When cognition overshadows affect, maladaptive information-seeking arises

Irene Cogliati Dezza [1,2,3], Gaia Molinaro [4], Tom Verguts [1]

1 Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, Ghent, Belgium
2 Department of Experimental Psychology, University College London, 26 Bedford Way, London, United Kingdom
3 The Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, 10-12 Russell Square, London, United Kingdom
4 Department of Psychology, University of California, Berkeley, 2121 Berkeley Way, Berkeley, USA

Abstract

In today's digital age, people have unprecedented access to vast amounts of negative information such as news about wars, environmental disasters, or economic collapse. While some people selectively avoid such distressing content, others persistently immerse themselves in such information, even to the detriment of their well-being and mental health – a phenomenon commonly known as doomscrolling. Here, we introduce a novel reinforcement learning model and empirically test its predictions to shed light on the origins of this maladaptive information-seeking behavior. First, by simulating human data, we showed that doomscrolling arises when the motivation to reduce uncertainty is higher than the motivation to improve affect through information-seeking; a pattern we term the "doomscrolling fingerprint". Next, we tested these predictions in the real population using data from an information-seeking task (N = 190) and a follow-up survey (N = 95). Consistent with our model's predictions, participants who presented a doomscrolling fingerprint sought more negatively valenced information when performing the information-seeking task. Most importantly, these participants reported a greater incidence of doomscrolling in real-life behavior. Our study takes a significant step toward unraveling the intricate psychological mechanisms behind maladaptive information-seeking behaviors prevalent in today's digital age, paving the way to future behavioral interventions.

Statement of Relevance

Why do people persistently consume negatively valenced information even if it jeopardizes their well-being and mental health? While this phenomenon, commonly known as "doomscrolling", is often observed, the psychological mechanisms underlying it are unknown. Nevertheless, gaining insights into these mechanisms is vitally important for developing better interventions and prevent the negative consequences of maladaptive information-seeking on people's well-being and mental health. In this study, we take a significant step toward unraveling the origins of doomscrolling by developing and testing a novel computational model. We used the model to make predictions on how this behavior may arise and we tested these predictions in human participants. Overall, our results suggest that doomscrolling arises when the motivation to reduce uncertainty is higher than the motivation to improve affect with the consumed information.

Key words
Information-seeking, doomscrolling, motivated behavior, reinforcement learning, uncertainty reduction, affect.

Word count = 7 000

Introduction

In today's digital age, people have unprecedented access to vast amounts of information that can influence their wellbeing and mental health. For example, watching daily news on international conflicts or the climate disasters may evoke feelings of sadness or anxiety. Reading a sweet message from a beloved one on a messaging application, on the contrary, may make one feel happy and appreciated. Recent research suggests that people generally consume information that elicits positive affect and avoid information that triggers negative affect [1-7]. However, in our contemporary society this adaptive mechanism sometimes fails. This is evident in the widespread phenomenon of doomscrolling [8], wherein people persistently consume negatively valenced information even at the expense of their well-being and mental health. Indeed, a growing body of literature shows that doomscrolling can enhance the risk for developing psychopathology, including post-traumatic stress, anxiety, and depression disorders [9-14]. Here, we shed light on the origins of this maladaptive information-seeking behavior by developing a novel computational model and empirically testing the model's predictions.

Our computational model is a generalization of the well-known and widely used reinforcement learning (RL) framework. In this framework, an agent interacts with the environment to learn how to maximize long-term rewards by estimating expected reward values, and using these estimates to choose among available options [15]. Two reasons motivate us to use such a framework. First, doomscrolling implies making sequential decisions about which information to consume and which to avoid. Second, recent information-seeking theories [16-18] and experimental findings [4,19-24] suggest that information one consumes can function as a reinforcer, akin to standard rewards (e.g., money, incentives) and the brain generates an information prediction error that parallels a reward prediction error [2,25]. Therefore, the decision to seek information might be reinforced by the information itself through learning mechanisms [18]. Indeed, RL has been previously used to study information-seeking [26-32].

Contrary to previous RL models for information-seeking, however, our RL model integrates multiple motivational factors to decide which information to seek. This is because recent advances from the information-seeking field suggest that various information-seeking motives jointly contribute to information-seeking behavior [2,3,6,7,23,33-35]. When deciding whether to seek or avoid information, people consider whether the information will positively or negatively impact their affective states [1-3]. However, people also consider whether the future information will help them both make a better sense of the world by reducing uncertainty [34,36-40], and collect future rewards [23,31,41,42]. In particular, recent information-seeking frameworks [43] and experimental findings [6,7,33] suggest that these factors are integrated into a single value of information, which then directs information-seeking. We refer to these factors as Affect, Cognition, and Action aligning with the terminology in the information-seeking literature discussed above (most prominently, [43]. A schematized summary of our generalized model is provided in **Figure 1**).

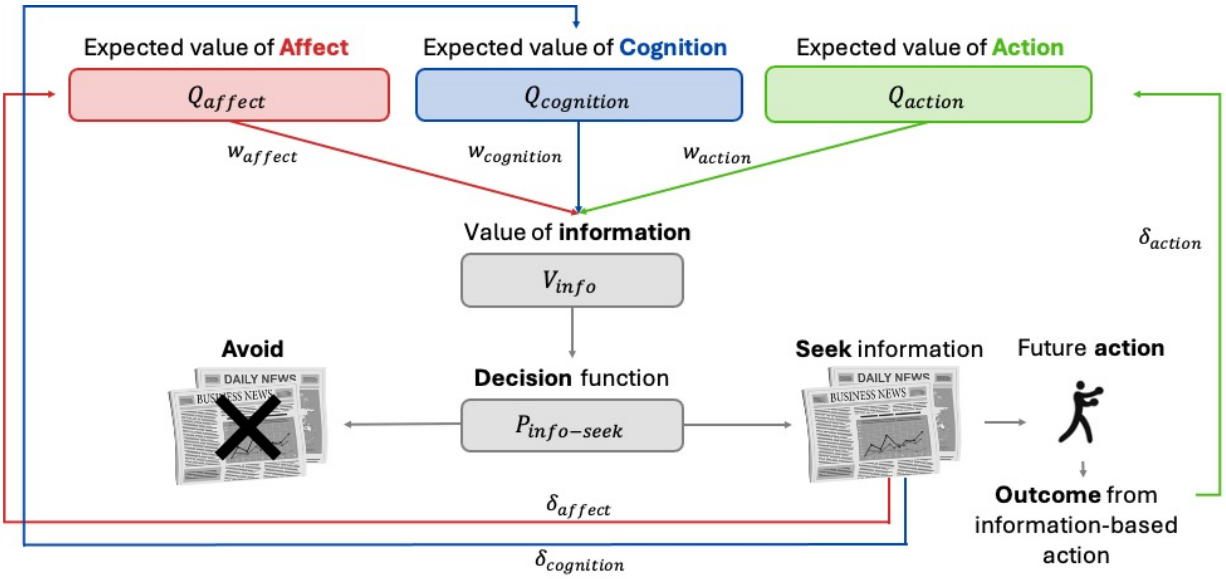


Figure 1. ***A generalized RL model of information-seeking.*** An agent's information-seeking decisions are determined by linearly combining the expected value of the information-seeking motivational factors (i.e., Q_{affect}, $Q_{cognition}$, Q_{action}) and their relative importance (i.e., w_{affect}, $w_{cognition}$, w_{action}) into the value of information (V_{info}). The greater the value of information, the higher the probability that the agent will decide to seek information $P_{info-seek}$. If the information is sought, the valence of information (i.e., whether it is positive or negative) and its uncertainty (i.e., whether the information has a higher or lower degree of uncertainty) are used to update Q_{affect} and $Q_{cognition}$ through their prediction errors (δ_{affect}, $\delta_{cognition}$), respectively. The information can then be used to select future actions whose outcome is similarly integrated into Q_{action} through the prediction error δ_{action}.

In order to shed light on how certain people persistently immerse themselves in distressing content (i.e., engage in doomscrolling), we introduce a novel computational model that integrates recent findings and current theories of human information-seeking into the RL framework and we empirically test its predictions. We first validate our model by simulating human data to computationally accounting for the finding (now reproduced several times) that people generally seek information that makes them feel good, reduces their uncertainty, and helps to collect future rewards. Next, we use the model to individuate the computational mechanisms of doomscrolling, and we empirically test the model's predictions in the real population using data from an information-seeking task and a follow-up survey. Overall, our results suggest that doomscrolling arises when the motivation to reduce uncertainty is higher than the motivation to improve affect through information-seeking. This computationally precise understanding of the widespread phenomenon of doomscrolling opens the way to novel interventions preventing the negative consequences of maladaptive information-seeking on people's well-being and mental health.

