Jennifer Nagel, University of Toronto
May 31, 2023

Lecture 2: Knowledge and surprise

SECTION 1: KNOWLEDGE, INFORMATION, AND THE TARGET OF CURIOSITY

1. Curiosity, or "the intrinsic desire to know" (Gottlieb & Oudeyer, 2018, 764), is said to appear in a great range of species, including primates, various birds, rats, octopuses, some say even fruit flies. To say that curiosity is an intrinsic desire for knowledge is to say that curious creatures in some sense want knowledge for its own sake, and not just for its contributions to securing other primary rewards, such as nourishment. But is there any reason to think that such a desire could be found in all these animals?

Daniel Berlyne: "Animals spend much of their time seeking stimuli whose significance raises problems for psychology" (1966: 25). He notes that (1) Playful and exploratory behavior emerges too early in life to be explicable by an association with the satisfaction of other basic needs. (2) Animals will sometimes attend to novel and complex stimuli even when it is hazardous to do so, and even in preference to satisfying immediate thirst or hunger.

2. What does curiosity aim at?

- **Knowledge** (e.g. Gottlieb & Oudeyer, 2018; Kang et al., 2009; Kobayashi, Ravaioli, Baranès, Woodford, & Gottlieb, 2019; Loewenstein, 1994)
- **Information** (e.g. Hsee & Ruan, 2016)
- **Both knowledge and information** "a reward-learning process of knowledge acquisition or information seeking" (Lau, Ozono, Kuratomi, Komiya, & Murayama, 2020); or, "a drive state for information" whose purpose is "to motivate the acquisition of knowledge and learning" (Kidd & Hayden, 2015, 450, 457); stimuli that enhance the value of knowledge (Dubey & Griffiths, 2020); "information-bearing stimulation" in animals, and both information-bearing stimulation and knowledge in humans (Berlyne, 1966).
- **How do knowledge and information relate?** Berlyne defines knowledge as "information stored in the form of ideational structures" (1966, 31); cf. "We define knowledge as information that is internal to the agent" (Silver, Singh, Precup, & Sutton, 2021).

3. Can we bring philosophical and empirical work together here?

Epistemologists (like me): Knowledge is the most general factive mental state, a state whose objective, reality-matching correctness is its essential core (Williamson, 2000). Non-factive states like belief can apply to true or false contents; knowledge, by contrast, can only be of the truth: if you actually know that *p,* then *p* must be the case (factivity).
Cognitive scientists: Curiosity researchers sometimes talk about varying levels of knowledge being "directly detected within the brain" (Baldassarre, 2011, 3). Q: Do we even mean the same thing by 'knowledge'?

4. My aim: to defend the idea that curiosity is a natural desire for gaining knowledge, understood as a factive mental state. This epistemological framework characterizes knowledge as a special kind of adaptation of the agent to reality. This adaptation can be driven by ordinary extrinsic motivations (such as thirst); it is accelerated by curiosity. Recent work in reinforcement learning (RL) helps us to see why biological agents like us need accelerated knowledge gain.

5. An instructive example: Jürgen Schmidhuber's (2010) vision-based agent encounters a screen broadcasting white noise. This is information-rich in the classic sense (Shannon, 1948); the signal cannot be compressed. If curiosity were just a drive to consume information, this screen should be fascinating. Schmidhuber notes that there are no patterns that the agent could make progress in compressing, no regularities that any agent could come to know (A dark room is also boring, because its regularity is learned at once, affording no further progress.) What the curious agent seeks is not just raw information, but information at a level that poses an appropriate challenge to the agent's current cognitive powers.

6. The core idea: Because our most basic form of learning runs on prediction-error correction, creatures like us gain knowledge when we are surprised, when events violate our expectations. Following a cue from reinforcement learning (RL), we can understand curiosity as a surprise-driven mechanism: surprise functions as a reward for the curious creature (in the technical sense of "reward"). Curious animals benefit from the adversarial interaction between their reactive prediction-error correction processes and the active surprise-seeking force of their curiosity.

