Learning From 'What Might Have Been': A Bayesian Model of Learning From Regret

Kate Petrova,* **James J. Gross, Tobias Gerstenberg**

Department of Psychology, Stanford University

*kpetrova@stanford.edu

Abstract

Regret is a common emotion that might either catalyze or impair decision-making. What determines whether regret will be helpful or harmful in a given situation? We test the hypothesis that regret is more likely to hinder decision-making during the early stages of learning, when information is limited, but help during later stages of learning, when the learner has a better understanding of the environment. We introduce a Bayesian model of learning from regret, in which the "counterfactual weight" parameter – reflecting how strongly individuals update their beliefs about foregone outcomes – predicts both learning outcomes and the intensity of subjective regret. We find that probing regret early in the learning phase leads to worse performance than probing regret later or not at all. This work has important implications for both cognitive and affective science, shedding light on the appraisal mechanisms by which regret influences decision-making.

Keywords: regret; emotion; affect; decision-making; Bayesian model

Introduction

Imagine standing in your favorite coffee shop, torn between trying the new seasonal latte or sticking with your tried-and-true regular coffee. You decide to branch out – only to discover the new latte is not to your liking. As you leave, you think, "if only I had ordered my usual coffee, I'd be much happier now." This is an example of regret – an emotion we feel when we realize that the outcome would have been better had we chosen differently (Zeelenberg et al., 2000).

Regret is a negatively valenced emotion characterized by three core components: (1) a representation of the actual outcome of one's action; (2) an appraisal of personal responsibility for the action that led to the actual outcome; and (3) a representation of a more desirable counterfactual outcome that would have resulted from an alternative action. A comparative evaluation of the actual and counterfactual outcomes is what elicits the distinct affective state of regret (Dijk & Zeelenberg, 2005; Zeelenberg et al., 2000).

The impact of regret on decision-making is double-edged (Inman, 2007). On the one hand, regret can serve an *adaptive* function by motivating individuals to learn from past mistakes and improve future decisions (O'Connor et al., 2014; Roese, 1997). In this way, regret fits into the broader framework of "affect as information," which argues that people use their own emotional experiences as vital sources of information in decision-making (Clore et al., 2012). For instance, regretting a poor investment decision may encourage more thorough research and risk assessment in future investments (Reb, 2008; Shani & Zeelenberg, 2007). On the other hand, regret can sometimes be *maladaptive*, leading people to avoid feedback, be excessively risk-averse, or get stuck in patterns of rumination (Reb & Connolly, 2009; Sijtsema et al., 2022; Zeelenberg et al., 1996).

The timing of regret

Crucially, whether regret acts as a beneficial or detrimental force may depend on *when* it arises. Early in learning, when information is scarce and uncertainty is high, experiencing regret might lead to impaired decision-making. In contrast, later in learning, when more information has accumulated and decisions are better informed, regret could have a corrective influence. One reason why regret might exert a negative influence on decision-making early in the learning process is because, without having sufficient information about the values of different choices, people are likely to imagine inaccurate counterfactual outcomes. For example, a customer visiting a new coffee shop for the first time might regret trying a new seasonal latte, imagining that a different drink would have tasted better. However, this counterfactual reasoning may be based on limited experience with the shop's full range of offerings rather than on a well-informed evaluation. Inaccurate counterfactuals can, in turn, lead to the formation of erroneous beliefs about the world, undermining future decision-making. This can lead to over-correction and stifle exploration in a novel environment. Thus, a customer who erroneously concludes that a black coffee is always superior may develop a bias against exploring the full menu and miss opportunities to discover potentially better options.

Broadly consistent with this view, past research shows that individuals who second-guess their initial judgments in high-uncertainty environments (which may be prompted by counterfactual thinking inherent in regret) show increased over-correction, which slows down their learning (Coali et al., 2022). Several studies also show that people frequently forego exploration as a way to minimize future regret (Bjälkebring et al., 2016; Coricelli et al., 2005). Early in the learning process, such regret minimization strategies might prevent optimal learning. In contrast, once the structure of the environment is better understood, this same regret-driven correction can become beneficial by reinforcing the exploitation of known high-value options.

