Title
A reinforcement learning framework for information-seeking and information-avoidance

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Cogliati Dezza, Irene
Molinaro, Gaia
Verguts, Tom

Publication Date
2023

Peer reviewed

A reinforcement learning framework for information-seeking and information-avoidance

Irene Cogliati Dezza (irene.cogliatidezza@gmail.com)
Department of Experimental Psychology, Ghent University,
2 Henri Dunantlaan, Ghent, Belgium
Department of Experimental Psychology, University College London,
26 Bedford Way, London, United Kingdom

Gaia Molinaro (gaiamolinaro@berkeley.edu)
Department of Psychology, University California, Berkeley,
2121 Berkeley Way, Berkeley, United States

Tom Verguts (tom.verguts@ugent.be)
Department of Experimental Psychology, Ghent University,
2 Henri Dunantlaan, Ghent, Belgium

Abstract

Every day, people are exposed to vast amounts of information that can impact how they feel, think about, and act upon the world. Here, we extend the computational reinforcement learning framework to explain how such an impact can shape future decisions to either seek or avoid information. By simulating human behavioral data, we showed that agents are more likely to seek information after exposure to information with a positive net impact on the agent's affect, cognition, and ability to make good decisions. The more the agent is exposed to this kind of information, the higher the probability that it will seek even more information in the future. On the contrary, decisions to remain ignorant are more likely to occur after repeated exposure to information with a negative net impact. Our model offers a novel computational framework within which maladaptive information-seeking and information-avoidance behaviors can be further investigated.

Keywords: information-seeking; motivation; reinforcement learning; computational modeling

Introduction

Thanks to modern advances in global communication through the Internet and social media, people have access to more information than ever before. Critically, each piece of information can impact how people feel, think about, and act upon the world (Cogliati Dezza, Maher, & Sharot, 2022; Sharot & Sunstein, 2020). For example, watching daily news on the climate crisis or the Ukraine war may make one feel sad or unsafe. At the same time, this information can enhance people's understanding of the world and might help taking actions that will yield future rewards and avoid losses.

In addition to impacting people's feelings, thoughts, and actions, consuming information may also shape future decisions to either seek further knowledge or remain ignorant. For example, if the consumed information causes a positive experience (e.g., by inducing happiness), one may be more likely to seek similar types of information than if it caused an overall negative experience (e.g., by inducing sadness). Indeed, recent theories and experimental findings suggest that the information can act as a reinforcer, similar to standard rewards (FitzGibbon, Lau, & Murayama, 2020; Marvin & Shohamy, 2016; Murayama, 2022). Here, we develop a novel computational framework to understand how the impact of information on how one feels, thinks about the world, and acts upon it can shape future decisions to seek or avoid information.

Our computational framework is a generalization of the standard reinforcement learning (RL) framework (Sutton & Barto, 1998). In RL, an agent interacts with the environment to learn how to maximize long-term rewards by estimating the expected reward value associated with available options via the calculation of reward prediction errors over a trial-and-error process. The agent then uses its acquired estimates to choose among available options. In our extended framework, the RL agent learns the likely impact of information on how it will feel (Affect), think about (Cognition), and act upon the world (Action) after consuming the information – similar to how state or state-action values are learned in the standard RL framework. A schematized summary of our extended framework is provided in **Figure 1**. To our knowledge, this is the first model that integrates multiple values to learn and direct information-seeking decisions.

To understand how the impact of information on how one feels, thinks, and acts can shape future information-seeking decisions, we expand the existing RL framework to incorporate recent findings (Cogliati Dezza, Maher, et al., 2022) and current theories of human information-seeking (Murayama, 2022; Sharot & Sunstein, 2020). We then use such extended framework to simulate behavioral data on an information-seeking task inspired by previously published experimental designs developed for human participants (Cogliati Dezza, Maher, et al., 2022). We further validated our framework by demonstrating it could capture real human behavior when adapted to a different information-seeking task (Hsee & Ruan, 2016).