SECTION 2: ANIMALS (INCLUDING US)

7. Rats "manifest curiosity apart from that curiosity which is directly associated with nutrition and reproduction" (Small, 1899: 90). Rats who are returned to their cages after being deprived of food for 24 hours will ignore openly available food if the bedding has been refreshed in their absence, and first "give themselves up for a while to explorations over and through their new bedding" (Dashiell, 1925: 208). Rats will also cross a floor that delivers a painful electric shock in order to explore new areas, and will do so when they are neither hungry nor thirsty (Warden, 1931).

8. Over 200 species of zoo animals in a classic study were given wooden blocks, steel chains, wooden dowels, rubber tubing and crumpled paper, for a series of six-minute test sessions (Glickman & Sroges, 1966). Pandemonium ensued. In a purely observational study, it can be hard to distinguish curiosity from various misapprehensions, amid interactions with the animal's current state of hunger, arousal, or fatigue. Upshot: controlled research is vital here.

9. Octopuses. Aristotle: "The octopus is a stupid creature, for it will approach a man's hand if it be lowered in the water" (*History of Animals* VIII(IX).37, 622a3-4). The approach bit is confirmed by researchers who drop plastic toys and clams into octopus tanks. Clams are touched more when the octopus was hungry as opposed to sated, but toys are explored with equal fervor in either condition (Kuba, Byrne, Meisel, & Mather, 2006a). Octopuses play with toys (Kuba, Byrne, Meisel, & Mather, 2006b), and can solve multi-step strategic problems (Richter, Hochner, & Kuba, 2016).

10. Monkeys also show strong tendencies towards sensorimotor exploration. Given a mechanical puzzle, solvable only by manipulating a series of six catches, clasps and levers in the right order, rhesus monkeys manipulate it avidly until they can solve it with high accuracy, producing a learning curve that "does not appear to differ in any way from a typical learning curve obtained on animals under hunger or thirst motivation" (Harlow, 1950).

11. Human ratings of surprise correlate with the degree to which the triggering experience runs contrary to current belief (for a review, see Reisenzein, Horstmann, & Schützwohl, 2019). Humans report higher curiosity to see the early resolution of gambles with greater outcome uncertainty (van Lieshout, Vandenbroucke, Müller, Cools, & de Lange, 2018). We are willing to endure physical pain to see trivial but highly chancy matters resolved. One experiment put human subjects in a room with a box containing some trick pens that would deliver a mildly painful electric shock when clicked, amid normal pens. Ten of the pens were marked with a red sticker, ten with a green sticker, and ten with a yellow sticker; subjects were told that the red stickers indicated a live battery which would always produce a painful shock on clicking the pen, the green stickers indicated a dead battery with no risk of shock, and the yellow ones were mixed. Left to their own devices, participants spontaneously clicked more of the yellow-sticker uncertain-outcome pens than both of the other colours put together. The unpredictability of the yellow-sticker pens seems to be part of their allure: those who simply wanted to alleviate boredom through variety could more easily have done so by systematically selecting green or red pens at will (Hsee & Ruan, 2016: 664).

12. What causes curiosity, and why? (1) On Carruthers's account, it is sparked by prediction error (e.g. the strange movement in the bushes), but this doesn't fit many cases (think of Hsee's pens). (2) Carruthers thinks that curiosity is just a diluted version of ordinary inquiry: "the core difference between intrinsically motivated states of questioning like curiosity and interest, on the one hand, and instrumentally motivated inquiries, on the other, is that the former are appraised against long-standing interests and values, whereas the latter are appraised against current goals"(2023: 26).