To examine this temporal dimension of regret, our study

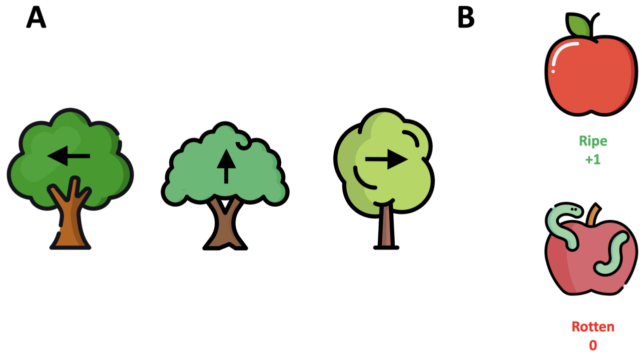

Figure 1: **Experimental procedure**. A) Participants select one of three trees on each trial using their keyboard. B) Participants can receive either a ripe fruit (+1 point) or a rotten fruit (0 points) after each pick. Participants only see the outcome of the chosen tree but not the outcomes they would have gotten from non-chosen trees.

investigates how the timing of regret influences decision-making within a controlled learning environment. Specifically, we examine whether regret is more likely to hinder decision-making during the early stages of learning, when information is limited, and whether it becomes more adaptive as more information becomes available. In doing so, we combine computational modeling and behavioral experiments, examining how the subjective experience of regret influences subsequent decisions.

Overview of current paradigm

To investigate how regret influences decision-making at different times in the learning process, we used a 3-armed bandit task with arms independent of each other. We used the same paradigm in both the experiment and the Bayesian model of regret described below. On each trial, the participant/model chooses from one of three trees and receives feedback (Figure 1a). The participant/model earns 1 point for each ripe fruit and 0 points for each rotten fruit (Figure 1b). The reward probabilities were defined for each tree as $P(\text{reward} \mid \text{tree}_1) = 0.7$, $P(\text{reward} \mid \text{tree}_2) = 0.5$, and $P(\text{reward} \mid \text{tree}_3) = 0.2$. The probabilities were randomly assigned to different trees and remained fixed throughout the task.

Bayesian model of learning from regret

Belief updating To represent its beliefs about each tree, the model maintains a Beta distribution parameterized by α and β for each tree. Initially, at trial $= 1$, all three trees have uniform priors ($\alpha = 1$, $\beta = 1$), yielding an estimated reward probability of 0.5 for each tree.

1. No Regret Model. Once the model chooses a tree, a reward of 1 or 0 is generated using a Bernoulli random process based on the true reward probability $P(\text{reward} \mid \text{tree}_i)$. After observing the outcome for the chosen tree, the model updates its α_i and β_i parameters. α is incremented by 1 when the reward is 1, and β is incremented by 1 when the reward is 0.

2. Grass-Is-Greener Model (GiG). In a standard Bayesian updating algorithm with independent options, only the chosen tree's parameters are updated and no information is gained about the non-chosen trees on each trial, because their values weren't observed. However, consistent with the definition of regret as an affective response that results when an individual believes that *the outcome would have been better had they chosen differently*, a regret-experiencing model must have the capacity to imagine that a different option might have yielded a reward when the chosen one did not. Thus, to model regret, we implement a mechanism that partially updates the values of *non-chosen* trees whenever the chosen tree returns 0.[1] Beliefs are updated as follows

$$\alpha_{\text{unchosen}} \leftarrow \alpha_{\text{unchosen}} + 1 \cdot \omega \quad (\text{if outcome} = 0), \qquad (1)$$

where ω is the *counterfactual weight*. Scaling the counterfactual update by ω allows us to implement partial belief updating for the non-chosen trees. A key conceptual reason for using a partial increment is that counterfactual outcomes (i.e., "imagined" outcomes) are, by definition, less certain and less informative than the actual observed outcomes. To provide a concrete example, $\omega = 0.5$ means that the model treats counterfactual outcomes as half as "valuable" as real ones. This partial update nudges the posterior mean of unselected trees upward, which simulates the cognitive mechanisms underlying experiences of regret.

3. Counterfactual Sampling Model (CS). Importantly, people might also engage in counterfactual simulation without assuming that a non-chosen option would have been better than the chosen one. To address this possibility, we implemented an alternative counterfactual belief updating mechanism in the *Counterfactual Sampling* model. After a failure to earn a reward, this model stochastically samples an alternative outcome for each non-chosen tree based on its current estimated probability of yielding a reward (represented by the mean of the Beta distribution for that tree). If a simulated outcome is a 1, the non-chosen tree is updated as though the model had actually observed a reward (scaled by ω). If a simulated outcome is a 0, the non-chosen tree is updated as though it had observed a negative outcome (scaled by ω).