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

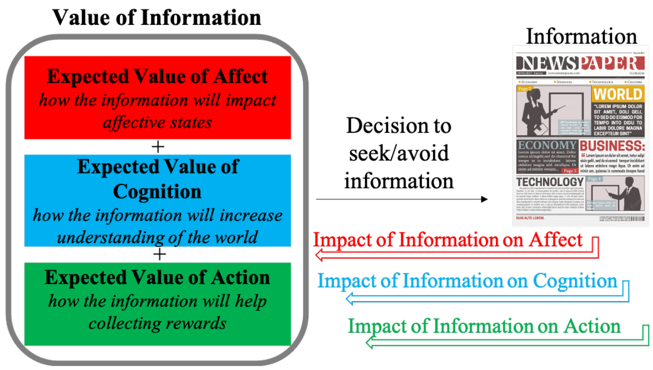

Figure 1: RL framework for information-seeking.

Methods

Computational framework

In our extended framework, the RL agent learns the expected value of Affect (i.e., the agent's prediction on how future information would impact its affective states), Cognition (i.e., the agent's prediction on how future information would impact its understanding of the environment in which it acts), and Action (i.e., the agent's prediction on how future information would impact the selection of reward-bearing actions) using the Rescorla-Wagner rule (Rescorla & Wagner, 1972):

$$Q_{motive}[\text{t}] = Q_{motive}[t-1] + \alpha_{motive}\ \delta_{motive}[t] \quad (1)$$

where $Q_{motive}[\text{t}]$ is the agent's prediction on how future information would impact the agent's affective states (if motive = Affect), its understanding of the environment in which it acts (if motive = Cognition) or the selection of reward-bearing actions (if motive = Action). $Q_{motive}[t-1]$ is the previous estimate, while $\delta_{motive}[t]$ is the current prediction error. $Q_{motive}[\text{t}]$ is only updated if the information is sought by the agent.

Following recent theories of information-seeking (Sharot & Sunstein, 2020), we allow the value of each motive to linearly integrate into the joint value of information:

$$V_{info}[\text{t}] = \beta_{affect} \times Q_{affect}[\text{t}] + \beta_{cognition} \times Q_{cognition}[t] + \beta_{action} \times Q_{action}[t] \quad (2)$$

where β_{affect}, $\beta_{cognition}$, β_{action} are the weights given to the expected values of Affect, Cognition, and Action, respectively. These weights determine the extent to which people are motivated to seek information that is expected to improve their affective states (β_{affect}), reduce their uncertainty ($\beta_{cognition}$) or help them collect future rewards (β_{action}). After estimating the joint value of information, the agent can use this overall estimate to direct its information-seeking decisions. This is formalized by entering V_{info} into a logistic probability function:

$$P_{info-seek}[\text{t}] = \frac{1}{1+\exp(cost - V_{info}[\text{t}])} \quad (3)$$

where the $cost$ term captures collective costs people incur when seeking information, such as money, time and physical/cognitive effort (Gottlieb, Cohanpour, Li, Singletary, & Zabeh, 2020; Horan, Daddaoua, & Gottlieb, 2019). We computed $P_{info-seek}$ in such a way to mimic non-deterministic choices, following many studies in the literature describing people's choices as a logistic function (Daw & Doya, 2006; Wilson & Niv, 2011; Wilson, Geana, White, Ludvig, & Cohen, 2014; Cogliati Dezza, Yu, Cleeremans, & Alexander, 2017; Cogliati Dezza, Noel, Cleeremans, & Yu, 2021; Cogliati Dezza, Cleeremans, & Alexander, 2022).

Each time a new piece of information is received, the agent integrates the valence of the consumed information into the prediction error as follows:

$$\delta_{affect}[t] = Info_{valence}[\text{t}] - Q_{affect}[t-1] \quad (4)$$

where $Info_{valence}$ is formalized as the expected utility of the information ($Info_{valence}[\text{t}] = \sum_{i=1}^{i=n} info_i \times p_i$). The same formulation has been shown to capture participants' information-seeking in past experiments and relates to participants' self-reported ratings of happiness (Cogliati Dezza, Maher, et al., 2022).

Similarly, the agent integrates its uncertainty after receiving the information as follows:

$$\delta_{cognition}[\text{t}] = Info_{uncertainty}[\text{t}] - Q_{cognition}[t-1] \quad (5)$$

where $Info_{uncertainty}$ is the inverse of the standard deviation of the received information($Info_{uncertainty}[\text{t}] = \frac{1}{\sqrt{\frac{\sum_{i=1}^{i=6}(c_i-\mu)^2}{N}}}$).

