2008) suggested that midbrain novelty responses in humans were similar to the responses elicited by reward-related-events. Thus, decisions that are eventually rewarded by external reward and decisions that seem to be driven by the novelty of the stimulus share an internal reward signal. Additionally, in the same study, it was found that novelty directed decisions were correlated with ventral striatal BOLD signals, which are typically evoked by rewarded events. Predictive information is also encoded by neurons during decision making. In (Blanchard et al., 2015) neurons in the orbitofrontal cortex of monkeys were found to code the value of decisions based on a combination of the offered extrinsic reward and predictive information. Moreover, neurons in the lateral habenula were found to encode information prediction errors (IPE's), similarly to RPE's (Bromberg-Martin & Hikosaka, 2011), and midbrain DA neurons were found to respond to expectation of information delivery in the same way that they respond to expectations of reward delivery (Bromberg-Martin & Hikosaka, 2009). Importantly, these results provide evidence for the integration of extrinsic and intrinsic value in brain regions that are known to be involved in processing extrinsic reward and decision making.

Nevertheless, it is hard to study intrinsic motivations. The main difficulty is a lack of a good operational definition (Oudeyer & Kaplan, 2007). As a possible solution, psychologists and neuroscientists are using ideas, tools and terminology from the field of reinforcement learning (RL) (Daw et al., 2006, 2011; Dubey & Griffiths, 2020). Notably, RL tools are effective in characterizing decision strategies of human and animal subjects. These methods have been applied for example for studying exploration in behavioral experiments, particularly in relatively simple tasks such as the k-armed bandit task (see L. Fox et al., 2020). However, exploration was introduced in a somewhat simplistic form in these studies, so that a more nuanced view of exploration, as well as experiments in more complex settings, are needed to study exploration as an intrinsically driven process (L. Fox et al., 2020; Shteingart & Loewenstein, 2014).

It can be argued that intrinsic motivations (exploration in particular) are drives that increase future extrinsic yield, and for this reason they should not be interpreted as intrinsic drives. To counter this argument, we claim that reducing behaviors to their role in increasing extrinsic yield could blind us to their inner mechanisms. Specifically, it would be extremely difficult to discern the specific drives of behaving agents in complex environments based solely on their resultant extrinsic yield. Many are fundamentally different in their function and mechanisms, and their exact effects on extrinsic vield (in complex environments) may be hard to disentangle. Therefore, when attempting to unravel the different intrinsic components that may contribute to the decisions-making process, and especially for non-optimal behaviors, it is useful to employ the concept of intrinsic rewards. This is particularly so since we have the tools to observe brain activity and therefore to find traces of the internal processes producing the behaviors in question. In particular, we suggest that the theoretical treatment of different intrinsic motivations is important, in order to improve our understanding of the similarities and differences in their underlying mechanisms and resulting performance.

While in the computational RL literature, different intrinsic reward mechanisms are sometimes studied independently of extrinsic rewards to directly compare them (e.g. Choshen et al., 2018; Little & Sommer, 2013), we do not claim that intrinsic rewards should be solely studied separately from extrinsic rewards. Indeed, the utility of endowing agents with intrinsic rewards is commonly evaluated in environments that do offer extrinsic rewards (e.g. Bellemare et al., 2016; Burda et al., 2018; Pathak et al., 2017). This is so because the complex tasks RL aims to solve often consist of complex environments providing sparse rewards (Burda et al., 2018; Duan et al., 2016; Dubey & Griffiths, 2020). Analyzing the interaction between intrinsic and extrinsic rewards in such environments led to better understanding of intrinsic motivations as well as to improvements in the performance of simulated agents (Burda et al., 2018; Duan et al., 2016; Lopes et al., 2012). Similarly, when studying animal behavior, natural environments certainly provide extrinsic rewards, so that intrinsic and extrinsic rewards seem to operate simultaneously (Bromberg-Martin & Monosov, 2020; Cervera et al., 2020; Deaner et al., 2005; Mehlhorn et al., 2015; Miller et al., 2017; Wittmann et al., 2008). Therefore, we suggest to adopt the successful approach of computational RL studies to better understand the function of extrinsic and intrinsic rewards in behavior and in the brain.