Action selection All models select actions in the same way. On each trial, the model decides which tree to pick from using Thompson sampling (Russo et al., 2018). As described in the section above, the model holds, for every tree i, a Beta distribution over its reward probability. On each trial, the model samples a single continuous value from each tree's Beta distribution and selects the tree with the largest sampled value.

[1] Though the trees in our paradigm are independent, evidence from past work suggests that the process of deliberation creates an inverse association among the representations of the values of considered options (Biderman & Shohamy, 2021). Consequently, it is possible that participants may update the values of non-chosen options based on imagined outcomes.

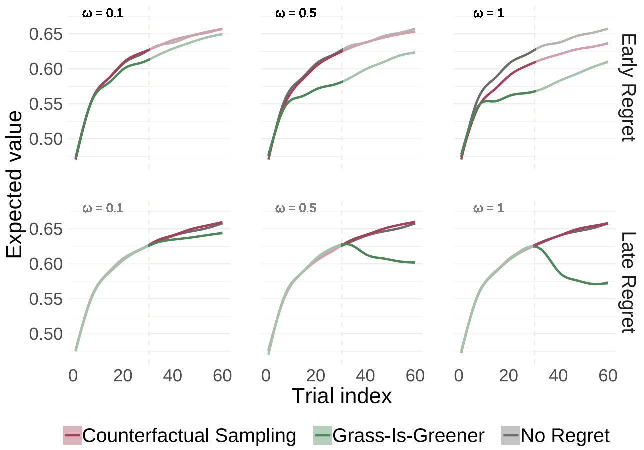

Figure 2: Simulated performance of the 3 models for different values of counterfactual weight based on 6,000 iterations per each combination of condition, model, and ω. Non-shaded areas highlight trials on which counterfactual belief updating took place in the CS and GiG models.

Theoretical simulation We simulated the performance of the 3 models, with 60 runs of 100 iterations per condition each, using different values of ω. As shown in Figure 2, the Grass-Is-Greener model performs worse on the task than the No Regret and the Counterfactual Sampling models across all conditions and ω values. This is because the model introduces a bias by assuming that the non-chosen options always would have returned a reward when the chosen one did not. Similarly, the Counterfactual Sampling model performs worse than the No Regret model in the 'early regret' condition. This is because the model's beliefs about the trees' distributions are uncertain early on, which means that sampling counterfactual outcomes from the non-chosen options mostly adds noise to the estimated reward probabilities. However, in the 'late regret' condition, the Counterfactual Sampling model sometimes outperforms the No Regret model (albeit only slightly). This is because, once the model has learned reliable estimates of the tree distributions, generating virtual observations by sampling counterfactual observations from non-chosen trees can lead to more robust estimates of those trees' expected values. Consistent with our hypotheses, these theoretical results show that there are situations, late in the learning process, in which updating one's beliefs based on simulated counterfactual observations can be helpful.

Experiment: Learning from regret

Consistent with the idea that the effects of regret on learning and decision-making are partially governed by the timing of *when* feelings of regret arise, we tested three pre-registered hypotheses[2]:

Hypothesis 1: When participants are prompted to reflect on their feelings of regret later in the task, they will exhibit better overall learning than participants who aren't asked to reflect on their feelings of regret at all.

Hypothesis 2: When participants are prompted to reflect on their feelings of regret earlier on in the task, they will exhibit worse overall learning than participants who aren't asked to reflect on their feelings of regret at all.

Hypothesis 3: The intensity of participants' self-reported feelings of regret on a given trial will depend on how good or bad their choice was compared to other options. Specifically, participants will feel more regret if the choice they made had a lower expected value compared to the alternatives.

Methods

Participants We recruited 176 participants ($M_{age} = 38, SD_{age} = 13$; 95 female, 75 male, 4 non-binary; 128 White, 24 Black or African American; 12 Multiracial; 11 Asian; 43 Hispanic) via Prolific. Participants were paid at a base rate of $0.65 ($8.57/hour) for a 5-minute experiment and received a bonus of 1 cent for each point they earned in the experiment, resulting in an average total compensation of $1.00 ($12.00/hour). Participants who resided in the United States, were fluent in English, had completed at least 10 previous submissions on Prolific, and had an approval rate of 95% or more were eligible to participate.