13. The "inverted-U-shape". Curiosity generally seems to be higher for questions on which we have a middling level of confidence (Loewenstein, 1994). For example, trivia questions produce lower reported curiosity among those who report either very low or high confidence of getting the answer right; peak reported curiosity appears in most people for questions on which they report a level of uncertainty near the midpoint (Kang et al., 2009). Something similar appears in the attentional patterns of infants: they are most attracted to moderately unpredictable visual events (Kidd, Piantadosi, & Aslin, 2012). This "Goldilocks effect" of seeking partial familiarity is seen as a way of keeping cognitive resources engaged in the zone between what is already known and what is currently unknowable.

SECTION 3: REINFORCEMENT LEARNING

14. The RL framework. (Haas, 2022; Sutton & Barto, 2020) In RL, the environment influences the agent by supplying a series of observations, and the agent influences the environment by producing a series of actions. Observations (chess board positions, pixel values in a video game, sensory signals) are a signal of the state of the environment. Given these observations, the agent computes the current state and selects an action according to its policy, leading to a new state supplying fresh observations, perhaps including an experience of reward. Reward is a special observation, a scalar signal that serves as a target for the agent, like points in a game for an agent who has the sole objective of maximizing his score; this reward signal may be densely or sparsely distributed among other observations.

15. Value and reward. A central puzzle in goal-directed action is how future reward can have a bearing on the present moment of choice. In RL this puzzle is solved by the agent's evolving representation of value, a measure of the extent to which available actions in each state tend to produce reward. The value of an action in a given state is the total average long-term reward it can be expected to lead to, conditional on the agent's policy, taking past experience as a guide. Crucially, when reward is finally encountered, the values assigned to the earlier state-action pairs that led to it are then updated to reflect the later payoff. Over time, in environments where there are meaningful state-action-reward relationships to be discovered, repeated experimentation with different state-action pairings allows the development of a meaningful representation of value. The computational back-tracking of assigning high value to actions on paths that end up yielding reward is a species of prediction-error learning; projected reward estimates get revised, over time, to match reality.

16. Exploration and knowledge. Good policies favor actions that have shown high value for the relevant situation, exploiting what has been learned to date, together with some exploration of other actions, to test whether undiscovered higher value elsewhere should prompt a further update to the policy. The simple ε-greedy approach chooses a random 'exploratory' action some small percentage of the time, and otherwise takes the 'greedy' action currently rated as having highest value, exploiting what has been learned to date. Viewed in isolation, exploratory actions might look like the antithesis of rational agency: they are independent of the agent's best current representation of value, even random. However, as long as the environment is not yet fully known, agents must incorporate some exploratory actions into their overall activity in order to avoid being trapped in local optima while larger reward lies elsewhere. When training converges to an optimal policy, the agent's valuations are factive mental states, factive because the truth of their contents is what causes their stable existence in the trained agent, and mental because of their role in guiding the trained agent's choice of action. RL agents really do gain knowledge, in the epistemologist's sense of the word.

17. Adding curiosity to an RL agent. Curiosity provides qualitative guidance for exploration—e.g. "explore what surprises you"' (Bellemare et al., 2016). There are several types of artificial curiosity (for a partial taxonomy, see Oudeyer & Kaplan, 2007). Agents who experience surprise as reward do better if they are making predictions involving lasting features of the environment, as opposed to predictions of raw sensory stimulation, where the state space is both very large and resistant to modelling (Pathak, Agrawal, Efros, & Darrell, 2017).

18. Biological agents respond to surprise: in particular, the phasic dopamine crucial to affective learning is produced by surprise (Barto 2013). If surprise functions as reward, reinforcement learning will teach agents to forgo small, short-term surprise (coin flips) in favor of actions with the potential to produce larger surprise in the long run.

19. If knowledge gain can be driven by pure extrinsic reward, why do we need curiosity? Knowledge can be gained by a completely incurious agent; AlphaZero is rewarded strictly for winning games. But chess is a fully observable environment, with games averaging 40 moves per side. The natural world is a partially observable environment, with extrinsic reward very sparsely distributed among a massive flood of sensory signals. However, animals can control which parts they observe, and the background causal structure of our relationship with the environment ensures that many observable features of reality ultimately have systematic relationships with extrinsic reward.