Indeed, machine learning applications of RL often make use of intrinsic rewards. RL agents that only use extrinsic reward struggle to reach optimal performance in complex environments that provide sparse reward (Burda et al., 2018; Dubey & Griffiths, 2020). These types of environments commonly require effective and efficient exploration in order to achieve good performance. The provision of intrinsic rewards noticeably improves performance in such tasks (e.g. videogames and locomotion (Burda et al., 2018; Duan et al., 2016)). Initially inspired by results on intrinsic motivations in humans and animals, algorithms that produce rewards based on the experience of the agent, independently of external rewards, have been developed (Baldassarre & Mirolli, 2013; Klyubin et al., 2005; Little & Sommer, 2013; Oudeyer & Kaplan, 2007). Importantly, intrinsic motivations in RL produce rewards that are integrated together with extrinsic rewards into the values that drive agent performance. Moreover, these intrinsic motivations are often based on concepts from the behavioral and biological sciences (e.g. exploration, empowerment and mental effort (L. Fox et al., 2020; R. Fox et al., 2017; Klyubin et al., 2005; Little & Sommer, 2013)).

Taxonomies of theoretical intrinsic rewards have been suggested in the literature (Baldassarre & Mirolli, 2013; Oudeyer & Kaplan, 2007). Here we review existing taxonomies with examples, and suggest a modification (Fig. 1).



- Rational inattention (Sims, 2010)
  Empowerment (Klyubin et al., 2005)
  Complexity (mental effort) (R. Fox et al., 2017; Still & Precup, 2012; Tishby & Polani, 2011)
- Predictive power (Ay et al., 2008; Still & Precup, 2012)

Fig. 1 Taxonomy of intrinsic motivations

In (Baldassarre & Mirolli, 2013; Oudeyer & Kaplan, 2007), a distinction was made between knowledge-based and competence-based intrinsic motivations. Knowledge-based motivations serve to sample, gather and integrate information about the interaction of the agent with the environment. In this category, Baldassarre & Mirolli (2013) made a sub-distinction between novelty-based and prediction-based mechanisms. Novelty-based mechanisms drive the agent to experience all the available state-action pairs through novelty of actions and/ or observations. The goal of *prediction*-based motivations is to build an accurate model of the environment. Usually, these motivations work with an internal model of the environment, and guide behavior towards those parts of the world that are badly predicted. Competence-based motivations focus on improving the agent's interaction with the environment, rather than refining its knowledge or increasing its experience. Typically, these mechanisms employ quantifications of the agent's challenge, performance and efficiency in episodes of behavior (Oudeyer & Kaplan, 2007). Here we suggest expanding this category to motivations that drive towards different types of environment interactions as well as different characteristics of the state-action trajectory of the agent (to include intrinsic motivations such as empowerment (Klyubin et al., 2005) and complexity or mental effort (R. Fox et al., 2017; Still & Precup, 2012; Tishby & Polani, 2011)). We will refer to this generalized category as Interactive-based motivations.

We believe that the extensive literature and experience with intrinsic motivations in artificial RL agents creates a reverse opportunity: to use the existing quantitative models of intrinsic motivations, developed and studied in the RL setting, as tools for analyzing human and animal behavior. We believe that theoretically-derived intrinsic motivations can be used as tools to effectively address suboptimal behavior and its brain mechanisms. With such tools, it becomes possible to use available intrinsic reinforcers to infer optimal policies and compare them with observed behavior. By comparing policies that optimize given forms of intrinsic reward with behavior, it may be possible to identify the operation of intrinsic motivations as well as their nature. Neural activity associated with action selection and value representation can be analyzed also with respect to such policies. In fact, such analysis can already be applied to existing behavioral datasets and experimental environments, with the hope of shedding light on intrinsic rewards that may have been present but were not previously incorporated into the analysis. When pursuing such a program, it is important to acknowledge the potential pitfall of rationalizing behaviors with a superficial combination of intrinsic and extrinsic rewards that post-hoc fit the behavior well.

In conclusion, by quantifying intrinsic motivations, behavioral drives that were previously unaccounted for could

elucidate structure and meaning of seemingly non-optimal behaviors.

Taken together with the existing work on extrinsic reward, this would allow a richer understanding of decision making in behavior and in brains.

Authors' contributions MK and IN conceived the paper; MK and IN wrote the paper.

**Funding** This work was supported by grant no. 2016688 from the CRCNS BSF-NSF program and by personal grant no. 1126/18 from the Israel Science Foundation (ISF). MK is funded by a fellowship awarded from the Neubauer Foundation.

Availability of data and material Not applicable.

#### Declarations

**Code availability** Not applicable.

Conflicts of interest/Competing interests None.