Results

To validate our model, we first aimed to computationally account for the general findings that people tend to seek information that makes them feel good, reduces their uncertainty, and helps them to collect future rewards (Simulations 1a, 1b). Next, we used the model to make predictions about doomscrolling (Simulation 2) and we tested these predictions in the real population (Study 1, Study 2).

To this end, we use our model to simulate behavior in a task described below and inspired by previous information-seeking experimental designs developed for human participants [7] **Figure 2**; **Methods**). Here, the RL agent learns the expected value of Affect (i.e., the agent's prediction on how future information will impact its affective states), Cognition (i.e., the agent's prediction on how future information will impact its understanding of the environment in which it acts), and Action (i.e., the agent's prediction on how future information will impact the selection of reward-bearing actions).The expected value of each motive is learned through the Rescorla-Wagner rule [44] (for details see **Methods** and **Figure 1**):

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

where $Q_{motive}[\mathrm{t}]$ is the agent's prediction on how future information will impact its 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 prediction, while $\delta_{motive}[t]$ is the current prediction error (for details see **Methods** and **Figure 1**). All motive expected values are then integrated through a value function of information:

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

where β_{affect}, $\beta_{cognition}$, and 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 use information-seeking to increase their affect (β_{affect}), reduce their uncertainty ($\beta_{cognition}$), and improve future choices (β_{action}) [43]. After estimating the joint value of information, the agent can use this value to direct its information-seeking decisions. This is formalized by entering V_{info} into a logistic probability function:

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

where $cost$ refers to the cost of seeking information and $P_{info-seek}$ indicates the probability of seeking the evaluated piece of information.

In the task, we represented the information hidden from the agent as three cards randomly selected from a six-card lottery (**Figure 2**). Cards from the lottery were drawn from an (integer) uniform distribution. On each trial, the agent decided whether to purchase the three cards or remain ignorant (**Figure 2**). This task design was intended to mimic real-life scenarios in which people visit information online or on social media and can 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 2**). If the agent decided to play the lottery, one card (i.e., the outcome card) was randomly chosen from the six-card lottery and the number printed on it was added to or removed from (depending on whether the outcome was positive or negative) 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 could help the agent maximize gains and reduce losses in the later choice stage (i.e., play the lottery or pass).

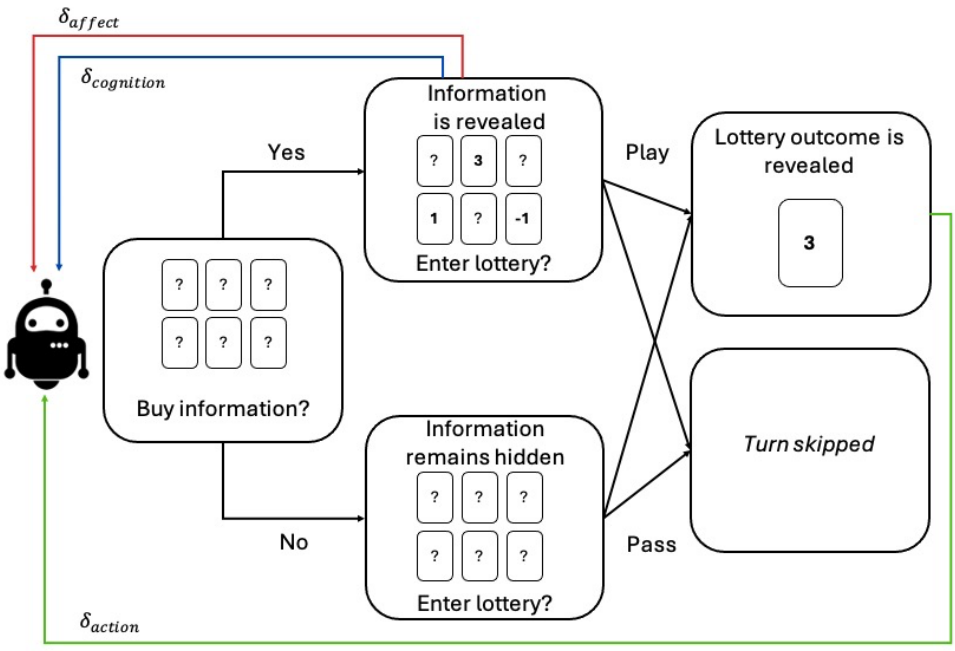


Figure 2. ***The Simulated environment.*** On each trial, a six-card lottery was presented to the agent. The agent could either decide to seek information ('Yes') or remain ignorant ('No'). If the information was sought, three cards were revealed to the agent, one of which was the outcome card (although the agent did not know which one this was). If the information was not sought, the three cards remained hidden to the agent. The revealed information was used to update the expected value of Affect and Cognition by integrating δ_{affect} and $\delta_{cognition}$ into those values. Next, the agent decided to either play the lottery or pass. If the agent played the lottery, a card was randomly drawn from the six-card lottery and added to the agent's pot of money. If the agent decided to pass, the turn was skipped and no value was added to the agent's pot of money. The outcome of such choice was integrated in the expected value of Action through δ_{action} solely if the agent decided to seek information in the previous step.

Simulation 1: The generalized RL model computationally accounts for previous findings in the information-seeking literature

We first asked whether our model could learn to seek information that helps it feel good, reduce uncertainty, and collect future rewards (in line with Affect, Cognition, and Action motives). To do so, we ran a set of simulations in a positively skewed environment where lottery cards were drawn from a uniform distribution within the range [-3, 5] (Simulation 1a). We simulated 50 agents playing with the information-seeking task described above for 100 trials. We set the initial values for $Q_{affect}[0]$, $Q_{cognition}[0]$ and $Q_{action}\ [0]$ to 0 and their relative weights β_{affect}, $\beta_{cognition}$, β_{action} to 1. The cost of information was set to $cost$ = 0.1, while the three learning rates α_{affect}, $\alpha_{cognition}$ and α_{action} were set to 0.5.

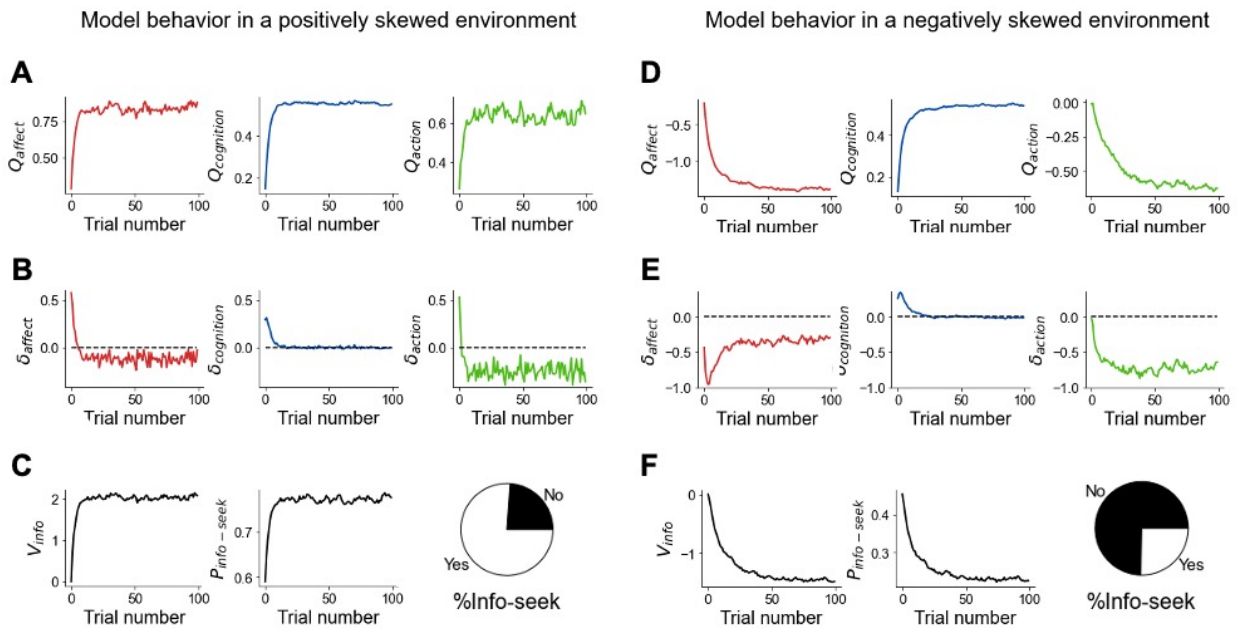


Figure 3. ***The impact of information is learned over time.*** **A**) Our model learned that information in the positively skewed environment increases affect, reduces uncertainty and leads to the selection of reward-yielding actions. **B**) As learning progresses, prediction errors for all Q values converge to zero **C**) This resulted in an increased probability of seeking information in the positively skewed environment. **D**) Our model learned that information in the negatively skewed environment decreased affect and did not always lead to rewarding actions. As information always reduces uncertainty in this simulated environment, $Q_{cognition}$ increases across trials. **E**) As learning progressed, prediction errors for all Q values converged towards zero. **F**) This resulted in a low probability of seeking information in this environment.

