Regret Induces Rapid Learning from Experience-based Decisions: A Model-based Facial Expression Analysis Approach

Nathaniel Haines[1], Olga Rass[2], Yong-Wook Shin[3], Jerome R. Busemeyer[2], Joshua W. Brown[2], Brian F. O'Donnell[2], & Woo-Young Ahn[4]

[1] Department of Psychology, The Ohio State University, Columbus, OH
[2] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN
[3] Department of Psychiatry, Ulsan University School of Medicine, Seoul, Korea
[4] Department of Psychology, Seoul National University, Seoul, Korea

Correspondence:
Woo-Young Ahn, Ph.D.
Department of Psychology
Seoul National University
Seoul, Korea 08826
Tel: +82-2-880-2538, Fax: +82-2-877-6428. Email: wahn55@snu.ac.kr

Keywords: Reinforcement learning; Bayesian data analysis; emotion; facial expression; counterfactual reasoning; computational modeling

Abstract

Regret—an emotion comprised of both a counterfactual, cognitive component and a negative, affective component—is one of the most commonly experienced emotions involved in decision making. For example, people often behave such that their decisions minimize potential regret and therefore maximize subjective pleasure. Importantly, functional accounts of emotion suggest that the experience and future expectation of regret should promote goal-oriented behavioral change. While many studies have confirmed the functional role of regret, the cognitive-emotional mechanisms through which regret facilitates behavioral change remain unclear. We hypothesize that a greater negative affective state accompanying the counterfactual component of regret will potentiate learning, value, and/or exploration of potential options—all of which could lead to behavioral change. Because prior studies on regret-driven decision making have focused almost exclusively on description-based paradigms, the potential role of learning from regret through experience is underexplored. Here, we leverage computational, emotion-driven models of decision making to determine the role of regret in risky decision making. Further, we use computer-vision to detect positive and negative affect (valence) intensity from subjects' faces in response to feedback, which was entered as direct input into each model to determine whether valence affects learning, valuation, or exploration/exploitation. Using multiple model comparison methods, we found that: (1) people weight regret by its expectedness when making experience-based decisions, and (2) people learn more rapidly as they experience increasingly intense negative affect. Our findings support functional accounts of regret and demonstrate the potential for model-based facial expression analysis to enhance our understanding of cognition-emotion interactions.

1. Introduction

1.1 The adaptive function of regret

Among the various emotions underlying our decisions (e.g., Davis, Love, & Maddox, 2009; Frey, Hertwig, & Rieskamp, 2014; Heilman, Crişan, Houser, Miclea, & Miu, 2010), *regret*—a counterfactual, aversive emotional state that occurs when we make a choice and then later wish to have made an alternative choice (Kahneman & Miller, 1986)—is one of the most frequently experienced and discussed emotions in everyday situations (Shimanoff, 1984). Importantly, people who experience regret either most often and/or most intensely endorse more severe symptoms of anxiety and depression in addition to lower life satisfaction than their peers (Kocovski, Endler, Rector, & Flett, 2005; Lecci, Okun, Karoly, 1994, 1994; Monroe, Skowronski, Macdonald, & Wood, 2005). Despite being associated with multiple negative functional outcomes, people look back on regretted decisions with appreciation, which is not true of other negative emotions like anger and disappointment (Saffrey, Summerville, & Roese, 2008). These seemingly contradictory findings are reconciled by functional accounts (see Keltner & Gross, 1999) of counterfactual thinking which posit that regret facilitates goal-oriented behavioral change by signaling us to avoid making unjustified choices that lead to negative, unwanted outcomes (Epstude & Roese, 2008; Roese, 1994; Zeelenberg & Pieters, 2007). Indeed, the intensity of regret that we experience is proportional to how active our role is in the regret-inducing decision, how justifiable our decision is, and the quality of our decision process (e.g., Inman & Zeelenberg, 2002; Pieters & Zeelenberg, 2005). Further, we experience the most regret in domains where we have opportunities for corrective action (Roese & Summerville, 2005). However, not all negatively-valenced counterfactual thinking (e.g., disappointment) leads to behavioral change, suggesting that regret serves a specific function in changing our behavior (e.g., Zeelenberg et al., 1998).

While observational and survey-based studies consistently implicate regret in goal-oriented behavioral change, the cognitive mechanisms responsible for regret-induced behavioral change remain underexplored. In particular, regret may facilitate behavioral change by modifying how we learn from, value, or explore/exploit our choices, and each of these mechanisms has different implications for regret-induced behavioral change. By determining the cognitive and emotional mechanisms involved in regret-based decision making, we may better understand why regret can sometimes lead to positive behavioral changes (e.g., corrective action), whereas other times it can lead to negative behavioral outcomes (e.g., depression, anxiety, lower life satisfaction, etc.).