## References

- Amir, N., Suliman-Lavie, R., Tal, M., Shifman, S., Tishby, N., & Nelken, I. (2020). Value-complexity tradeoff explains mouse navigational learning. *PLOS Computational Biology*, *16*(12), e1008497. https:// doi.org/10.1371/journal.pcbi.1008497
- Baldassarre, G., & Mirolli, M. (Eds.). (2013). Intrinsically Motivated Learning in Natural and Artificial Systems. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-32375-1
- Bellemare, M. G., Srinivasan, S., Ostrovski, G., Schaul, T., Saxton, D., & Munos, R. (2016). Unifying count-based exploration and intrinsic motivation. *Proceedings of the 30th International Conference on Neural Information Processing Systems*, 1479–1487.
- Berlyne, D. E. (1966). Curiosity and Exploration. Science. https://doi. org/10.1126/science.153.3731.25
- Blanchard, T. C., Hayden, B. Y., & Bromberg-Martin, E. S. (2015). Orbitofrontal Cortex Uses Distinct Codes for Different Choice Attributes in Decisions Motivated by Curiosity. *Neuron*, 85(3), 602–614. https://doi.org/10.1016/j.neuron.2014.12.050
- Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain Dopamine Neurons Signal Preference for Advance Information about Upcoming Rewards. *Neuron*, 63(1), 119–126. https://doi.org/10. 1016/j.neuron.2009.06.009
- Bromberg-Martin, E. S., & Hikosaka, O. (2011). Lateral habenula neurons signal errors in the prediction of reward information. *Nature Neuroscience*, 14(9), 1209–1216. https://doi.org/10.1038/nn.2902
- Bromberg-Martin, E. S., & Monosov, I. E. (2020). Neural circuitry of information seeking. *Current Opinion in Behavioral Sciences*, 35, 62–70. https://doi.org/10.1016/j.cobeha.2020.07.006
- Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science*, 340(6128), 95–98. https://doi.org/10.1126/ science.1233912
- Burda, Y., Edwards, H., Pathak, D., Storkey, A., Darrell, T., & Efros, A. A. (2018). Large-Scale Study of Curiosity-Driven Learning. [Cs, Stat]. http://arxiv.org/abs/1808.04355

- Cervera, R. L., Wang, M. Z., & Hayden, B. Y. (2020). Systems neuroscience of curiosity. *Current Opinion in Behavioral Sciences*, 35, 48–55. https://doi.org/10.1016/j.cobeha.2020.06.011
- Choshen, L., Fox, L., & Loewenstein, Y. (2018). DORA The Explorer: Directed Outreaching Reinforcement Action-Selection. [Cs, Stat]. http://arxiv.org/abs/1804.04012
- Constantinople, C. M., Piet, A. T., & Brody, C. D. (2019). An Analysis of Decision under Risk in Rats. *Current Biology*, 29(12), 2066-2074.e5. https://doi.org/10.1016/j.cub.2019.05.013
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, 69(6), 1204–1215. https://doi.org/10. 1016/j.neuron.2011.02.027
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095), 876–879. https://doi.org/10.1038/nature04766
- Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys Pay Per View: Adaptive Valuation of Social Images by Rhesus Macaques. *Current Biology*, 15(6), 543–548. https://doi.org/10.1016/j.cub.2005.01.044
- Duan, Y., Chen, X., Houthooft, R., Schulman, J., & Abbeel, P. (2016). Benchmarking deep reinforcement learning for continuous control. Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48, 1329–1338.
- Dubey, R., & Griffiths, T. L. (2020). Understanding exploration in humans and machines by formalizing the function of curiosity. *Current Opinion in Behavioral Sciences*, 35, 118–124. https:// doi.org/10.1016/j.cobeha.2020.07.008
- Fox, L., Dan, O., Elber-Dorozko, L., & Loewenstein, Y. (2020). Exploration: From machines to humans. *Current Opinion in Behavioral Sciences*, 35, 104–111. https://doi.org/10.1016/j.cobeha.2020.08.004
- Fox, R., Pakman, A., & Tishby, N. (2017). Taming the Noise in Reinforcement Learning via Soft Updates. [Cs, Math]. http://arxiv. org/abs/1512.08562
- Garenne, A., Pasquereau, B., Guthrie, M., Bioulac, B., & Boraud, T. (2011). Basal ganglia preferentially encode context dependent choice in a two-armed bandit task frontiers in systems. *Neuroscience*, 5. https://doi.org/10.3389/fnsys.2011.00023
- Horvitz, J. C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*, 96(4), 651– 656. https://doi.org/10.1016/S0306-4522(00)00019-1
- Kakade, S., & Dayan, P. (2002). Dopamine: Generalization and bonuses. *Neural Networks*, 15(4–6), 549–559. https://doi.org/10. 1016/S0893-6080(02)00048-5
- Klyubin, A. S., Polani, D., & Nehaniv, C. L. (2005). Empowerment: A Universal Agent-Centric Measure of Control. 2005 IEEE Congress on Evolutionary Computation, 1, 128–135. https://doi.org/ 10.1109/CEC.2005.1554676
- Little, D. Y., Sommer, F. T. (2013). Learning and exploration in actionperception loops. *Frontiers in Neural Circuits*, 7. https://doi.org/ 10.3389/fncir.2013.00037
- Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 116(1), 75–98. https:// doi.org/10.1037/0033-2909.116.1.75
- Lopes, M., Lang, T., Toussaint, M., & Oudeyer, P. (2012). Exploration in Model-based Reinforcement Learning by Empirically Estimating Learning Progress. Advances in Neural Information Processing Systems, 25. https://papers.nips.cc/paper/2012/hash/ a0a080f42e6f13b3a2df133f073095dd-Abstract.html
- Mehlhorn, K., Newell, B., Todd, P., Lee, M., Morgan, K., Braithwaite, V., Hausmann, D., Fiedler, K., Gonzalez, C. (2015). Unpacking the exploration-exploitation tradeoff: a synthesis of human and animal literatures. *Decision*, 2. https://doi.org/10.1037/dec0000033
- Miller, K. J., Botvinick, M. M., & Brody, C. D. (2017). Dorsal hippocampus contributes to model-based planning. *Nature Neuroscience*, 20(9), 1269–1276. https://doi.org/10.1038/nn.4613