As shown in **Figure 3A**, the model learned that information positively impacted its affect, increased its knowledge about the hidden numbers, and helped to collect rewards during the lottery choice. This is shown by increased value for Affect, Cognition and Action across trials. As the model learned the true value of Affect, Cognition and Action, its prediction errors δ_{affect}, $\delta_{cognition}$ and δ_{action} converged to zero (**Figure 3B)**. In this simulation and the following ones, similar results were obtained by varying the learning rates for Affect, Cognition, and Action. As expected, smaller learning rates slowed down the learning process (**Figure S1**). In line with recent findings and theories on information-seeking [2,17,18,20,25], the model successfully learned the benefits of obtaining information, as shown by increased information value and probability to seek information throughout the task (**Figure 3C**). In other words, in this first positively skewed environment, the agent chose to seek information on most of the trials.

Next, we asked whether our model could also learn to avoid unhelpful, negatively charged information. In Simulation 1b, we therefore simulated 50 agents playing the information-seeking task for 100 trials using the above model parametrization, but in a negatively skewed environment where the lottery cards were drawn from a uniform distribution within the range [-5, 3]. As shown in the first panel of **Figure 3D,** the model learned that information negatively impacted its affect in most of the trials. Indeed, the value for Affect (Q_{affect}) decreased across trials. The model also learned to reduce the value of Action Q_{action} as, in this environment, information did not help to collect future rewards (**Figure 3D**). In stark contrast, the value of Cognition ($Q_{cognition}$) remained

positive, as even in this negatively skewed environment information-seeking reduced the agent's uncertainty. As before, prediction errors for all Q values converged to zero over time, showing that the agent learned to predict its environment (**Figure 3E**). Most importantly, the value of information, as well as the probability of information-seeking, decreased across trials (**Figure 3F**).

Taken together, these simulations suggest that our RL model can computationally account for findings in the literature which suggest that people seek information that makes them feel good and avoid information that makes them feel bad [1-3], reduces their uncertainty [34,36-40], and helps them select actions that will yield future rewards and avoid those that will yield future losses [23,31,41,42]. In the next section, we use this model to make predictions about doomscrolling.

Simulation 2: Predictions about doomscrolling

Simulations 1a and 1b suggest that the impact of information on Affect, Cognition, and Action can be learned over time and the agent's behavior can be adjusted depending on the type of information received. Here, we use the above model to dig into scenarios in which people decide to persistently seek negatively valenced information (i.e., doomscrolling).

As shown in Simulation 1b, where the environment was negatively skewed, the value of Cognition remained positive while the values of Affect and Action became negative (**Figure 3D**). This is because, in our simulated environment, information always reduces uncertainty, even if the degree to which it does varies across contexts, and even if the valence of information is negative. Here, we test whether simulating the model with higher weights on Cognition would result in persistently seeking negatively valenced information by overemphasizing the uncertainty-reduction aspect of information. To test this, we ran a set of simulations in the same environment of Simulation 1b (lottery cards were drawn from a negatively skewed uniform distribution with values in the [-5,3] range). The parametrization of the model was the same as in Simulation 1 except that the weight assigned to Cognition ($\beta_{cognition}$) was set to 1 and the relative weights assigned to Affect (β_{affect}) and Action (β_{action}) to 0.1. As expected, after running the model for 50 independent simulations (each consisting of 100 trials), we observed that the value of information increased over time (**Figure 4C**). This resulted in frequent information-seeking choices despite the information had negative consequences on Affect. Therefore, although information tended to be negatively valenced and did not lead to rewarding actions (as shown by decreased Q_{affect} and Q_{action}; **Figure 4A**), placing a higher weight on Cognition biased the model towards more information-seeking. Next, we ran the same simulations but varying β_{affect} , $\beta_{cognition}$, and β_{action} across a wide parameter space. As shown in **Figure 4 D**, lower values of β_{affect} and β_{action} compared to $\beta_{cognition}$ resulted int frequent information-seeking choices.

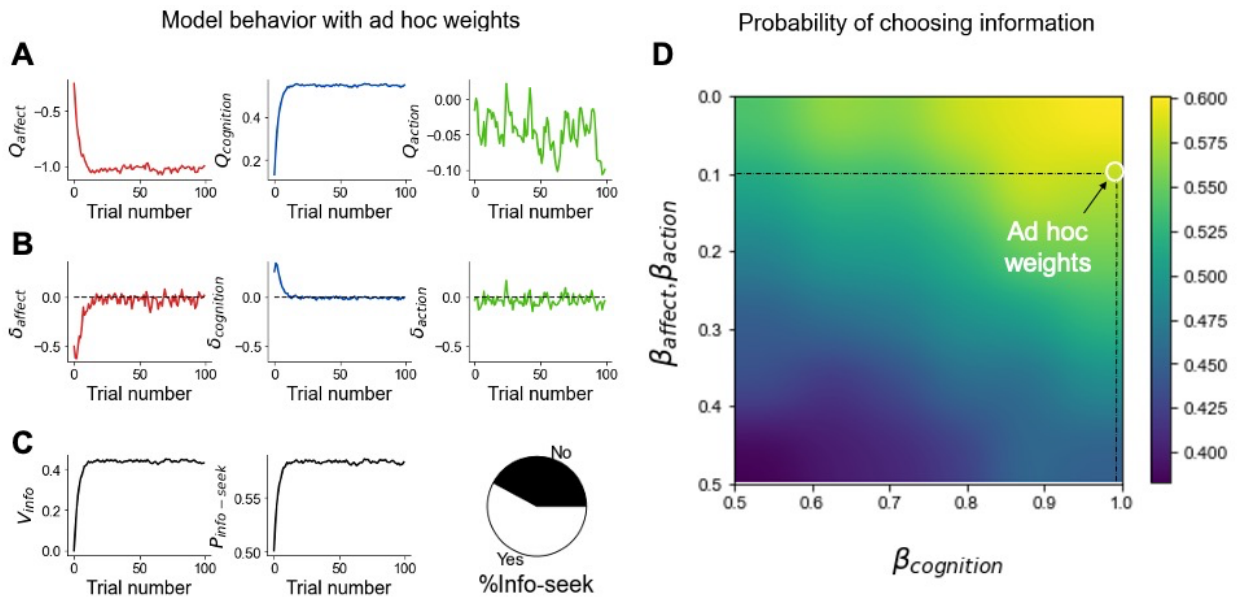


Figure 4. ***Predictions about doomscrolling.*** **A**) Information-seeking behavior and value learning for an agent withe β_{affect}=0.1, $\beta_{action} = 0.1$, and $\beta_{cognition} = 1$. The value for Affect and Action decreased across trials, while the value of Cognition increased. **B**) As learning progressed, prediction errors for all Q-values converged to zero. **C**) Despite the fact that the value for Affect and Action decreases across trials, the increased importance given to Cognition directs the agent to seek information. **D**) The probability to seek information as a function of different values for β_{affect} , $\beta_{cognition}$, β_{action}. In every simulation, $\beta_{cognition}$and β_{action} had the same value.

Taken together, our simulations show that doomscrolling can occur when $\beta_{cognition}$ is sufficiently high relative to β_{affect} and β_{action}. We refer to this configuration of weights as the "doomscrolling fingerprint". The greater the difference between β_{action} and β_{affect}, the more marked the fingerprint. In the next section, we check whether the doomscrolling fingerprint can be identified in human participants. We then compare the prominence of such a fingerprint with the consumption of negatively valenced information in an information-seeking task.

Study 1: A subset of human participants displays the doomscrolling fingerprint, which predicts consumption of negatively valenced information during an information-seeking task

Based on the simulations described above, our model predicts that doomscrolling arises when $\beta_{cognition}$ is sufficiently high relative to β_{affect} and β_{action}. Here, we sought to identify the doomscrolling fingerprint and predict consumption of negatively valenced information in human participants playing an information-seeking task. To do so, we obtained data from a previously published study [7] in which, on each trial, participants decided whether to purchase information about a lottery or remain ignorant about the most likely future outcomes (**Methods**). In the original study, a linear regression model predicted each participants' information-seeking decisions from objective measures of Affect, Cognition, and Action – thus yielding standardized coefficients for each factor. In the following analyses, we interpret such coefficients as the weights participants assigned to the three factors and use them to identify patterns of weight value-motive combinations.