Procedure Following the informed consent procedure and instructions, participants went through 60 trials of the fruit-picking game (Figure 1). The game was programmed with jsPsych (de Leeuw, 2015). At the beginning of the experiment, participants were told that though the trees are independent, some trees are more likely to have ripe fruit than others, but that they would still occasionally receive rotten fruit from good trees and ripe fruit from bad trees. Their goal is to earn as many points as they can throughout the game.

Past research in affective science suggests that simply engaging with one's emotions during self-report enhances the prominence of those emotions in individuals' awareness, intensifying their experience (Robinson & Clore, 2002; Van Boven et al., 2010). Thus, to manipulate the timing of regret and the associated counterfactual thinking, we randomly assigned participants to one of three between-person conditions: *Control* (proceeding through the game with no questions about regret); 'early regret' (every time participants get a 0 they are asked to report how much they regret their choice in the first 30 trials); or 'late regret' (in the last 30 trials). Participants rated "How much do you regret selecting this tree?" on a $0--100$ slider scale ranging from "Not at all" to "Extremely".

Overall learning was operationalized as the sum of the expected values of the options that participants chose across all 60 trials. For example, if a participant chose a sequence of a 0.2 option and a 0.7 option on two consecutive trials, the total expected value over these trials would be calculated as $0.2+0.7=0.9$. This way of operationalizing learning provides a more direct measure than the total number of points

[2] All materials and preregistration links can be found at https://github.com/kateptrv/learning-from-what-might-have-been.

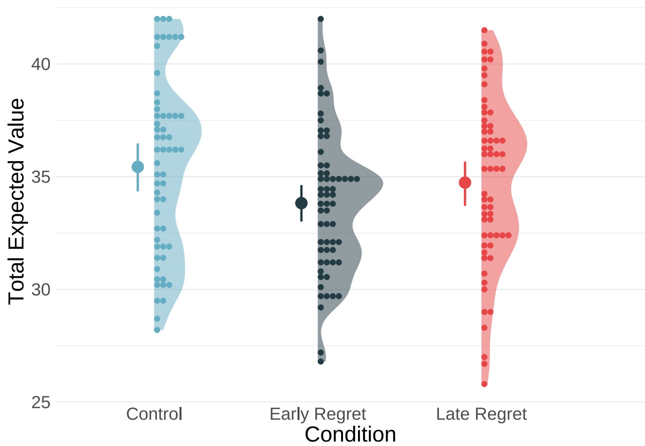

Figure 3: Participants' performance (total expected value of chosen trees across 60 trials) by condition. *Note*: Large points show condition-level means with 95% bootstrapped confidence intervals; small points show individual participants' data points.

earned by participants, since point allocation in the experiment is determined probabilistically.

Results

Empirical results To test Hypotheses 1 and 2 (Figure 3), we fit Bayesian linear mixed-effects models in *brms* (Bürkner, 2017) to evaluate the average difference in the total expected values (EV) of options selected by participants in different conditions. The model was parameterized as follows

$$\text{Total EV}_p = \text{B}_0 + \text{B}_1 \cdot \text{condition}_p + \varepsilon_p, \tag{2}$$

where p = participant. We report unstandardized values B throughout.

We found partial support for our hypotheses. Contrary to our expectations, participants in the 'late regret' condition did not perform better than participants in the 'control' condition, $B_1 = -.69$, 95% CrI $[-2.12, 0.72]$. However, as hypothesized, participants in the 'early regret' condition performed credibly worse than participants in the 'control' condition, $B_1 = -1.62$, 95% CrI $[-2.90, -0.32]$. These results indicate that while early regret can hinder learning by disrupting the acquisition of necessary information, late regret does not necessarily translate into improved learning performance; at least in our paradigm. This is somewhat contrary to the results of the theoretical simulation summarized in Figure 2, in which the CS model (slightly) outperformed the No Regret model in the 'late regret' condition when $\omega = 0.5$. Part of the reason for this discrepancy might be that participants in our experiment may engage in biased belief updating like the GiG model, or give more or less weight to their counterfactual belief updates than what is optimal.