Nöe, A. (2004). Action in Perception. MIT press.

- Oudeyer, P. Y., Kaplan, F. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, 1. https://doi.org/10.3389/neuro.12.006.2007
- Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2017). Curiosity-Driven Exploration by Self-Supervised Prediction. *IEEE Conference* on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, 488–489. https://doi.org/10.1109/CVPRW.2017.70
- Reggev, N., Chowdhary, A., & Mitchell, J. P. (2021). Confirmation of interpersonal expectations is intrinsically rewarding. *Social Cognitive and Affective Neuroscience*, *nsab081*. https://doi.org/ 10.1093/scan/nsab081
- Rochat, P. (1989). Object manipulation and exploration in 2- to 5-month-old infants. *Developmental Psychology*, 25(6), 871–884. https://doi.org/10.1037/0012-1649.25.6.871
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54–67. https://doi.org/10.1006/ ceps.1999.1020
- Shteingart, H., & Loewenstein, Y. (2014). Reinforcement learning and human behavior. Current Opinion in Neurobiology, 25, 93–98. https://doi.org/10.1016/j.conb.2013.12.004
- Sims, C. A. (2010). Chapter 4—Rational Inattention and Monetary Economics. In B. M. Friedman & M. Woodford (Eds.), Handbook of Monetary Economics (Vol. 3, pp. 155–181). Elsevier. https:// doi.org/10.1016/B978-0-444-53238-1.00004-1
- Still, S., & Precup, D. (2012). An information-theoretic approach to curiosity-driven reinforcement learning. *Theory in Biosciences*, 131(3), 139–148. https://doi.org/10.1007/s12064-011-0142-z
- Tamir, D. I., & Mitchell, J. P. (2012). Disclosing information about the self is intrinsically rewarding. *Proceedings of the National Academy of Sciences*, 109(21), 8038–8043. https://doi.org/10. 1073/pnas.1202129109
- Thompson, W. R. (1953). Exploratory behavior as a function of hunger in "bright" and "dull" rats. *Journal of Comparative and Physiological Psychology*, 46(5), 323–326. https://doi.org/10.1037/ h0054134
- Tishby, N., & Polani, D. (2011). Information Theory of Decisions and Actions. In V. Cutsuridis, A. Hussain, & J. G. Taylor (Eds.), *Perception-Action Cycle: Models, Architectures, and Hardware* (pp. 601–636). Springer. https://doi.org/10.1007/978-1-4419-1452-1\_19
- Trujano, R. E., & Orduña, V. (2015). Rats are optimal in a choice task in which pigeons are not. *Behavioural Processes*, 119, 22–27. https://doi.org/10.1016/j.beproc.2015.07.010
- van Lieshout, L. L., de Lange, F. P., & Cools, R. (2020). Why so curious? Quantifying mechanisms of information seeking. *Current Opinion in Behavioral Sciences*, 35, 112–117. https://doi.org/10. 1016/j.cobeha.2020.08.005
- Wittmann, B. C., Daw, N. D., Seymour, B., & Dolan, R. J. (2008). Striatal Activity Underlies Novelty-Based Choice in Humans. *Neuron*, 58(6), 967–973. https://doi.org/10.1016/j.neuron. 2008.04.027
- Woods, P. J., & Bolles, R. C. (1965). Effects of current hunger and prior eating habits on exploratory behavior. *Journal of Comparative* and Physiological Psychology, 59(1), 141–143. https://doi.org/ 10.1037/h0021614
- Zimbardo, P. G., & Miller, N. E. (1958). Facilitation of exploration by hunger in rats. *Journal of Comparative and Physiological Psychology*, 51(1), 43–46. https://doi.org/10.1037/h0048820

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.