First, we checked whether the doomscrolling fingerprint was present in the real population by running a cluster analysis (**Methods**) on these participants' weights. The analysis revealed two clusters of participants (**Figure 5A, B**). The *Cognition Cluster* was characterized by high weights on Cognition and low weights on Affect (average β_{affect} = 0.028, average $\beta_{cognition}$ = 0.619, average β_{action} = 0.321). On the contrary, the *Affect Cluster* was characterized by high weights on Affect and low weights on Cognition (averaged β_{affect} = 0.978, averaged $\beta_{cognition}$ = 0.124, averaged β_{action} = 0.306). We also obtained similar results when analyzing the data from a replication of the above information-seeking study [7] (**Supplementary**). These results suggest that a subset of people, namely the Cognition Cluster, display a doomscrolling fingerprint.

As a further validation step, we simulated our model and task with the averaged weights for Affect, Cognition, and Action associated with each cluster. Both simulations were run in the same environment as Simulation 1b (lottery cards were drawn from a negatively skewed uniform distribution with values ranging between [-5,3]). As shown in **Figure 5C**, the model sought more negatively valenced information when simulated with the weights obtained from the Cognition Cluster than when simulated with the weights obtained from the Affect Cluster, essentially replicating what our previous model simulations suggested (**Figure 4**).

We then asked whether participants from the Cognition Cluster would seek more negatively valenced information in the original task, as our model predicts (**Method**). To show this, we first check whether participants from the Cognition Cluster were more likely to seek information with lower valence (regardless of whether its sign was positive or negative) compared to the Affect Cluster. Indeed, the Cognition Cluster (M = - 0.22; SD = 3.92) sought more information with lower valence compared to the Affect Cluster (M = 4.79, SD = 3.92; t(45) = -4.61, $p < 10^{-5}$). We obtained similar results when analyzing the data of the replication experiment (**Supplementary**). Next, we asked whether participants from the Cognition Cluster were more likely to seek negatively valenced information (therefore selecting only information with a negative sign). To do so, we compared the percentage of information-seeking decisions between the two clusters solely in trials in which the information had negative valence. Results revealed that participants from the Cognition Cluster sought more often negative valenced information (M = 51.4, SD = 21) compared to the Affect Cluster (M = 43.1, SD = 19.4; t(47)= 1.5, p = 0.1; **Figure 5D**) as predicted by our model simulations (**Figure 5B** and **C**). Although this result was only a tendency, we obtained a statistically significant result in the replication experiment with a larger number of participants (**Supplementary**).

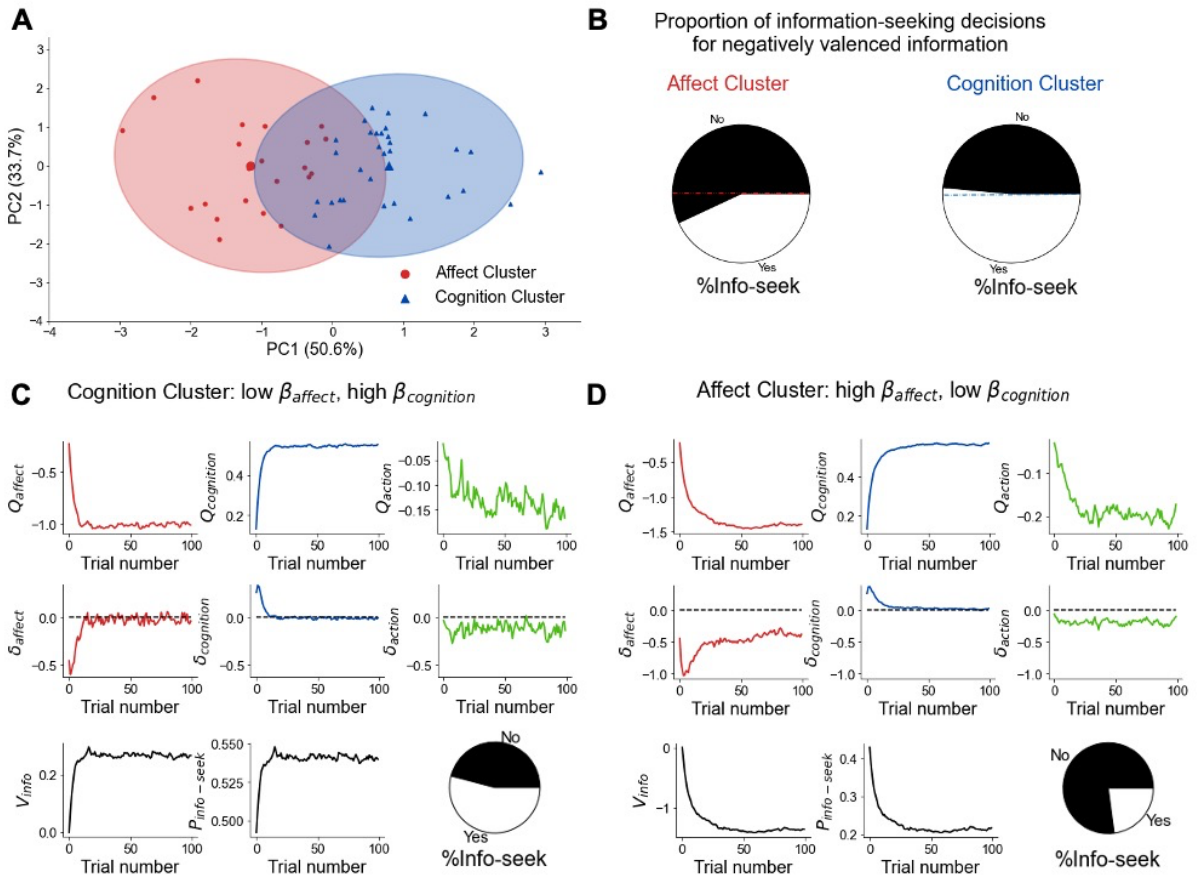


Figure 5. ***Doomscrolling fingerprint in human participants.*** **A**) The best two clusters (explained variance in parentheses) plotted by projecting their values into a 2-D space. **B**) Human participants' information-seeking decisions for negatively valanced information plotted separately for each cluster. **C**) Simulations of the model using the average weights for the Cognition Cluster. The model extensively sought information in spite of the negative impact on Affect. **D**) Simulations of the model using the averaged weights for the Affect Cluster. The model sought negative information less often compared to the simulations with the Cognition Cluster.

In sum, these results show that a subset of human participants displays a doomscrolling fingerprint, as identified by the Cognition Cluster above. As predicted by our model, the Cognition Cluster sought more negatively valenced information when performing an information-seeking task. In the next section, we test whether the doomscrolling fingerprint also predicts doomscrolling in real life.

Study 2: The doomscrolling fingerprint predicts doomscrolling in real life

To test whether the doomscrolling fingerprint predicts real-life doomscrolling, we collected data from a larger follow-up survey in which the same participants who played with the information-seeking task were asked to fill out questionnaires about doomscrolling, general information-seeking tendency, news consumption, and social media consumption, among others (**Methods**). As the response rate to the follow-up survey only reached roughly 50% across the main experiment (N = 28) and its replication (N = 67), and because experimental conditions were identical in the two subsets, we ran the subsequent analyses by combining the two datasets (N_{total} = 95).

For each participant, we define the prominence of their doomscrolling fingerprint as the difference between the weights assigned to Cognition and the weights people assigned to Affect while performing the original information-seeking task ($\beta_{cognition}$ - β_{affect}). As predicted by our model (**Figure 4D**), the higher the doomscrolling fingerprint the more participants reported a high Doomscrolling Score (r = 0.22, p = 0.032; **Figure 6A**), obtained from the responses to a doomscrolling scale (see **Methods**). The doomscrolling fingerprint did not correlate with the Information Consumption Score (r = -0.1, p = 0.22), obtained from a general information-seeking tendency scale (see **Methods**). However, both the Doomscrolling Score (r = 0.26, p = 0.012) and the Information Consumption Score (r = 0.24, p = 0.02) positively correlated with the weights participants assigned to Cognition. That is, the higher the weights assigned to Cognition, the more people engage in real-life doomscrolling and often get informed through the Internet and social media.