20. Exploration and sparse reward. While AlphaZero's successor MuZero is superhuman at chess and many Atari games, humans beat it easily at games of long-range strategic exploration (Schrittwieser et al., 2020). In Montezuma's Revenge, the player must navigate an extended 99-room labyrinth, searching for keys, amulets and other devices to unlock doors and defeat enemies later. The first reward is a key that can be grasped only through navigating a precise path along ladders and ropes, and jumping over a skull; it is estimated that random action sequences will attain this first reward only once every 500,000 attempts (Salimans & Chen, 2018). Curious RL does well here (Burda et al., 2018). Some RL agents even have a self-model enabling more strategic pursuit of prediction error (Haber, Mrowca, Wang, Fei-Fei, & Yamins, 2018).

21. Intrinsic reward scaffolds the pursuit of extrinsic reward. "Learning mechanisms of animals fail to function when there are long delays between the performed behaviours and the learning signals they cause. Moreover, there are few chances to produce, by trial-and-error, complex behaviours and long action chains that result in a positive impact on homeostatic needs. As a consequence complex behaviours and chains would never be learned based only on extrinsic motivations." (Baldassarre, 2011) Intrinsic motivations drive the agent to acquire "readily available building blocks" that can support complex actions later, as needed (Singh, Lewis, Barto, & Sorg, 2010). Curiosity is a thirst for knowledge gain, essential for animals who can learn more effectively, in the very limited time available, by actively striving to learn.

22. Curious humans ask questions. We spend large portions of our waking time in interactive conversation, much of which serves no obvious instrumental purpose (Dessalles, 2020; Mehl & Pennebaker, 2003). We also express our surprise to each other: for example, addressees say "oh" more than once a minute in ordinary conversation (Reece et al., 2023). This is the default epistemic change-of-state marker, typically signalling knowledge gain (Heritage, 2018). Are curious humans mapping out what others know, in the same way curious rats map out their spatial environment?

References:

Baldassarre, G. (2011). *What are intrinsic motivations? A biological perspective.* Paper presented at the 2011 IEEE international conference on development and learning (ICDL).

Bellemare, M., Srinivasan, S., Ostrovski, G., Schaul, T., Saxton, D., & Munos, R. (2016). Unifying count-based exploration and intrinsic motivation. *Advances in neural information processing systems, 29*.

Berlyne, D. E. (1966). Curiosity and Exploration. *Science, 153*(3731), 25-33.

Burda, Y., Edwards, H., Pathak, D., Storkey, A., Darrell, T., & Efros, A. A. (2018). Large-scale study of curiosity-driven learning. *arXiv preprint arXiv:1808.04355*.

Carruthers, P. (2023). The Contents and Causes of Curiosity. *British Journal for the Philosophy of Science*.

Dashiell, J. F. (1925). A quantitative demonstration of animal drive. *Journal of Comparative Psychology, 5*(3), 205.

Dessalles, J.-L. (2020). Language: The missing selection pressure. *Theoria et Historia Scientiarum, XVII*, 7-57.

Dubey, R., & Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity. *Psychological review, 127*(3), 455.

Glickman, S. E., & Sroges, R. W. (1966). Curiosity in zoo animals. *Behaviour, 26*(1-2), 151-187.

Gottlieb, J., & Oudeyer, P.-Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience, 19*(12), 758-770.

Haas, J. (2022). Reinforcement learning: A brief guide for philosophers of mind. *Philosophy Compass*. doi:10.1111/phc3.12865

Haber, N., Mrowca, D., Wang, S., Fei-Fei, L. F., & Yamins, D. L. (2018). Learning to play with intrinsically-motivated, self-aware agents. *Advances in neural information processing systems, 31*.

Harlow, H. F. (1950). Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *Journal of Comparative and Physiological Psychology, 43*(4), 289.