1.2 Regret theory

Regret theory encompasses a variety of computational models that describe how people incorporate expected regret into their subsequent decisions (Bell, 1982; Loomes & Sugden, 1982b), thus offering a potential way of understanding the role of regret in behavioral change. According to regret theory, counterfactual comparisons between chosen and foregone outcomes lead to cognitive-emotional states of *regret* or *rejoicing* if the comparison is negative or positive in value, respectively. By assuming that people make decisions to maximize their emotional expectations, regret theory can capture shortcomings of expected utility theory while maintaining a simple form with relatively few assumptions compared to other models (Loomes & Sugden, 1982a; 1982b). Disappointment theory was later developed and incorporated into models of regret (Bell, 1985; Loomes & Sugden, 1986). Similar to regret, disappointment is a counterfactual emotion that arises when a choice could have turned out better (e.g., a gamble could have resulted in a win but instead resulted in a loss). Unlike regret, disappointment does not involve the comparison of an alternative choice to the received outcome.

Decision Affect Theory (DAT) is an instantiation of regret theory which assumes that decisions in the face of uncertainty are driven by emotions including regret, disappointment, and their counter-parts of rejoicing and elation (Mellers, Schwartz, Ho, & Ritov, 1997; Mellers, Schwartz, & Ritov, 1999). Specifically, DAT assumes that individuals make decisions by maximizing their emotional expectations (R), or subjective emotional pleasure, which follow the general form:

$$\begin{aligned} R \propto\ & \text{Chosen Outcome Utility} + \\ & \text{Regret / Rejoice} + \\ & \text{Disappointment / Elation} \end{aligned} \tag{1}$$

Here, regret and rejoice both involve counterfactual comparisons across possible choice alternatives (i.e. "*I am upset [glad] that I chose this option.*"), whereas disappointment and elation reflect counterfactual comparisons across alternative states of the world (i.e. "*I am upset [glad] that my choice turned out this way.*"). Notably, these counterfactual comparisons occur frequently in daily life. For example, imagine attending your favorite restaurant and ordering something new off the menu. Upon finishing your meal, you would experience disappointment if the meal was worse than you expected (i.e. "*That could have been much better.*"). Further, you would also experience regret if you compared the new meal to previous meals that you enjoyed

much more (i.e. "*I should have ordered my favorite meal.*"). In addition to including terms for disappointment and elation, DAT further expands on regret theory by assuming that people weight their emotional expectations in proportion to how *unexpected* (i.e. surprising) an outcome is. For example, when choosing between two novel meals with either a high (meal A) versus low (meal B) expected probability of tasting better than your favorite meal, DAT predicts that you would experience *more regret* by choosing and disliking meal A because it has a *high expected probability* of tasting better than your favorite meal (i.e. the negative outcome is more unexpected for meal A compared to B).

1.3 Integrating emotion and decision models

DAT has been used to successfully model how individuals make decisions based on description (i.e., probabilities associated with outcomes are given; Coricelli et al., 2005; Mellers et al., 1997; 1999), yet the majority of decisions that we make in everyday life are based on experience (i.e., probabilities are learned from experience). Given the large and reliable differences observed in how people make description- versus experience-based decisions (e.g., Barron & Erev, 2003; Hertwig, Barron, Weber, & Erev, 2004; Ungemach, Chater, & Stewart, 2009; Wulff, Mergenthaler-Canseco, & Hertwig, 2018), it is unclear if DAT as applied to description-based tasks generalizes to experience-based tasks. Namely, people act *as if* they overweight low probabilities when making description-based choices yet overweight low probabilities when making experience-based decisions. This difference is important because DAT assumes that regret intensity is proportional to surprise, rather than expectedness (Mellers et al., 1997; 1999). Since DAT uses the probability of an outcome as a proxy for surprise (see eq. 4–5), and past applications of DAT have focused on description-based tasks, the model may need modification to account for differential probability weighting in experience-based tasks. More importantly, surprise-weighted regret is inconsistent with functional theories of regret—such as decision justification theory and regret regulation theory (Connolly & Zeelenberg, 2002; Zeelenberg & Pieters, 2007)—which show that people experience more intense regret when they make unjustified or low quality decisions (Inman & Zeelenberg, 2002; Pieters & Zeelenberg, 2005). Using the meal example from section 1.2, an unjustified/low quality decision would occur when you order a new meal despite knowing from your prior dining experience that there is a *low probability* of the new meal being better than what you typically order. Therefore, regret should increase in proportion to the probability of a regretful outcome (i.e. regret expectedness)

to facilitate behavioral changes that maximize emotional expectations. This is a notable contrast with theories of surprise-based learning such as predictive coding, supervised learning with the delta rule, and temporal difference models of reinforcement learning.