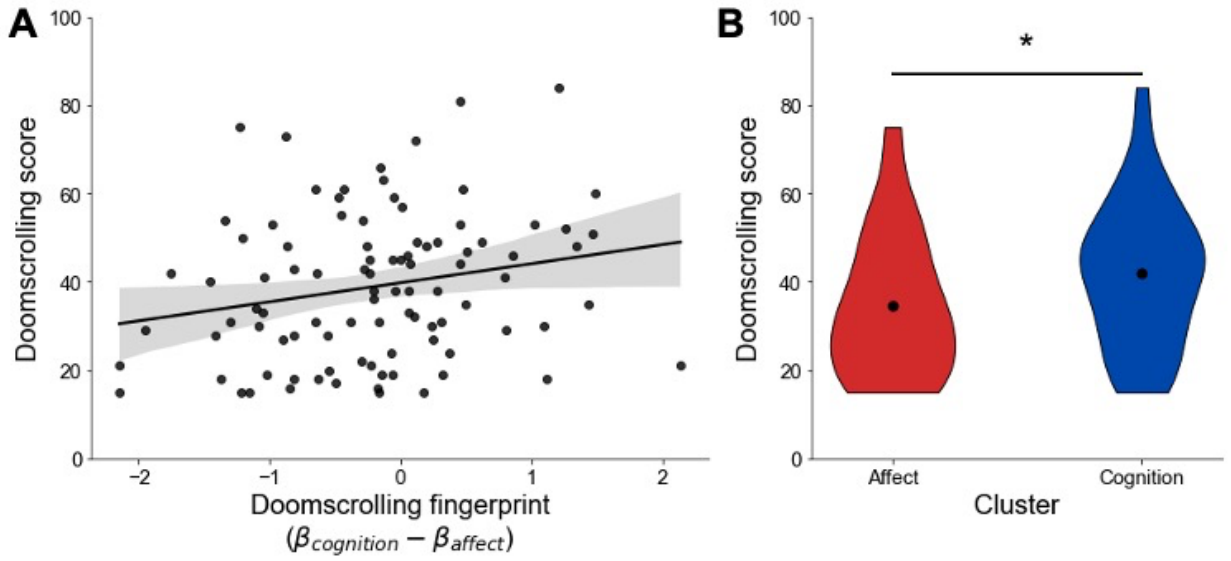


Figure 6. ***The doomscrolling fingerprint predict doomscrolling in real life.*** **A**) The doomscrolling Score correlated positively with the doomscrolling fingerprint computed as the difference between the weights assigned to Cognition and the weights assigned to Affect ($\beta_{cognition}$ - β_{affect}). In other words, the more prominent the doomscrolling fingerprint the greater the incidence of doomscrolling participants reported perceiving in real life. **B**) The Cognition Cluster reported to doomscroll more compared to the Affect Cluster. *: $p < 0.05$

Subsequently, we compared participants' responses on the doomscrolling scale between the Cognition cluster and the Affect cluster. As predicted by our model, participants from the Cognition cluster reported higher Doomscrolling Scores (M = 41.9, SD = 16.3) compared to the Affect cluster (M = 34.5, SD = 16.2; t(80)= 2.2, p = 0.033; **Figure 6B**). We did not observe differences in average daily consumption of information as well as average daily use of social media between the two clusters (both p > 0.5).

Taken together, these results suggest that individuals' doomscrolling fingerprints (i.e., their tendency to focus on uncertainty-reducing information above and beyond positively valenced information) predicts real-life doomscrolling. In particular, the higher the difference between the

weights assigned to Cognition and those assigned to Affect, the greater the reported incidence of doomscrolling in participants' reports. Moreover, participants from the Cognition Cluster reported more real-life doomscrolling compared to the Affect Cluster despite not differing in overall news consumption and social media use.

Discussion

In the contemporary society, people have unprecedented access to vast amounts of information, and each piece of information can influence their wellbeing and mental health. Some people selectively avoid distressing content, such as news about wars, pandemics, or environmental disasters. In contrast, others persistently immerse themselves in such information, even to the detriment of their well-being and mental health. The latter phenomenon is known as doomscrolling. To shed light on the origins of this maladaptive information-seeking behavior, we have introduced a novel computational model that integrates recent findings and current theories of human information-seeking into the RL framework and we have empirically tested its predictions. Overall, our model simulations suggest that assigning higher weights on the agent's prediction on how future information will impact its understanding of the environment in which it acts (Cognition), relative to the agent's prediction on how future information will impact its affective states (Affect) can lead to doomscrolling. Using data from an information-seeking task and a follow-up survey, we found evidence for these predictions in the real population.

First, we generalized the RL framework to account for new theories and findings from the information-seeking literature. Within this extended framework, information can act as a reinforcer to determine the occurrence and frequency of information-seeking in the future [18,25]. Based on the findings that multiple motivational factors jointly motivate people to seek information [2,3,6,7,23,33-35], we developed an RL agent that seeks to maximize the weighted sum of the expected value of these factors. By doing so, our model was able to computationally account for findings suggesting that people seek information that makes them feel good and avoid information that makes them feel bad [1-3], reduces their uncertainty [34,36-40], and helps them to obtain rewards in the future [23,31,41,42]. These simulations suggest that RL algorithms can be generalized to maximize other variables besides the expected reward value – echoing previous attempts of extending RL algorithms beyond reward-learning including working memory [45], effort investments [46], and intrinsic motivation [47].

Next, through simulations and analyses of empirical data, we shed light on the potential computational mechanisms behind doomscrolling. Our model simulations suggest that attributing higher weights to Cognition and lower weights to Affect in the value of information may lead to doomscrolling. We refer to this configuration of weights as the doomscrolling fingerprint. By analyzing data from an information-seeking task [7], we found evidence that this fingerprint can be found in a subset of human participants – the "Cognition Cluster" . By contrast, the remaining group of participants – the Affect Cluster – assigned higher weights to Affect relative to Cognition. Such individual differences in how people display interest for negative content are consistent with findings in the literature [2,48-50]. For example, some people are highly curious about images of death and violence [48,49] while others explicitly avoid negative information [43,51].

Subsequent analyses also showed that participants from the Cognition Cluster sought negatively valenced information more than the Affect Cluster when performing the information-seeking task. These results suggest that variations in how people weigh the motivational factors can lead to doomscrolling. We tested this hypothesis by analyzing data from a follow-up survey. Results showed that the higher the doomscrolling fingerprint the more participants engaged in real-life doomscrolling. Interestingly, we measure the doomscrolling fingerprint by using data from a very simple information-seeking task [7], and nonetheless, it was a valid predictor of a complex behavior such as doomscrolling. Moreover, as predicted by our model, participants from the Cognition Cluster reported greater engagement in real-life doomscrolling compared to the Affect Cluster. In line with these findings on individual differences, recent studies showed that variations in the preference for reducing uncertainty about negatively valenced information are reflected in the monkey brain [52]. In particular, monkeys that prefer to reduce uncertainty about negative information, compared to those that do only reduce uncertainty for positive information, have specific populations of neurons in the ventrolateral Prefrontal Cortex and Anterior Cingulate Cortex that anticipate the reduction of uncertainty about negative information. On the contrary, monkeys who do not prefer to reduce uncertainty about negative information do not have such neural anticipation. Future research is needed to understand whether individual differences in doomscrolling are also reflected in the human brain.

However, our study also prompts further questions. For example, throughout the study, we focus on individual variation in the weights given to the information-seeking motivational factors. However, changes in the weights might occur as a function of contexts, internal states or a mix of the two. This raises the further question of whether people have fixed (trait-like) or variable (state-dependent) weights. Moreover, the question of whether and how people learn these weights remains. Indeed, in our model, we assumed that people learn only the value of each motivational factor, while the weights assigned to them were manually set in the model. However, at a hierarchically higher level, it is possible that also the weights are learned across time using reinforcement learning mechanisms. If so, can people re-learn their weights to avoid maladaptive information-seeking? For example, can people learn to reduce the weight assigned to Cognition and increased the weight on Affect to avoid persistently seeking negative valenced information? Providing answers to these questions could help develop future interventions to prevent the negative consequences of maladaptive information-seeking on people's well-being and mental health.

Taken together, by integrating recent findings and current theories of human information-seeking into the RL framework we shed light on the origins of maladaptive information-seeking behaviors. We showed that doomscrolling arises when the motivation to reduce uncertainty is higher relative to the motivation to improve affect through information-seeking. Our study takes a significant step forward in unraveling the intricate psychological mechanisms behind maladaptive information-seeking behaviors prevalent in today's digital age, potentially paving the way for future behavioral interventions.

Methods

The computational model

To select the information that maximizes the linear combination of Affect, Cognition, and Action (i.e., the value of information; V_{info}), the agent has to first learn their values (Q_{affect}, $Q_{cognition}$, Q_{action}; **Figure 1**). In our model, Q_{affect} refers to the agent's prediction on how future information would impact its affective states; $Q_{cognition}$ refers to the agent's prediction on how future information would impact its understanding of the environment in which it acts; and Q_{action} refers to the agent's prediction on how future information would impact the selection of reward-bearing actions. The initial values of Q_{affect}, $Q_{cognition}$, and Q_{action} are initialized to zero and updated recursively using the Rescorla-Wagner rule [44]. Here and across the manuscript we use this simplest formulation of RL as it has been proved to explain many learning and decision-making mechanisms [15,31,53], and it has a biological validity [54,55]. However, we note that more complex RL algorithms can be integrated into our model to account for behavioral complexity (e.g., model-based algorithms; Sutton & Barto, 2018).