We entered the inverse of the standard deviation in the calculation of $\delta_{cognition}$ as high uncertainty corresponds to higher values of the standard deviation, and low uncertainty corresponds to lower values. We used the standard deviation as such formulation has been shown to adequately explain human information-seeking motivated by uncertainty reduction (Cogliati Dezza, Maher, et al., 2022; Bromberg-Martin et al., 2022) and relates to people's subjective uncertainty (Cogliati Dezza, Maher, et al., 2022).

Lastly, the outcome of the choice made after receiving the information is integrated as follows:

$$\delta_{action}[\text{t}] = Info_{action}[t] - Q_{action}[t-1] \quad (6)$$

where $Info_{action}$ is the outcome obtained from a choice made after receiving the information. Higher the outcome obtained from a choice, more the information did help in selecting rewarding-bearing actions.

Simulated Environments

Lottery Task Our first simulated environment was a lottery task (**Figure 2**). This task was inspired by existing information-seeking tasks developed for human participants (Cogliati Dezza, Maher, et al., 2022). In the lottery task, the agent was presented with six hidden cards drawn from an (integer) uniform distribution (depending on simulations, the range was either [-1, 5] or [-5, 1]). On each trial, the agent decided whether to uncover three of the six cards at a known cost or remain ignorant about the entire set of cards (**Figure 2A**). This task design was intended to mimic real-life scenarios in which people search for information online or on social media and decide whether to click on a provided link or open an app to retrieve initially hidden information. To include an instrumental component to this task, at the end of each trial, the agent had the choice to either enter a lottery or pass (**Figure 2B**). If the agent decided to play the lottery, one of the six cards (the "outcome card") was randomly chosen and the number printed on it determined the outcome of the trial, as the same amount was added to the agent's pot of money. If the agent decided to pass, no money was added to the agent's pot. By design, the three cards available for information purchase contained the outcome card among two decoy cards. Therefore, retrieving information about the cards would not only impact how the agent feels and its uncertainty but could also help the agent to maximize gains and reduce losses in future play or pass choices.

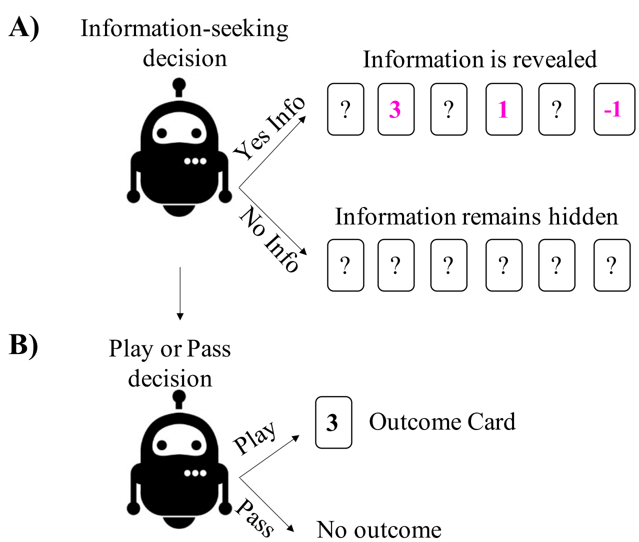

Figure 2: A trial of the lottery task. A) The agent decides whether to uncover three cards at a known cost or remain ignorant. B) The agent decides whether to play the lottery or pass.

Each time the agent decides to seek information, δ_{affect}, $\delta_{cognition}$, and δ_{action} are computed and integrated into Eq. 1. In this task, $Info_{valence}$ is the sum of the product of each uncovered card and the probability of its occurrence ($Info_{valence}[t] = \sum_{i=1}^{3} c_i \times p_i$), $Info_{uncertainty}$ is the inverse of the standard deviation of the information, and $Info_{action}$ is the outcome obtained from the lottery choice, which is zero if the lottery is not played.

After the information-seeking choice is made, the agent decides to play the lottery or pass. On trials in which the agent seeks information, the play and pass choices are determined by entering the expected utility of the information into a probability function:

$$P_{choice} = \frac{1}{1+\exp(cost - Info_{valence})} \quad (7)$$

On trials in which the agent decides not to seek information, it randomly decides whether to play or pass.

Shock Task Our second simulated environment was based on the Hsee & Ruan's task (Hsee & Ruan, 2016) in which participants were presented with a set of cues and they could decide to click on each of them to know whether an electric shock would be delivered. We, therefore, named it the "Shock Task". To mimic such scenario, we simulated the model in an environment that could deliver either negative ("yes shock") or positive outcomes ("no shock") using a uniform distribution of values within the range [-5, 5] (**Figure 3**), with -5 being very negative and 5 very positive.