To test Hypothesis 3, we fit a Bayesian linear mixed-effects model to evaluate the average difference in self-reported regret intensity following trials on which participants selected the options with different expected values. The model was parameterized as follows:

$$\text{Regret}_{p,t} = \text{B}_0 + \text{B}_1 \cdot EV_{p,t-1} + \varepsilon_{p,t} \tag{3}$$

where p = participant and t = trial index. Consistent with our hypothesis, participants across conditions reported lower subjective regret intensity following trials on which they selected trees with higher expected values, $B_1 = -28.85$, 95% CrI $[-37.67, -20.36]$.

Bayesian models of learning from regret

Parameter fitting. We implemented a set of Bayesian models (described in the *Bayesian model of learning from regret* section) to capture how different values of counterfactual weight (ω) might influence choice behavior at different stages of learning. We estimated ω as well as a learning rate parameter η (a coefficient used to scale belief updating for observed but not counterfactual data (Wilson et al., 2013)) separately for each participant with a Metropolis–Hastings Markov-Chain Monte-Carlo (MCMC) sampler, implemented in Python. The initial values of ω (for non-Control participants only) and η were randomly chosen from uniform priors ranging from 0.01 to 1.00. We ran chains of 2,000 iterations per participant, discarding the first 500 as burn-in. On each iteration, we used a log-normal random walk (to ensure non-negative values) to propose a new candidate value of ω and η. All 60 trials of the task were run for each candidate value of ω and η for each participant, with the GiG- and CS-specific belief-updating mechanisms. The loss function minimized the negative log-likelihood across 60 trials.

Because action selection in Thompson sampling involves choosing the arm with the highest value of the three draws, we cannot directly compute the probability of tree i being chosen in closed form. This probability, however, is required to calculate the log-likelihood of the model. We estimated this probability in several steps. First, we calculated the probability density function (PDF) of the chosen tree i for the value of the random draw x_i. Second, we calculated the cumulative density functions (CDF) for each of the two non-chosen trees j. Each CDF in this case represents the probability that the random draw from the Beta distribution of tree j is $\leq x_i$. Third, we derived the joint probability of the two non-chosen trees producing a draw $\leq x_i$ by taking the product of the two CDFs. Fourth, by multiplying the PDF of the chosen tree i with the product of the CDFs of the non-chosen trees, we got the probability that the chosen tree i returns a draw $= x$, and that both non-chosen trees produce draws lower than x. Finally, we integrated this product over all possible values of x (0 to 1, since x represents a probability of receiving a reward from a given tree) to derive the overall probability that tree i returns the highest draw and is, therefore, chosen on a given trial. Formally,

$$P(\text{tree } i \text{ is chosen}) = \int_0^1 PDF_i(x_i) \cdot \prod_{j \neq i} CDF_j(x_i)\, dx. \tag{4}$$

We used the resulting probability estimates to compare participants' choices to those of the model by calculating the log-likelihood of the data under each model for each trial and each value of ω and η. We used the standard Metropolis-Hastings rule (Chib & Greenberg, 1995) to compare the log-likelihoods of each new candidate value to the current value.

The means and standard deviations of each participant's posterior ω and η distributions based on the 1,500 post-burn-in iterations are shown in Table 1. Counterfactual weight ω was significantly higher in the 'early regret' condition compared to the 'late regret' condition in both the GiG model ($t(116) = 2.25, p = .03$) and the CS model ($t(116) = 2.21, p = .03$). The values of ω were also higher in the CS model compared to the GiG model, across conditions, possibly owing to the fact that the CS model bases counterfactual updates on the agent's own beliefs, making such updates more informative and, therefore, carrying more weight. The magnitude of belief-updating η , though not significantly different across conditions or models, varied significantly across participants, indicating that participants differed in how readily they incorporated new information into their beliefs.

Experiment simulation. We used the final parameter estimates of ω and η (the means of the posterior distributions) for each participant to simulate the decision-making process. We ran three simulations: (1) No Regret null model; (2) Grass-Is-Greener (GiG) model; and (3) Counterfactual Sampling (CS) model. The simulations generated predicted choices according to the model-specific procedures described in the *Bayesian model of learning from regret* section. We ran 100 independent simulations of 60 trials per model for each participant.