Although computational models provide a formal way to test theories of emotion-based decision making and have led to novel insights into human behavior, they are typically developed to account for choice data alone without making explicit assumptions about the dynamic relationship between experienced emotions and behavior. Indeed, several studies have applied counterfactual models to experience-based tasks (Boorman, Behrens, & Rushworth, 2011; Hayden, Pearson, & Platt, 2009; Lohrenz, McCabe, Camerer, & Montague, 2007; Yechiam & Rakow, 2012); yet few studies have related components of these models to specific emotional processes as clearly as DAT (Mellers et al., 1997), and still fewer have used independent measures of emotion (e.g., self-reports, skin conductance response, etc.) to inform cognitive model development (Jian Li, Schiller, Schoenbaum, Phelps, & Daw, 2011; Rutledge, Skandali, Dayan, & Dolan, 2014). In one of few studies of its kind, Coricelli et al. (2005) showed that people become increasingly regret-averse as they experience more regret-inducing outcomes. Notably, the effect of cumulative regret aversion on subsequent choices was mediated by BOLD signaling in the medial orbitofrontal cortex and amygdala, which are part of a network that plays a crucial role in the development of emotionally relevant cue-outcome contingencies (e.g., Sharpe & Schoenbaum, 2016). However, the task used in Corcielli et al. (2005) was description-based, so the model used could not specifically test whether regret affected learning versus valuation or for outcome evaluation mechanisms. Additionally, no external measures of emotion were included in the model, so it is unclear if subjects' emotional responses dynamically modulated their learning or valuation of regret throughout the task. The lack of studies incorporating real-time emotions into regret-based decision models is a potentially important oversight, given both the importance of negative affect in motivating behavioral change and recent movements toward linking the traditionally separated components of cognition and emotion (Barrett, 2009; Duncan & Barrett, 2007; Eldar et al., 2016; Etkin et al., 2015).

1.4 The current study

Here, we use: (1) a decision-making task that evokes counterfactual thinking; (2) automated computer-vision emotion coding (e.g., Haines et al., 2019); and (3) computational, emotion-driven reinforcement learning models to identify the cognitive mechanisms responsible for

regret-induced behavioral change in experience-based decision making. Computer-vision allows us to objectively measure trial-by-trial positive and negative affect during decision-making, which fills a critical gap in previous studies of regret; namely, prior studies have focused almost exclusively on the cognitive—but not the affective—component of regret, in part because objective emotion coding is difficult to implement during behavioral tasks. Further, we use hierarchical Bayesian analysis and Bayesian model comparison techniques to test a series of competing hypotheses regarding the relationship between cognitive and affective components of regret, which we formalize as separate computational models. Specifically, we hypothesize that people should weight regret by its expectedness (Ahn et al., 2012) rather than surprise, as in the original DAT instantiation; this approach is in line with the literature on the functional value of regret. Additionally, we hypothesize that emotional facial expressions that people produce while undergoing decision making will represent either: (1) learning, (2) valuation, or (3) exploration/exploitation, and that this effect will be specific to regret rather than other counterfactual emotions.

2. Method

2.1 Subjects

The current study included 50 subjects' data collected from two different research sites. Of these subjects, 31 had facial expressions recorded, while 19 are from a previous study which did not record facial expressions (Ahn et al., 2012). We aggregated both datasets for model comparison purposes (see sections 2.6 and 3.1 for details). All subjects gave informed consent prior to participating in the study. The study protocol was approved by local Institutional Review Boards at both research sites.

2.2 Behavioral Task

All subjects completed four separate gambling games in randomized order. Subjects were told that each game was independent of all other games, and they were given an opportunity for a break between each game. Each game consisted of 90 trials, where subjects were asked to choose between two options (see Fig. 1). Throughout each game, selecting one of the options won a fixed amount of points (i.e. safe option), whereas the other option had some probability of winning a high or low amount of points (i.e. risky option). Locations of the safe and risky options were randomized across subjects but remained fixed within games. The probability of

winning a high number of points for the risky option was fixed but unknown, and subjects had to learn the probability from experience. After making a choice, point values for both the chosen and unchosen options were revealed (i.e. "full-information"), which allowed subjects to make counterfactual comparisons between the choices they made and could have made. Subjects were instructed to make choices that maximized their points. Unbeknownst to subjects, the expected value for each option was identical (see Table 1 for payoff distributions for each game).