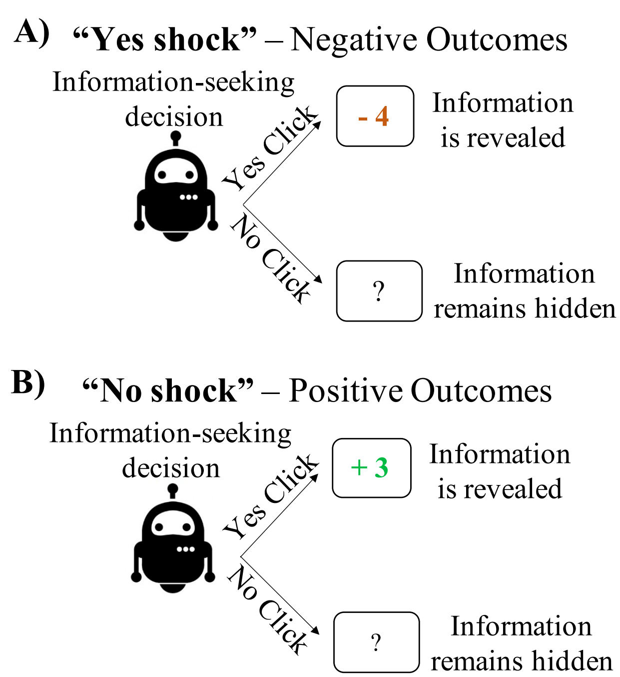

Figure 3: A, B) A trial of the shock task, where the information can either deliver a shock (A) or not (B).

In the original task, participants were exposed to two conditions: the certain condition and the uncertain condition. In the certain condition, participants knew in advance whether the cues would deliver shocks. Therefore, in this condition, participants would not reduce their uncertainty by clicking on the cues, but they would only experience the outcomes of their decisions (either positive or negative). To simulate the certain condition, we ran the above model where only Q_{affect} was entered into Eq. 2, such that only the degree to which the information feedback (no shock vs. shock) would influence the agent's affect would matter in its information-seeking decision (i.e., either clicking on the cues or not clicking on the cues):

$$V_{info}[t] = \beta_{affect} \times Q_{affect}[t] \quad (8)$$

On the contrary, in the uncertain condition, participants did not know in advance whether the cues would deliver shocks. Therefore, in this condition, participants would both reduce their uncertainty and experience the outcomes of their decisions (either positive or negative) by clicking on the cues. To simulate the uncertain condition, we ran the above model where $\boldsymbol{Q_{affect}}$ and $\boldsymbol{Q_{cognition}}$ were both entered into Eq. 2, such that the degree to which the information feedback would influence the agent's affect and the gain in its understanding of the cue both matter for the agent's final information-seeking decision:

$$V_{info}[t] = \beta_{affect} \times Q_{affect}[t] + \beta_{cognition} \times Q_{cognition}[t] \quad (9)$$

In both sets of simulations, we did not include Q_{action} as in this task the information has not an instrumental value (i.e., it can therefore not be used to improve future choices).

Results

The impact of information on Affect, Cognition and Action influences future information-seeking decisions

We first simulated our RL agent in a positively skewed environment in which the lottery cards were drawn from a uniform distribution within the range [-1, 5]. In this environment, information is most likely to be positive, reduce uncertainty about the future outcome, and help making choices that will yield future rewards. Therefore, in this environment, information has a *positive net impact*. The definition of the net impact concerns only the averaged current value of Q_{affect}, $Q_{cognition}$ and Q_{action} and does not include the weights given to these values in Eq. 2.

We simulated 100 agents, each playing the lottery task for 100 trials. We set the initial values for Q_{affect}, $Q_{cognition}$ and Q_{action} to 0 and their relative weights β_{affect}, $\beta_{cognition}$, β_{action} to 1. The cost of information was set to 0.1, while the three learning rates α_{affect}, $\alpha_{cognition}$ and α_{action} were set to 0.2.

As shown in **Figure 4A**, consuming information with positive net impact on Affect, Cognition, and Action increases the probability of seeking information in future trials. This translated into frequent decisions to seek information in such environment.