Q_{affect} is updated as follows:

$$Q_{affect}[\mathrm{t}] = Q_{affect}[t-1] + \alpha_{affect} \times \delta_{affect}[t] \quad (4)$$

where $Q_{affect}[t]$ is the previous prediction on how future information would impact the agent's affective states. The parameter α_{affect} is the learning rate at which the agent integrates its prediction error δ_{affect} in the Affect value function (**Figure 1**). δ_{affect} is defined as the difference between the agent's affect after the information is received and the previous prediction such that:

$$\delta_{affect}[t] = Info_{valence}[\mathrm{t}] - Q_{affect}[t-1] \quad (5)$$

$Info_{valence}$ is the valence of the information (i.e., whether the information is positive or negative). For example, in the simulated environment (**Figure 2**) $Info_{valence}$ of the information feedback corresponds to the sum of the product of each number provided to the agent and its occurrence probability (see below for more details). We have already shown that people's affect during information-seeking correlates with a such computation of $Info_{valence}$ [7].

$Q_{cognition}$ is updated as follows:

$$Q_{cognition}[\mathrm{t}] = Q_{cognition}[t-1] + \alpha_{cognition} \times \delta_{cognition}[t] \quad (6)$$

where $Q_{cognition}[t]$ is the previous prediction on how future information would impact the agent's understanding of the environment in which it acts. $\alpha_{cognition}$ is the learning rate at which the agent integrates its prediction error $\delta_{cognition}$ in the Cognition value function (**Figure 1**). $\delta_{cognition}$ is defined as the difference between the agent's uncertainty after the information is received and the previous prediction such that:

$$\delta_{cognition}[\mathrm{t}] = Info_{uncertainty} - Q_{cognition}[t-1] \quad (7)$$

where $Info_{uncertainty}$ is the inverse of the standard deviation (i.e., $1/SD_{info}$) and it captures the agent's uncertainty after the information is received. SD_{info} is computed across lottery cards. People's uncertainty is often formalized as SD [56-58] and we have recently shown that this formulation relates to people's subjective uncertainty [7]. Here, we entered the inverse of SD in the calculation of $\delta_{cognition}$ as high uncertainty corresponds to higher values of SD, and low uncertainty corresponds to lower values. We note, however, that other formulations can be used such as entropy (which is (a linear function of) the log of the SD in case of a Gaussian distribution) or precision (which is the inverse of variance) [59,60].

Lastly, Q_{action} is updated as follows:

$$Q_{action}[\mathrm{t}] = Q_{action}[t-1] + \alpha_{action} \times \delta_{action}[t] \quad (8)$$

where $Q_{action}[t]$ is the previous prediction on how future information would impact the selection of reward-bearing actions, and α_{action} is the learning rate at which the agent integrates its prediction error δ_{action} into the Action value function (**Figure 1**). δ_{action} is defined as the difference between how information impacts the outcome of future choices and the previous prediction such that:

$$\delta_{action}[\mathrm{t}] = Info_{outcome\ action}[t] - Q_{action}[t-1] \quad (9)$$

where $Info_{outcome\ action}$ is the outcome obtained from a choice made if the information is received. $Info_{outcome\ action}$ links Q_{action} to the future choice made by the agent with the help of information. We note that in our formalization the instrumental utility of the information is computed within the update of the Action value function. This means that if the information helps to collect rewards during future choices, the prediction error is positive and Q_{action} increases. On the contrary, if information does not yield future rewards, the prediction error is negative and Q_{action} decreases.

Next, the above values are combined into the information value (V_{info}) through the following function:

$$V_{info}[\mathrm{t}] = \beta_{affect} \times Q_{affect}[\mathrm{t}] + \beta_{cognition} \times Q_{cognition}[t] + \beta_{action} \times Q_{action}[t] \quad (10)$$

where β_{affect}, $\beta_{cognition}$, β_{action} are the relative weights given to Affect, Cognition and Action values, respectively. These weights determine to what extent people are motivated to use information-seeking to increase affect, reduce uncertainty and improve future choices [43].

After estimating the joint value of information, the agent can use it to direct its information-seeking choices. This is formalized by entering V_{info} into a logistic probability function as follows:

$$P_{info-seek} = \frac{1}{1+\exp\left(cost - V_{info}\right)} \quad (11)$$

where the $cost$ term captures the many types of costs people incur when seeking information [61,62] – whether it is money (e.g., to purchase a journal subscription), time (spent on finding and consuming the information), or physical/cognitive effort (needed to access and consume the information). Moreover, we used the above formulation to mimic non-deterministic choices, in accordance with many studies in the literature describing people's choices between available options as a logistic function (or its generalization: the softmax function; [42,63,64]. We note that a logistic noise (i.e., temperature) parameter could also be added to the function.

The simulated environment

The simulated environment consisted of a set of trials in which the RL agent chose whether to seek or avoid information (**Figure 2**). The task's trial-by-trial dynamics allows the specified model to learn the ensemble of Affect, Cognition, and Action values (see below). We represented the information hidden from the agent as three cards drawn from a six-card lottery (**Figure 2**). If, on a given trial, the agent decided to receive information, it would pay a fixed cost to see three out of six cards. To include an instrumental component in this task, at the end of each trial the agent had the choice to play with the lottery or pass (**Figure 2**). If the agent decided to play the lottery, one card was randomly chosen from the six-card lottery and the number printed on the card (which could be positive or negative) determined the outcome of the trial (i.e., the outcome card) which was added to the agent's pot of money. If the agent decided to pass, 0 was added to the agent's pot. As information contains the outcome cards and two decoy cards, retrieving information about the cards would help the agent maximize gains and reduce losses in the later decision stage (i.e., play the lottery or pass).

The six-card lottery was drawn from an (integer) uniform distribution, whose range was fixed in each set of simulations. We manipulated this range to create different types of environments in which the simulated agent could learn. For example, the range of the distribution was set to [-3, 5] to create a positively skewed environment. Conversely, the range was set to [-5, 3] to create a negatively skewed environment. In the positively skewed environment, if the agent decided to obtain information, its valence was positive in most of the trials. In the second negatively skewed environment, instead, if the agent decided to obtain information, its valence was usually negative. In both environments, the information always reduced the model's uncertainty about the hidden numbers. Lastly, as the distribution underlying the numbers also determined the outcome card, in the first environment information improved future choices most of the time, while in the second the information did not improve future choices most of the time.

To update Q_{affect}, the agent computes $Info_{valence}$ from Eq. 5 as follows:

$$Info_{valence}\,[\mathrm{t}] = \sum_{i=1}^{3} c_i \times p_i \quad (12)$$

where c_i is the value of the number *i* displayed on a card at trial *t* and p is the probability that the card with number *i* will be the outcome of the lottery. Note that here, p = $\frac{1}{3}$ on every trial *t* in which information was sought.

To update $Q_{cognition}$, $Info_{uncertainty}$ from Eq. 7 is computed as follows:

$$Info_{uncertainty}[t] = 1/SD_{info}[t] = \frac{1}{\sqrt{\frac{\sum_{i=1}^{3}(c_i-\mu)^2}{N}}} \quad (13)$$

where c_i is the value of the card *i* displayed at trial *t,* μ is the mean of the 3 information cards and *N* is their total number (i.e., N=3).

Q_{action} was only updated on trials in which the agent decided to seek information. The outcome of the play/pass choice was used to compute $Info_{outcome\ action}$ from Eq. 9. Each time the agent passed, the $Info_{outcome\ action}$ was set to zero, and each time the agent entered the lottery the $Info_{outcome\ action}$ was equal to the number printed on the outcome card. On trials in which the agent sought information, the play/pass choice (P_{choice}) was determined by entering the $Info_{valence}$ into a probability function as follows:

$$P_{choice} = \frac{1}{1+\exp(cost-Info_{valence})} \quad (14)$$

where $cost$ refers to the cost of entering the lottery (which was fixed and known to the agent). If the information was avoided, the agent made the lottery choice at random (i.e., 50% of the times it will choose to play), as none of the cards were displayed. In this later case, however, the outcome of such choice was not used to update Q_{action}.

Human data: information-seeking task

To obtain the information-seeking weights (β_{affect}, $\beta_{cognition}$, β_{action}), we used a previously collected dataset (available at https://github.com/affective-brain-lab/People-adaptively-use-information-to-improve-their-internal-states-and-external-outcomes; [7]. In particular, we used data from two independent experiments (N = 54 and N = 136, respectively). The second experiment was a replication of the first experiment. In these experiments, on each of 90 trials, a lottery depicted as 6 cards was shown to participants. Each card had a number on it, indicating the magnitude of associated loss or gain. Participants could decide to purchase advance information, where 3 cards from the lottery were shown to them – the outcome card plus two decoys. Prior to committing to an information-seeking decision, participants were also asked to report their subjective ratings for Affect, Cognition, and Action. At the end of each trial, participants were presented with a choice to play the lottery or pass. If participants decided to play, the outcome card determined the payoff on that trial – either a loss or a gain.