Figure 1

Behavioral task time course

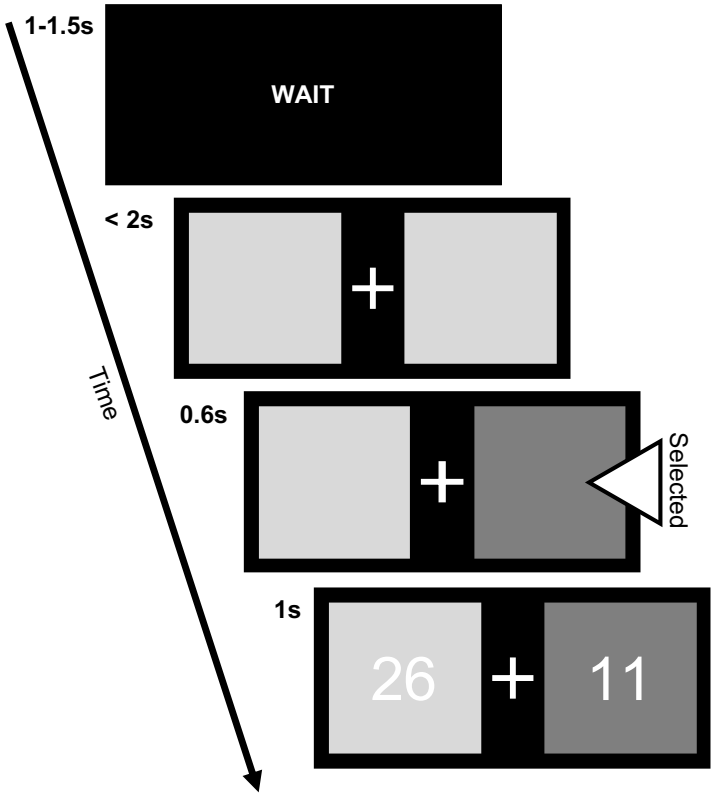


Table 1

Payoff structure for each game

Game	*M*	Risky Option				
		L	*H*	Pr(*H*)	EV	SD
1	12	1	56	0.2	12	22.0
2	11	1	26	0.4	11	12.3
3	10	1	16	0.6	10	7.4
4	9	1	11	0.8	9	4.0

Note. M = Payoff for the safe option. *L* = Low payoff for the risky option. *H* = High payoff for the risky option. Pr(*H*) = Probability of receiving high payoff given choosing the risky option. EV = Expected value of the risky option (note that the EV of the risky option is equivalent to the value of the safe option within each game). SD = Standard deviation of the risky option.

2.3 Computational Models

We conducted model comparison on five different models to test our hypothesis that regret is weighted by its expectedness, rather than surprise. Specifically, we fit models which assumed that subjects engaged in decisions based on: (1) fictive learning alone, (2) surprise-weighted regret, (3) expectedness-weighted regret (Ahn et al., 2012), (4) unweighted emotional terms, or (5) expectedness weights for all emotional terms. All five models assumed that subjects learned the probability of the high payoff of the risky option according to the delta rule (Rescorla & Wagner, 1972):

$$\Pr_H(t+1) = \Pr_H(t) + \alpha \cdot [Y(t) - \Pr_H(t)] \tag{2}$$

where $\Pr_H(t)$ is the expected probability of the high payoff of the risky option occurring on trial t, α $(0 < \alpha < 1)$ is the learning rate, and $Y(t)$ is an indicator for whether the risky option resulted in the high ($Y(t) = 1$) or low ($Y(t) = 0$) payoff on trial t. Values of α close to 1 (0) indicate rapid (slow) updating of the expected high payoff for the risky option. Additionally, all models assumed that subjects made choices according to the Luce choice rule (a.k.a. softmax function; Luce, 1959) with a trial-independent inverse temperature parameter (Yechiam & Ert, 2007):

$$\Pr_{Risky}(t+1) = \frac{e^{\theta \cdot Q_R(t+1)}}{e^{\theta \cdot Q_R(t+1)} + e^{\theta \cdot Q_S(t+1)}} \tag{3}$$

where $\Pr_{Risky}(t+1)$ is the probability of choosing the risky option on trial t, θ is the inverse temperature parameter determined by $\theta = 3^c - 1$ $(0 < c < 5)$, and $Q_R(t+1)$ and $Q_S(t+1)$ are the action values for the risky and safe options, respectively. Higher (lower) values of θ indicate that subjects are making choices that are more deterministic (random) with respect to their risky and safe action values. Note that risky and safe action values are determined differently for each of our three models, which we describe below.