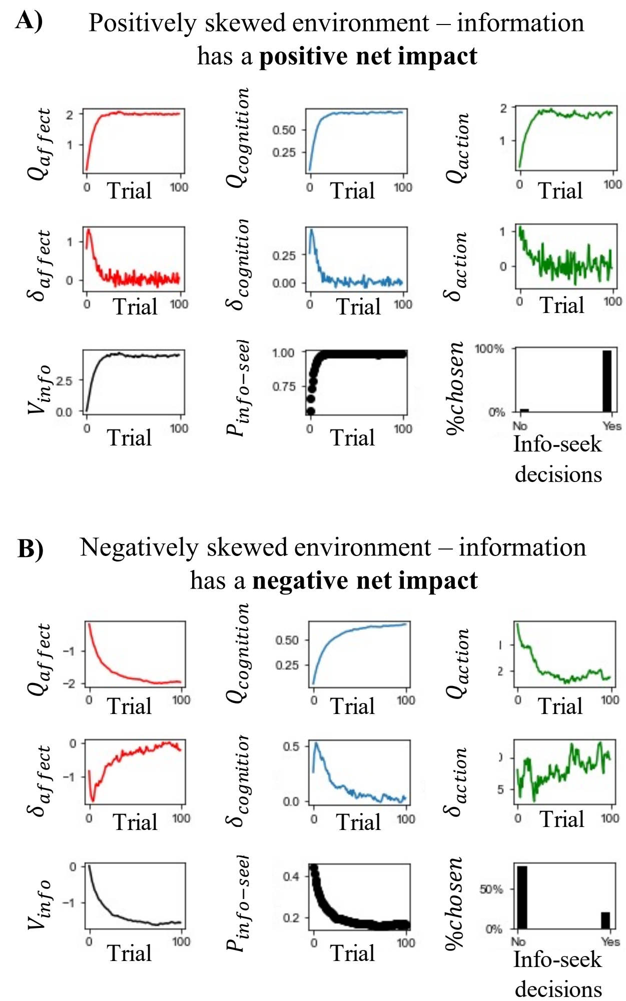

Figure 4: A) Consuming information with a positive net impact increases the probability of seeking information in future trials. B) Consuming information with a negative net impact decreases the probability of seeking information.

We then simulated the RL agent in a negatively skewed environment in which the lottery cards were drawn from a uniform distribution within the range [-5, 1]. In this environment, information is often negatively valenced and does not help in making lottery choices that will yield future rewards, but it still reduces uncertainty about the outcome card. In this environment, information has a *negative net impact*. As in the previous simulation, the definition of the net impact concerns only the averaged value of Q_{affect}, $Q_{cognition}$ and Q_{action} in such environment and does not include the weights given to these values in Eq. 2.

We again simulated 100 agents for 100 trials following the same model parametrization reported above. As shown in **Figure 4B**, consuming information with a negative net impact decreases the probability of seeking information in future trials. This translated into increased decisions to remain ignorant in such an environment.

To note, we obtained similar results with different values of learning rate (e.g., 0.5, 0.9) and cost (e.g., 0.5, 1.5) in both sets of simulations.

Taken together, the impact of information on Affect, Cognition, and Action influences future agent's information-seeking decisions in the same fashion as more standard rewards influence choices. In particular, when the consumed information has a positive net impact, the agent is more likely to seek information in the future, while when the consumed information has a negative net impact the agent is more likely to remain ignorant in the future.

Repeated exposure to information influence future information-seeking decisions more strongly

We then simulated the RL agent in both environments for a variable number of trials. In the positively skewed environment, we observed that the higher the number of exposures to information, the higher the probability that agents would seek information (**Figure 5A**). The opposite was true in the negatively skewed environment: the higher the number of exposures to information the more the agents decided to avoid information (**Figure 5B**).

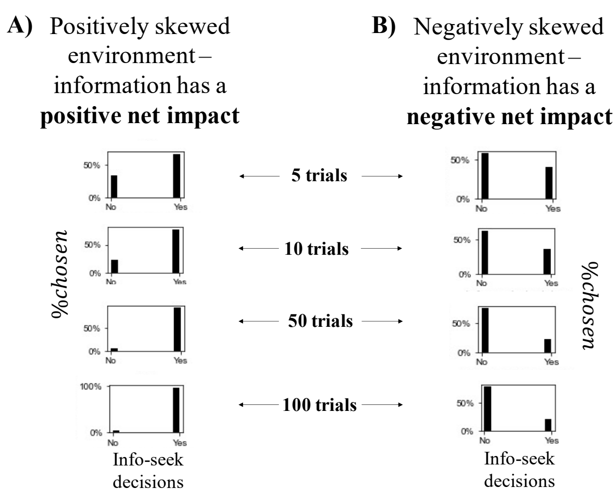

Figure 5: Impact of information on future information-seeking decisions across variable number of trials.