Human data: Cluster analysis

In the above task, participants' information-seeking was best described by a linear regression model with the subjective ratings for Affect, Cognition, and Action as predictors. We obtained the weights for Affect, Cognition, and Action from the model fit and entered them into our cluster

analysis. We used the NbClust function in R [65] to compute the clusters which best grouped the weights for Affect, Cognition, and Action. The NbClust function allows to test 30 clustering indices (e.g., varying all combinations of number of clusters, distance measures and clustering methods) to determine the best number of clusters in a given dataset, and it returns the cluster number best explained by the majority of these indices.

Human data: valence of information

We computed the valence of information in the information-seeking task similarly as described in Eq. 12. However, in this case we considered the value of all the six cards shown to participants as follows:

$$Valence\ of\ Information\ [\mathrm{t}] = \sum_{i=1}^{6} c_i \times p_i \quad (15)$$

where c_i is the value of the number *i* displayed on a card at trial *t* and p is the probability that the card with number *i* will be the outcome of the lottery. Note that here, p = $\frac{1}{6}$ in every trial *t*. To compute the percentage of information-seeking decisions with negative valence, we selected trials with valence < 0 and computed the percentage of information-seeking decisions.

Human data: surveys

Participants from the two experiments were contacted again and asked to fill out a follow-up survey. In particular, participants were presented with questions about their daily information consumption and psychological feelings related to that. Here, we report and analyze only 4 questionnaires of the survey:

1) Doomscrolling scale – this is a 15-item questionnaire on participants' general consumption of negatively valenced information (e.g., "I feel the urge to seek bad news/content online, more and more often"; [8].
2) General information-seeking tendency scale – this is an adaptation of the above Doomscrolling scale. In particular, participants were presented with 12 items from the above scale where the word "bad" was removed (e.g., I feel the urge to seek news/content online, more and more often"), thus measuring people's tendency to seek information overall, rather than specifically negatively valenced information.
3) Social media use – participants were asked "On average, how many hours per day do you usually spend on social media (excluding text messaging)?". They provided their answers on a 5-point Likert scale ("<30 minutes", "30 minutes to 2 hours", "2 to 4 hours", "4 to 6 hours", "> 6 hours").
4) News consumption – participants were asked "On average, how many hours do you usually spend reading/listening news/content (including TV, radio, newspaper, online news, social media)?". They provided their answers on a 5-point Likert scale ("<30 minutes", "30 minutes to 2 hours", "2 to 4 hours", "4 to 6 hours", "> 6 hours").

Responses to the doomscrolling scale were summed for each participant, returning a Doomscrolling Score for each individual. Similarly, answers to the general information-seeking tendency were summed for each participant returning an Information Consumption Score.

N = 28 participants from the main experiment and N = 67 participants from the replication experiment filled out the survey resulting in a response rate of 50% in both experiments. As experimental conditions were exactly the same in both experiments, we decided to combine responses in all the analyses reported in the manuscript.

Acknowledgments

I.C.D. was funded by the FWO junior postdoctoral fellowship.

References

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

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

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

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

5 van Lieshout, L. L. F., de Lange, F. P. & Cools, R. Uncertainty increases curiosity, but decreases happiness. *Sci Rep* **11**, 14014, doi:10.1038/s41598-021-93464-6 (2021).

6 Kelly, C. A. & Sharot, T. Individual differences in information-seeking. *Nat Commun* **12**, 7062, doi:10.1038/s41467-021-27046-5 (2021).

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

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

9 Silver, R. C. *et al.* Mental- and physical-health effects of acute exposure to media images of the September 11, 2001, attacks and the Iraq War. *Psychol Sci* **24**, 1623-1634, doi:10.1177/0956797612460406 (2013).

10 Thompson, R. R., Jones, N. M., Holman, E. A. & Silver, R. C. Media exposure to mass violence events can fuel a cycle of distress. *Sci Adv* **5**, eaav3502, doi:10.1126/sciadv.aav3502 (2019).

11 Garfin, D. R., Silver, R. C. & Holman, E. A. The novel coronavirus (COVID-2019) outbreak: Amplification of public health consequences by media exposure. *Health Psychol* **39**, 355-357, doi:10.1037/hea0000875 (2020).

12 Zhao, N. & Zhou, G. Social Media Use and Mental Health during the COVID-19 Pandemic: Moderator Role of Disaster Stressor and Mediator Role of Negative Affect. *Appl Psychol Health Well Being* **12**, 1019-1038, doi:10.1111/aphw.12226 (2020).

13 Buchanan, K., Aknin, L. B., Lotun, S. & Sandstrom, G. M. Brief exposure to social media during the COVID-19 pandemic: Doom-scrolling has negative emotional consequences, but kindness-scrolling does not. *PLoS One* **16**, e0257728, doi:10.1371/journal.pone.0257728 (2021).

14 Price, M. *et al.* Doomscrolling during COVID-19: The negative association between daily social and traditional media consumption and mental health symptoms during the COVID-19 pandemic. *Psychol Trauma*, doi:10.1037/tra0001202 (2022).

15 Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction*. Second edn, (MIT Press, 2018).

16 Gottlieb, J., Oudeyer, P. Y., Lopes, M. & Baranes, A. Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends Cogn Sci* **17**, 585-593, doi:10.1016/j.tics.2013.09.001 (2013).
17 FitzGibbon, L., Lau, J. K. L. & Murayama, K. The seductive lure of curiosity: information as a motivationally salient reward. *Current Opinion in Behavioral Sciences* (2020).
18 Murayama, K. A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic-extrinsic rewards. *Psychol Rev* **129**, 175-198, doi:10.1037/rev0000349 (2022).
19 Bromberg-Martin, E. S. & Hikosaka, O. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* **63**, 119-126, doi:10.1016/j.neuron.2009.06.009 (2009).
20 Bromberg-Martin, E. S. & Hikosaka, O. Lateral habenula neurons signal errors in the prediction of reward information. *Nat Neurosci* **14**, 1209-1216, doi:10.1038/nn.2902 (2011).
21 Blanchard, T. C., Hayden, B. Y. & Bromberg-Martin, E. S. Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* **85**, 602-614, doi:10.1016/j.neuron.2014.12.050 (2015).
22 Ligneul, R., Mermillod, M. & Morisseau, T. From relief to surprise: Dual control of epistemic curiosity in the human brain. *Neuroimage* **181**, 490-500, doi:10.1016/j.neuroimage.2018.07.038 (2018).
23 Kobayashi, K. & Hsu, M. Common neural code for reward and information value. *Proc Natl Acad Sci U S A* **116**, 13061-13066, doi:10.1073/pnas.1820145116 (2019).
24 Lau, J. K. L., Ozono, H., Kuratomi, K., Komiya, A. & Murayama, K. Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. *Nat Hum Behav* **4**, 531-543, doi:10.1038/s41562-020-0848-3 (2020).
25 Marvin, C. B. & Shohamy, D. Curiosity and reward: Valence predicts choice and information prediction errors enhance learning. *J Exp Psychol Gen* **145**, 266-272, doi:10.1037/xge0000140 (2016).
26 Bennett, D., Bode, S., Brydevall, M., Warren, H. & Murawski, C. Intrinsic Valuation of Information in Decision Making under Uncertainty. *PLoS Comput Biol* **12**, e1005020, doi:10.1371/journal.pcbi.1005020 (2016).
27 Baranes, A. & Oudeyer, P.-Y. Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robotics and Autonomous Systems* **61** (2013).
28 Iigaya, K., Story, G. W., Kurth-Nelson, Z., Dolan, R. J. & Dayan, P. The modulation of savouring by prediction error and its effects on choice. *Elife* **5**, doi:10.7554/eLife.13747 (2016).
29 Santucci, V. G., Oudeyer, P. Y., Barto, A. & Baldassarre, G. Intrinsically motivated open-ended learning in autonomous robots (editorial). *Frontiers in Neurorobotics* **13**, 115 (2020).
30 Daddaoua, N., Lopes, M. & Gottlieb, J. Intrinsically motivated oculomotor exploration guided by uncertainty reduction and conditioned reinforcement in non-human primates. *Sci Rep* **6**, 20202, doi:10.1038/srep20202 (2016).
31 Cogliati Dezza, I., Yu, A. J., Cleeremans, A. & Alexander, W. Learning the value of information and reward over time when solving exploration-exploitation problems. *Sci Rep* **7**, 16919, doi:10.1038/s41598-017-17237-w (2017).
32 Cogliati Dezza, I., Cleeremans, A. & Alexander, W. H. Independent and interacting value systems for reward and information in the human brain. *Elife* **11**, doi:10.7554/eLife.66358 (2022).
33 Molinaro, G., Cogliati Dezza, I., Buhler, S. K., Moutsiana, C. & Sharot, T. Multifaceted information-seeking motives in children. *Nat Commun* **14**, 5505, doi:10.1038/s41467-023-40971-x (2023).