Model fit. We computed the empirical probability of the observed human responses by comparing the distributions of simulated outcomes with the human data (Figure 4). The cumulative log-likelihood for participant i was calculated by summing the log-likelihoods across all 60 trials.

We found that the GiG model provided a significantly better fit to the data compared to the No Regret null model. On average across participants, the Log-Likelihood Ratio (LLR) of the two models ($2 \cdot \Delta\text{LL}$) was 2.84, 95% CI $[1.29, 4.52]$ in favor of the GiG model (global LRT = 463.17, df = 120, $p < .0001$). The GiG model provided a better fit than the No Regret model for 74 out of the 120 participants in the 'early regret' and 'late regret' conditions. The CS model similarly provided a better fit to the data compared to the null model (average LLR = 2.25, 95% CI $[0.82, 3.87]$; global LRT = 290, df = 120, $p < .0001$), providing a better fit than the No Regret model for 71 out of the 120 participants in the two experimental conditions. In contrast, the GiG model and the CS model were not meaningfully different (average LRT = 2.37, 95% CI $[-0.61, 5.34]$. Taken together, these findings provide initial support for the idea that the addition of the counterfactual weight ω offers a meaningful improvement in the prediction of behavioral data collected under the regret paradigm over and above individual learning rate η. However, the current paradigm does not distinguish between the two hypothesized methods of counterfactual belief updating.

Table 1: Estimates of counterfactual weight (ω), learning rate (η) and model fit (AIC vs. the AIC of the Null model) for different models in different experimental conditions.

Condition	Mean ω (SD)	Mean η (SD)	AIC (Null)
	Grass-Is-Greener Model		
Early Regret	0.57 (0.83)	0.51 (0.31)	9,461 (9,763)
Late Regret	0.31 (0.63)	0.46 (0.31)	11,178 (11,339)
Control	N/A	0.50 (0.30)	10,215 (10,250)
	Counterfactual Sampling Model		
Early Regret	0.92 (1.62)	0.47 (0.32)	9,522 (9,763)
Late Regret	0.48 (0.79)	0.43 (0.31)	11,289 (11,339)
Control	N/A	0.47 (0.32)	10,223 (10,240)

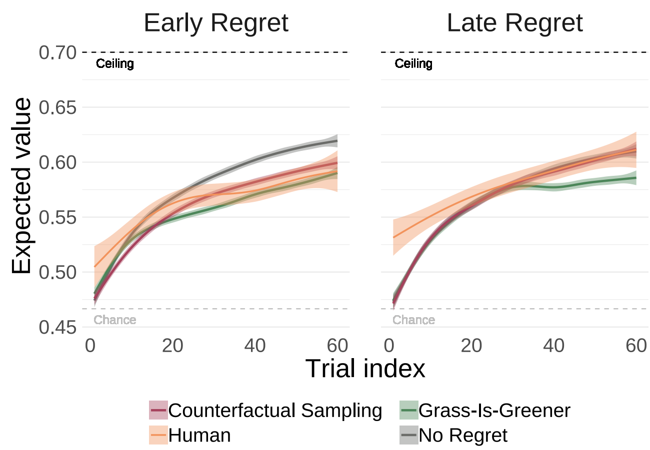

Figure 4: Learning curves of the three models compared to human behavioral data.

Model performance. Paralleling the empirical analyses of the human data, we compared how the models performed in the task. We found that, consistent with our hypothesis, the 'early regret' condition in the GiG model performed significantly worse on the task (i.e., selecting options with lower total expected values across 60 trials) than the 'control' condition, $B_1 = -2.06$, 95% CrI $[-2.56, -1.56]$. Contrary to our hypothesis but consistent with the results of the behavioral experiment, the 'late regret' condition similarly performed worse than the 'control' condition, $B_1 = -1.70$, 95% CrI $[-2.29, -1.09]$. The results of the CS model simulation revealed no significant differences in the task performance across conditions ('early regret' vs. 'control': $B_1 = -0.82$, 95% CrI $[-1.80, 0.12]$; 'late regret' vs. 'control': $B_1 = -0.22$, 95% CrI $[-1.11, 0.69]$).