Interestingly, we observed that under the same model parametrization, in a positively skewed environment the agent required less exposure to information to learn that information had a positive net impact, while more trials were needed to learn that information had a negative impact in the negatively skewed environment (**Figure 6**). This is because in this latter environment, although information negatively impacts affect and action, it always reduces uncertainty about the lottery card.

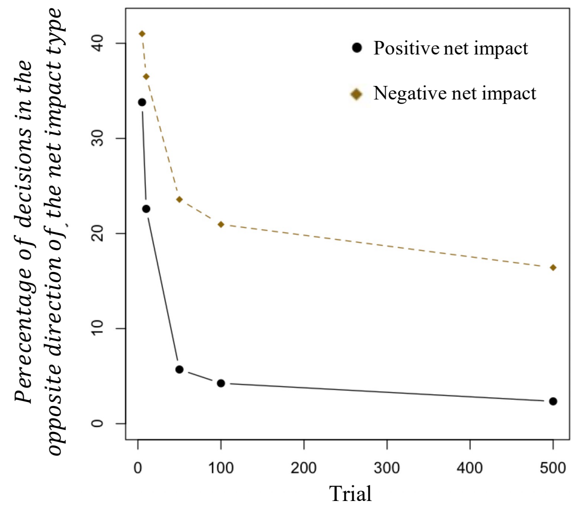

Figure 6: Percentage of decisions in the opposite direction of the impact type – that is, seeking information when it has a negative net impact and remaining ignorant when the information has a positive net impact.

Taken together, repeatedly exposing the agent to information with a positive or a negative impact influences future information-seeking more strongly.

RL-guided information-seeking can predict human decisions

To assess whether our model could capture real human behavior in the shock task, we simulated 100 RL agents for 100 trials of the shock task. The results of these simulations showed that the agents were seeking information more often in the uncertain condition compared to the certain condition (**Figure 7**). These results replicate the original findings, according to which people clicked on cues more often in the uncertain condition, where uncertainty could be reduced by seeking information, than in the certain condition (Hsee & Ruan, 2016).

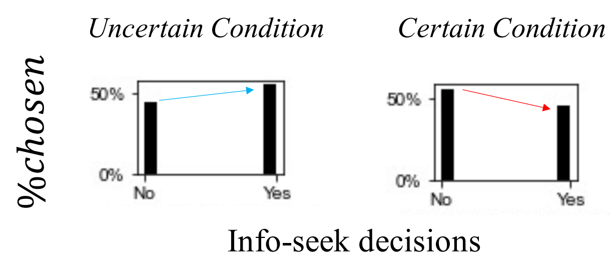

Figure 7: The agent's information-seeking decisions in the certain and uncertain conditions.

Discussion

Humans are exposed to vast amounts of information on a daily basis, and each piece of information can influence how they feel, think about, and act upon the world. Here, we have introduced a novel computational framework to understand how the impact of information on affect, cognition, and action can shape future decisions to seek information or remain ignorant. Overall, our simulations suggest that values related to information properties and learned via RL influence future information-seeking decisions, similar to how more standard rewards influence other behaviors.

In our first set of simulations with the lottery task, we showed that consuming information with a positive net impact on one's feelings, uncertainty reduction, and ability to make decisions that will yield rewards increases the probability of seeking further information. On the contrary, when one consumes information with a negative net impact, the probability of seeking information in the future decreases. These results are in line with recent theories and experimental findings that collectively consider information to be a reinforcer, similar to standard rewards (FitzGibbon et al., 2020; Marvin & Shohamy, 2016; Murayama, 2022). The impact of information can therefore be learned and influences future decisions to seek or avoid information.