34 van Lieshout, L. L. F., de Lange, F. P. & Cools, R. Why so curious? Quantifying mechanisms of information seeking. *Current Opinion in Behavioral Sciences* **35**, 112-117 (2020).

35 Golman, R., Loewenstein, G., Molnar, A. & Saccardo, S. The Demand for, and Avoidance of, Information. *Management Science*, doi:10.1287/mnsc.2021.4244 (2021).

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

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

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

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

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

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

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

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

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

45 O'Reilly, R. C. & Frank, M. J. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput* **18**, 283-328, doi:10.1162/089976606775093909 (2006).

46 Verguts, T., Vassena, E. & Silvetti, M. Adaptive effort investment in cognitive and physical tasks: a neurocomputational model. *Front Behav Neurosci* **9**, 57, doi:10.3389/fnbeh.2015.00057 (2015).

47 Singh, S., Barto, A. G. & Chentanez, N. Intrinsically motivated reinforcement learning. *Adv. Neural Inform. Process. Syst. 17* (2005).

48 Oosterwijk, S. Choosing the negative: A behavioral demonstration of morbid curiosity. *PLoS One* **12**, e0178399, doi:10.1371/journal.pone.0178399 (2017).

49 Niehoff, E. & Oosterwijk, S. To know, to feel, to share? Exploring the motives that drive curiosity for negative content. *Current Opinion in Behavioral Sciences* **35**, 56-61, doi:10.1016/j.cobeha.2020.07.012 (2020).

50 Oosterwijk, S., Snoek, L., Tekoppele, J., Engelbert, L. H. & Scholte, H. S. Choosing to view morbid information involves reward circuitry. *Sci Rep* **10**, 15291, doi:10.1038/s41598-020-71662-y (2020).

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

52 Jezzini, A., Bromberg-Martin, E. S., Trambaiolli, L. R., Haber, S. N. & Monosov, I. E. A prefrontal network integrates preferences for advance information about uncertain rewards and punishments. *Neuron* **109**, 2339-2352 e2335, doi:10.1016/j.neuron.2021.05.013 (2021).

53 Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876-879, doi:10.1038/nature04766 (2006).

54 Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593-1599 (1997).
55 Silvetti, M., Vassena, E., Abrahamse, E. & Verguts, T. Dorsal anterior cingulate-brainstem ensemble as a reinforcement meta-learner. *PLoS Comput Biol* **14**, e1006370, doi:10.1371/journal.pcbi.1006370 (2018).
56 Auer, P., Cesa-Bianchi, N. & Fischer, P. Finite-time analysis of the multiarmed bandit problem. *Machine Learning* **47**, 235-256 (2002).
57 Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D. & Meder, B. Generalization guides human exploration in vast decision spaces. *Nat Hum Behav* **2**, 915-924, doi:10.1038/s41562-018-0467-4 (2018).
58 Bromberg-Martin, E. S. *et al.* A neural mechanism for conserved value computations integrating information and rewards. *bioRxiv* (2022).
59 Friston, K. *et al.* Active inference and learning. *Neurosci Biobehav Rev* **68**, 862-879, doi:10.1016/j.neubiorev.2016.06.022 (2016).
60 Oudeyer, P. Y. & Kaplan, F. What is Intrinsic Motivation? A Typology of Computational Approaches. *Front Neurorobot* **1**, 6, doi:10.3389/neuro.12.006.2007 (2007).
61 Horan, M., Daddaoua, N. & Gottlieb, J. Parietal neurons encode information sampling based on decision uncertainty. *Nat Neurosci* **22**, 1327-1335, doi:10.1038/s41593-019-0440-1 (2019).
62 Gottlieb, J., Cohanpour, M., Li, Y., Singletary, N. & Zabeh, E. Curiosity, information demand and attentional priority. *Current Opinion in Behavioral Sciences* **35**, doi:10.1016/j.cobeha.2020.07.016 (2020).
63 Daw, N. D. & Doya, K. The computational neurobiology of learning and reward. *Curr Opin Neurobiol* **16**, 199-204, doi:10.1016/j.conb.2006.03.006 (2006).
64 Wilson, R. C. & Niv, Y. Inferring relevance in a changing world. *Front Hum Neurosci* **5**, 189, doi:10.3389/fnhum.2011.00189 (2011).
65 Package'NbClust' (R, 2022).

Supplementary Material

Simulations 1a, 1b with different learning rates

Here, we ran Simulations 1a and 1b with learning rates α_{affect}, $\alpha_{cognition}$ and α_{action} = 0.1. These simulations replicate simulations reported in the main text, with the only difference that the model took more time to learn the value of each motive (**Figure S1**).

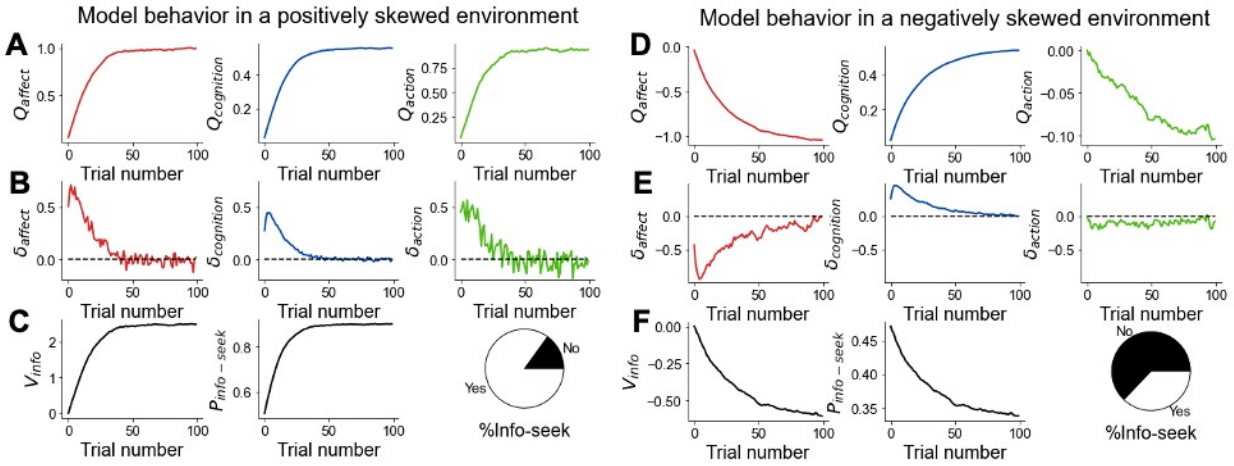


Figure S1 ***The impact of information is learned over time.*** **A**) Our model learned that information in a positively skewed environment increases affect, reduced uncertainty and leads to select rewarding actions. **B**) As the learning progresses, prediction errors for all Q values converged towards zero. **C**) This resulted in increased probability to seek information in this environment. **D**) Our model learned that information in a negatively skewed environment decreases affect and did not always lead to rewarding actions. As information always reduces uncertainty in the simulated environment, $Q_{cognition}$ increases across trials. **E**) As the learning progresses, prediction errors for all Q values converged towards zero. **F**) This resulted in decreased probability to seek information in such environment.

Analysis of human data in the replication experiment

We ran a cluster analysis to determine the number of clusters which best explained participants weights for Affect, Action and Cognition. As for the main experiment, we observed two clusters: the Cognition Cluster (averaged β_{affect} = 0.165, averaged $\beta_{cognition}$ = 0.467, averaged β_{action} = 0.288) and the Affect Cluster (averaged β_{affect} = 1.09, averaged $\beta_{cognition}$ = 0.078, averaged β_{action} = 0.284). Next, we asked whether participants from the Cognition Cluster were also more likely to seek lower valenced information. Overall, participants from such cluster sought information with lower valence (M = 0.452; SD = 3.83) compared to the Affect Cluster (M = 4.02, SD = 5.87; t(63)= -4.61, $p < 10^{-3}$). Then, we compared the percentage of information-seeking decisions solely in trials in which the information had negative valence. Results revealed that participants from the Cognition Cluster sought more often negative valenced information (M = 51.9, SD = 24.9) compared to the Affect Cluster (M = 41.3, SD = 20.6; t(104)= -2.65, p = 0.009) as predicted by our model simulations.