Exploratory analyses. We conducted additional exploratory analyses into how counterfactual weight in the GiG and the CS models relates to subjective feelings of regret. We found that participants with higher estimated values of ω report, on average, more intense feelings of regret across the task, $r = 0.24$. Further, participants with higher values of ω chose options with lower total expected values across 60 trials, $r = -0.39$.

General Discussion

The present work investigated whether regret shapes decision-making and learning over time. The results of our behavioral experiment show that experiencing regret earlier in the learning process – when people have little information about which options are better – can impair learning, whereas experiencing regret later does not necessarily confer additional benefits (at least in our paradigm).

The fact that participants did not perform better when prompted to reflect on their feelings of regret later in the task might be due to the fact that they engaged in biased counterfactual belief updating, always assuming that the non-chosen option would have been better (consistent with the *Grass-Is-Greener* model). Future experimental work should explicitly instruct participants to estimate the probability of reward from non-chosen options rather than nudging participants towards feelings of regret, as we did here. Another possibility is that participants might have generated realistic counterfactuals, grounded in their prior beliefs about the expected values of different options (consistent with the *Counterfactual Sampling* model), but that they gave either not enough or too much weight to these counterfactually-driven belief updates. Findings from our theoretical simulations suggest that a more realistic counterfactual belief updating which characterizes the *Counterfactual Sampling* model does, in fact, lead to better learning – but only if it is applied selectively late in the learning process *and* if the counterfactual updates are appropriately weighted (specifically, if the counterfactual outcomes are treated as approximately half as "valuable" as the observed outcomes; in other words, ω = 0.5).

Another core contribution of this work lies in our modeling of what we consider to be a core component underlying the cognitive appraisal of regret – the "counterfactual weight" (ω) that partially updates beliefs about non-chosen options. This conceptualization aligns with the widely accepted view that regret involves counterfactual thinking – imagining "what might have been" if a different decision had been made. Notably, higher values of ω in our study predicted more intense subjective feelings of regret, lending support to the idea that the emotional experience of regret is closely tied to the cognitive appraisal of better alternatives. Another intriguing result is that individual counterfactual weight estimates often exceeded 1, suggesting that some people in our study may have weighted imagined counterfactual information stronger than actual observations. One possible explanation for this surprising pattern of results is that individuals might feel disproportionately distressed when they think they missed a better outcome – effectively interpreting the forgone gain as a "loss". This account would be consistent with the proposition of the prospect theory that losses loom larger than gains (Tversky & Kahneman, 1981). Surprisingly, we found that ω was higher in the 'early regret' condition across models, indicating that participants are more likely to view foregone gains as losses when they have less certainty about the reward structure of the environment. A deeper investigation into counterfactual weighting would be a promising direction for future research – both as a means of evaluating which model best captures human learning from regret, and as a window into individual cognitive processes. Examining variability in ω across individuals could help explain why some people learn adaptively from regret while others become stuck in maladaptive patterns of indecision.

Limitations and future directions

One limitation of this work is that the relatively short task (60 trials) may have constrained the window in which regret could exert a distinct influence at later stages of learning. On the other hand, there are relatively few situations in real life where people make such repeated decisions, so future work might focus on more naturalistic or even one-shot decision paradigms to better approximate real-world contexts. Additional work is also required to model more diverse or dynamic environments featuring varying reward structures to capture the complex ways regret might unfold in everyday decision-making. Second, our regret prompt manipulated self-reported regret following negative outcomes. While such self-reports can intensify the salience of regret, this manipulation may not capture real-world contexts where regret is experienced spontaneously and continuously rather than prompted in discrete moments. Finally, the current model primarily addresses the cognitive component of regret – namely, how counterfactual thinking inherent in the appraisal of regret influences belief-updating – while not fully capturing its affective component. The subjective experience of regret may play an important role in driving the *motivation* to learn. Additional work is needed to develop models that capture both the cognitive and affective underpinnings of emotions and their roles in behavior.

Conclusion

There is growing interest in applying computational cognitive approaches to study affective phenomena (e.g. Bedder et al., 2024; Berwian et al., 2020). Still, computational models of emotion remain relatively scarce. We believe that our work makes a meaningful contribution to the emerging field of *computational affective science*, helping clarify verbal theories by offering a testable framework for how emotions such as regret arise and evolve. By illustrating that a formally defined parameter (counterfactual weight) can predict not only observable behaviors but also the subjective emotional experience, this work bridges the gap between theory and the lived experience of emotion. In doing so, it highlights a path forward for computational cognitive science – one in which affect is not simply an afterthought, but rather an essential component of comprehensive models of human behavior.