In our second set of simulations, we showed that the number of exposures to positive or negative information affects future information-seeking decisions. In particular, the higher the number of exposures to information with a positive net impact, the more an agent will seek further information. On the contrary, the higher the number of exposures to information with a negative net impact the more an agent will avoid information. However, the amount of exposure required to stabilize information-seeking preferences varied across the two different environments. In particular, a smaller number of trials was required to learn that information had a positive net impact, and therefore should be sought, compared to the number of trials needed to learn that information had a negative net impact, and therefore should be avoided. These findings might explain why positive information is learned faster than negative information (Unkelbach, Fiedler, Bayer, Stegmüller, & Danner, 2008; Unkelbach et al., 2010). Because even negative information reduces uncertainty to some extent, the impact of negatively valenced information on Cognition is positive even when the environment frequently delivers negative news, thus driving the overall value of information up and pushing agents to seek more information (and slowing down the learning of the negative impact of such information on other dimensions – such as Affect and Action).

In line with recent findings, our simulations suggest that the agent seeks more information when the information enhances its affect, understanding of the world, and ability to make future rewarding choices. This is in line with recent findings that suggest people avoid information that makes them feel bad and seek information that makes them feel good (Charpentier, Bromberg-Martin, & Sharot, 2018; Karlsson, Loewenstein, & Seppi, 2009; Kobayashi, Ravaioli, Baranes, Woodford, & Gottlieb, 2019; Vellani, de Vries, Gaule, & Sharot, 2021), reduces their uncertainty (Chater & Loewenstein, 2016; Cogliati Dezza et al., 2021; Golman & Loewenstein, 2018; Singh & Manjaly, 2021; van Lieshout, Traast, de Lange, & Cools, 2021), and helps them select actions that will yield future rewards (Cogliati Dezza et al., 2017; Kobayashi & Hsu, 2019; Stigler, 1961; Wilson et al., 2014).

To further validate our model predictions, we simulated our RL agent using the shock task (Hsee & Ruan, 2016). We showed that, as observed in human participants who performed the same task, our RL agents were more likely to seek information when presented with the uncertain condition, in which the agent's uncertainty could be reduced by seeking more information, compared to the certain condition, in which no new knowledge could be gained by actively seeking information. These results show that our framework that can be used to explain how people seek information maladaptively (e.g., doomscrolling; Sharma, Lee, & Johnson, 2022) or avoid information when it could be useful (e.g., deliberative ignorance; Hertwig & Engel, 2021). For example, by assigning greater weights to uncertainty reduction and smaller weights to the affective outcomes of information, the model could persistently seek negative valanced information regardless of its impact on its affective states, yielding doomscrolling-like behaviors (Cogliati Dezza, Molinaro, & Verguts, under review).

In sum, by extending the RL framework to include parameters relevant to information alongside standard rewards, we showed that the impact of information can be learned in the same fashion as the value of state-action pairs learned via typical reinforcers (e.g., food or money). Such learned values can then influence future decisions to seek or avoid information, similar to how standard reinforcers strengthen or dampen other behavioral responses. Our model offers a novel computational framework within which maladaptive information-seeking and avoidance behaviors can be further understood and investigated.

References

Bromberg-Martin, E. S., Feng, Y., Ogasawara, T., White, J. K., Zhang, K., & Monosov, I. E. (2022). A neural mechanism for conserved value computations integrating information and rewards. *bioRxiv*.

Charpentier, C. J., Bromberg-Martin, E. S., & Sharot, T. (2018). Valuation of knowledge and ignorance in mesolimbic reward circuitry. *Proc Natl Acad Sci U S A, 115*(31), E7255-E7264. doi:10.1073/pnas.1800547115

Chater, N., & Loewenstein, G. (2016). The under-appreciated drive for sense-making. *Journal of Economic Behavior & Organization, 126*, 137-154.

Cogliati Dezza, I., Cleeremans, A., & Alexander, W. H. (2022). Independent and interacting value systems for reward and information in the human brain. *Elife, 11*. doi:10.7554/eLife.66358

Cogliati Dezza, I., Maher, C., & Sharot, T. (2022). People adaptively use information to improve their internal and external states. *Cognition, 228*. doi:10.1016/j.cognition.2022.105224

Cogliati Dezza, I., Molinaro, G., & Verguts, T. (under review). When cognition overshadows affect maladaptive information-seeking arises.