Heritage, J. (2018). Turn-initial particles in English. In J. Heritage & M.-L. Sorjonen (Eds.), *Between turn and sequence: Turn-initial particles across languages* (pp. 155-189). Amsterdam: John Benjamins.

Hsee, C. K., & Ruan, B. (2016). The Pandora Effect: The Power and Peril of Curiosity. *Psychological Science, 27*(5), 659-666.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-y., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science, 20*(8), 963-973.

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron, 88*(3), 449-460.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one, 7*(5), e36399.

Kobayashi, K., Ravaioli, S., Baranès, A., Woodford, M., & Gottlieb, J. (2019). Diverse motives for human curiosity. *Nature Human Behaviour, 3*(6), 587-595.

Kuba, M. J., Byrne, R. A., Meisel, D. V., & Mather, J. A. (2006a). Exploration and habituation in intact free moving Octopus vulgaris.

Kuba, M. J., Byrne, R. A., Meisel, D. V., & Mather, J. A. (2006b). When do octopuses play? Effects of repeated testing, object type, age, and food deprivation on object play in Octopus vulgaris. *Journal of Comparative Psychology, 120*(3), 184.

Lau, J. K. L., Ozono, H., Kuratomi, K., Komiya, A., & Murayama, K. (2020). Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. *Nature Human Behaviour, 4*(5), 531-543.

Liquin, E. G., & Lombrozo, T. (2020). Explanation-seeking curiosity in childhood. *Current Opinion in Behavioral Sciences, 35*, 14-20.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75.

Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of social life: a psychometric analysis of students' daily social environments and natural conversations. *Journal of personality and social psychology, 84*(4), 857.

Oudeyer, P.-Y., & Kaplan, F. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in neurorobotics, 1*, 1-14.

Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2017). *Curiosity-driven exploration by self-supervised prediction.* Paper presented at the International conference on machine learning.

Reece, A., Cooney, G., Bull, P., Chung, C., Dawson, B., Fitzpatrick, C., . . . Marin, S. (2023). The CANDOR corpus: Insights from a large multimodal dataset of naturalistic conversation. *Science advances, 9*(13), eadf3197.

Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The cognitive-evolutionary model of surprise: A review of the evidence. *Topics in Cognitive Science, 11*(1), 50-74.

Richter, J. N., Hochner, B., & Kuba, M. J. (2016). Pull or push? Octopuses solve a puzzle problem. *PloS one, 11*(3), e0152048.

Salimans, T., & Chen, R. (2018). Learning montezuma's revenge from a single demonstration. *arXiv preprint arXiv:1812.03381*.

Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE transactions on autonomous mental development, 2*(3), 230-247.

Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., . . . Graepel, T. (2020). Mastering atari, go, chess and shogi by planning with a learned model. *Nature, 588*(7839), 604-609.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal, 27*(3), 379-423.

Silver, D., Singh, S., Precup, D., & Sutton, R. S. (2021). Reward is enough. *Artificial Intelligence, 299*, 103535.

Singh, S., Lewis, R. L., Barto, A. G., & Sorg, J. (2010). Intrinsically motivated reinforcement learning: An evolutionary perspective. *IEEE transactions on autonomous mental development, 2*(2), 70-82.

Small, W. S. (1899). Notes on the psychic development of the young white rat. *The American Journal of Psychology, 11*(1), 80-100.

Sutton, R. S., & Barto, A. G. (2020). *Reinforcement Learning: An Introduction, Second Edition*. Cambridge: MIT Press.

van Lieshout, L. L., Vandenbroucke, A. R., Müller, N. C., Cools, R., & de Lange, F. P. (2018). Induction and relief of curiosity elicit parietal and frontal activity. *Journal of Neuroscience, 38*(10), 2579-2588.

Warden, C. J. (1931). Animal motivation experimental studies on the albino rat. In *Animal Motivation Experimental Studies on the Albino Rat*: Columbia University Press.

Williamson, T. (2000). *Knowledge and its Limits*. New York: Oxford University Press.