Acknowledgments

TG was supported by grants from Stanford's Human-Centered Artificial Intelligence Institute (HAI) and Cooperative AI.

References

Bedder, R. L., Hitchcock, P. F., & Sharp, P. B. (2024). Unraveling repetitive negative thinking with reinforcement learning [Manuscript in preparation].

Berwian, I. M., Wenzel, J. G., Collins, A. G. E., Seifritz, E., Stephan, K. E., Walter, H., & Huys, Q. J. M. (2020). Computational mechanisms of effort and reward decisions in patients with depression and their association with relapse after antidepressant discontinuation. *JAMA Psychiatry*, *77*(5), 513–522.

Biderman, N., & Shohamy, D. (2021). Memory and decision making interact to shape the value of unchosen options. *Nature Communications*, *12*(1), 4648.

Bjälkebring, P., Västfjäll, D., Svenson, O., & Slovic, P. (2016). *Emotion*, *16*(3), 381–386.

Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28.

Chib, S., & Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician*, *49*(4), 327–335.

Clore, G. L., Gasper, K., & Garvin, E. (2012). Affect as information. In J. P. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 122–145). Psychology Press.

Coali, A., Gambardella, A., & Novelli, E. (2022). Understanding probabilistic reasoning in innovation. *Academy of Management Proceedings*.

Coricelli, G., Critchley, H., Joffily, M., O'Doherty, J., Sirigu, S., & Dolan, R. (2005). Regret and its avoidance: A neuroimaging study of choice behavior. *Nature Neuroscience*, *8*, 1255–1262.

de Leeuw, J. R. (2015). Jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12.

Dijk, E. V., & Zeelenberg, M. (2005). On the psychology of 'if only': Regret and the comparison between factual and counterfactual outcomes. *Organizational Behavior and Human Decision Processes*, *97*(2), 152–160.

Inman, J. J. (2007). Regret regulation: Disentangling self-reproach from learning. *Journal of Consumer Psychology*, *17*(1), 19–24.

O'Connor, E., McCormack, T., & Feeney, A. (2014). Do children who experience regret make better decisions? a developmental study of the behavioral consequences of regret. *Child Development*, *85*(5), 1995–2010.

Reb, J. (2008). Regret aversion and decision process quality: Effects of regret salience on decision process carefulness. *Organizational Behavior and Human Decision Processes*, *105*(2), 169–182.

Reb, J., & Connolly, T. (2009). Myopic regret avoidance: Feedback avoidance and learning in repeated decision making. *Organizational Behavior and Human Decision Processes*, *109*(2), 182–189.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological bulletin*, *128*(6), 934.

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, *121*(1), 133.

Russo, D. J., Van Roy, B., Kazerouni, A., Osband, I., & Wen, Z. (2018). A tutorial on thompson sampling. *Foundations and Trends in Machine Learning*, *11*(1), 1–96.

Shani, Y., & Zeelenberg, M. (2007). When and why do we want to know? how experienced regret promotes post-decision information search. *Journal of Behavioral Decision Making*, *20*(3), 207–222.

Sijtsema, J. J., Zeelenberg, M., & Lindenberg, S. M. (2022). Regret, self-regulatory abilities, and well-being: Their intricate relationships. *Journal of Happiness Studies*, *23*, 1189–1214.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Van Boven, L., Kane, J., McGraw, A. P., & Dale, J. (2010). Feeling close: Emotional intensity reduces perceived psychological distance. *Journal of personality and social psychology*, *98*(6), 872.

Wilson, R. C., Nassar, M. R., & Gold, J. I. (2013). A mixture of delta-rules approximation to bayesian inference in change-point problems. *PLoS Computational Biology*, *9*(7), e1003150.

Zeelenberg, M., Beattie, J., der Pligt, J. V., & Vries, N. K. D. (1996). Consequences of regret aversion: Effects of expected feedback on risky decision making. *Organizational Behavior and Human Decision Processes*, *65*(2), 148–158.

Zeelenberg, M., Dijk, W. W. V., Manstead, A. S., & van der Pligt, J. (2000). On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition & Emotion*, *14*(4), 521–541.