Cogliati Dezza, I., Noel, X., Cleeremans, A., & Yu, A. J. (2021). Distinct motivations to seek out information in healthy individuals and problem gamblers. *Transl Psychiatry, 11*(1), 408. doi:10.1038/s41398-021-01523-3

Cogliati Dezza, I., Yu, A. J., Cleeremans, A., & Alexander, W. (2017). Learning the value of information and reward over time when solving exploration-exploitation problems. *Sci Rep, 7*(1), 16919. doi:10.1038/s41598-017-17237-w

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Curr Opin Neurobiol, 16*(2), 199-204. doi:10.1016/j.conb.2006.03.006

FitzGibbon, L., Lau, J. K. L., & Murayama, K. (2020). The seductive lure of curiosity: information as a motivationally salient reward. *Current Opinion in Behavioral Sciences*.

Golman, R., & Loewenstein, G. (2018). Information gaps: a theory of preferences regarding the presence and absence of information. *Decision*.

Gottlieb, J., Cohanpour, M., Li, Y., Singletary, N., & Zabeh, E. (2020). Curiosity, information demand and attentional priority. *Current Opinion in Behavioral Sciences, 35*. doi:10.1016/j.cobeha.2020.07.016

Hertwig, R., & Engel, C. (2021). *Deliberate Ignorance*: MIT press.

Horan, M., Daddaoua, N., & Gottlieb, J. (2019). Parietal neurons encode information sampling based on decision uncertainty. *Nat Neurosci, 22*(8), 1327-1335. doi:10.1038/s41593-019-0440-1

Hsee, C. K., & Ruan, B. (2016). The Pandora Effect: The Power and Peril of Curiosity. *Psychol Sci, 27*(5), 659-666. doi:10.1177/0956797616631733

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: selective attention to information. *J Risk Uncertain 38*(2), 95–115.

Kobayashi, K., & Hsu, M. (2019). Common neural code for reward and information value. *Proc Natl Acad Sci U S A, 116*(26), 13061-13066. doi:10.1073/pnas.1820145116

Kobayashi, K., Ravaioli, S., Baranes, A., Woodford, M., & Gottlieb, J. (2019). Diverse motives for human curiosity. *Nat Hum Behav, 3*(6), 587-595. doi:10.1038/s41562-019-0589-3

Marvin, C. B., & Shohamy, D. (2016). Curiosity and reward: Valence predicts choice and information prediction errors enhance learning. *J Exp Psychol Gen, 145*(3), 266-272. doi:10.1037/xge0000140

Murayama, K. (2022). A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic-extrinsic rewards. *Psychol Rev, 129*(1), 175-198. doi:10.1037/rev0000349

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. . *Classical conditioning: Current research and theory*, 64–99.

Sharma, B., Lee, S., & Johnson, B. (2022). The dark at the end of the tunnerl: Doomscrolling on social media newsfeeds. *Technology, Mind, and Behavior, 3*(1).

Sharot, T., & Sunstein, C. R. (2020). How people decide what they want to know. *Nat Hum Behav, 4*(1), 14-19. doi:10.1038/s41562-019-0793-1

Singh, A., & Manjaly, J. A. (2021). The effect of information gap and uncertainty on curiosity and its resolution. *Journal of Cognitive Psychology, 33*(4), 403-423.

Stigler, G. (1961). The economics of information. *J Polit Econ* (69), 213–225.

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

Unkelbach, C., Fiedler, K., Bayer, M., Stegmüller, M., & Danner, D. (2008). Why positive information is processed faster: the density hypothesis. . *J Pers Soc Psychol*. doi:https://doi.org/10.1037/0022-3514.95.1.36.

Unkelbach, C., von Hippel, W., Forgas, J. P., Robinson, M. D., Shakarchi, R. J., & Hawkins, C. (2010). Good things come easy: Subjective exposure frequency and the faster processing of positive information. *Social Cognition, 28*(4), 538–555. doi:https://doi.org/10.1521/soco.2010.28.4.538

van Lieshout, L. L. F., Traast, I. J., de Lange, F. P., & Cools, R. (2021). Curiosity or savouring? Information seeking is modulated by both uncertainty and valence. *PLoS One, 16*(9), e0257011. doi:10.1371/journal.pone.0257011

Vellani, V., de Vries, L. P., Gaule, A., & Sharot, T. (2021). A selective effect of dopamine on information-seeking. *Elife*.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore-exploit dilemma. *J Exp Psychol Gen, 143*(6), 2074-2081. doi:10.1037/a0038199

Wilson, R. C., & Niv, Y. (2011). Inferring relevance in a changing world. *Front Hum Neurosci, 5*, 189. doi:10.3389/fnhum.2011.00189