2.3.1 Fictive Learning Alone (Fictive). The fictive learning model (Fictive) functions as a baseline model (i.e. no emotional expectations) to compare the DAT-derived models to. We use the term "fictive" to refer to the updating of $\Pr_H$ on each trial, irrespective to subjects' choice of the safe or risky option. Therefore, the Fictive model assumes that subjects update their action values for each option separately and then make their choices according to equation 3. Specifically, the action value for the safe option is determined by $Q_S(t+1) = M^{\omega}$, where M is the objective outcome of the safe option and ω $(0 < \omega < 1.5)$ is a shape parameter controlling how sensitive subjects are to rewards. Note that the safe option is the same across trials within each game, so $Q_S(t+1)$ is always the same value in the Fictive model. Conversely, the action value for the risky option is determined by $Q_R(t+1) = H^{\omega} \cdot \Pr_H(t+1) + L^{\omega} \cdot [1 - \Pr_H(t+1)]$, where H and L are the potential high and low payoffs for the risky option, respectively.

2.3.2 Surprise-weighted Regret (Original). For all regret-based models, we assumed that individuals make choices to maximize their subjective emotional pleasure (Ahn et al., 2012; Mellers et al., 1999). Therefore, the safe and risky action values are determined by weighted expected emotional responses (see equation 1) rather than utilities, as in the Fictive model.

Following DAT, the Original model assumes that regret/rejoice and disappointment/elation are determined by:

$$sgn(A - B) \cdot |A - B|^{\omega} \cdot W \tag{4}$$

where A and B represent values being counterfactually compared and W indicates the weight of the counterfactual comparison on the emotional response (Mellers et al., 1999). For regret and rejoice, the counterfactual comparison is across choice alternatives such that A is the received payoff and B is the payoff that would have been received if option B was chosen. For disappointment and elation, the comparison is between potential states of the choice such that A is the received payoff and B is the alternative payoff to option A—because the safe option only has a single payoff, disappointment and elation only apply to the risky option. Importantly, the

Original model weights all emotional responses by their surprisingness, which we captured by setting W to 1 minus the probability of the outcome. For example, if a subject chooses the risky option, they will receive either L (low payoff) or H (high payoff), and the foregone payoff will always be M (medium payoff). If they receive L, they will experience disappointment by comparing state L to alternative state H (e.g., $sgn(L-H) \cdot |L-H|^{\omega} \cdot W$). In addition to disappointment, they will experience regret by comparing their received payoff L to foregone payoff M (e.g., $sgn(L-M) \cdot |L-M|^{\omega} \cdot W$). For both terms, the emotional weight (W) is equal to the surprisingness of receiving L instead of H (e.g., $W = 1 - \Pr_L(t+1)$)[1].

Because our task offers full-information on two choice alternatives, there are 4 possible paired-outcomes that can occur on each trial which lead to different emotional responses: (1) received M, foregone H ($R_{M(H)}$), (2) received M, foregone L ($R_{M(L)}$), (3) received H, foregone M ($R_{H(M)}$), and (4) received L, foregone M ($R_{L(M)}$)[2]. According to equation 1, these emotional expectations along with the utilities of each outcome are updated on each trial according to the following set of (simplified) equations:

$$\begin{aligned}
R_{M(H)}(t+1) &= M^{\omega} - |M-H|^{\omega} \cdot \Pr_L(t+1), \\
R_{M(L)}(t+1) &= M^{\omega} + |M-L|^{\omega} \cdot \Pr_H(t+1), \\
R_{H(M)}(t+1) &= H^{\omega} + |H-M|^{\omega} \cdot \Pr_L(t+1) + |H-L|^{\omega} \cdot \Pr_L(t+1), \\
R_{L(M)}(t+1) &= L^{\omega} - |L-M|^{\omega} \cdot \Pr_H(t+1) - |L-H|^{\omega} \cdot \Pr_H(t+1)
\end{aligned} \tag{5}$$

Note that the probability of a low/high payoff from the risky option is updated (i.e. learned) on each trial, which in turn is used to weight the emotional response terms. Finally, the expected emotional responses to both the safe and risky options are a weighted sum of the above emotional responses:

$$\begin{aligned}
Q_S(t+1) &= R_{M(H)}(t+1) \cdot \Pr_H(t+1) + R_{M(L)}(t+1) \cdot \Pr_L(t+1) \\
Q_R(t+1) &= R_{H(M)}(t+1) \cdot \Pr_H(t+1) + R_{L(M)}(t+1) \cdot \Pr_L(t+1)
\end{aligned} \tag{6}$$

---

[1] $\Pr_L(t+1) = 1 - \Pr_H(t+1)$

[2] We use the notation $R_{A(B)}$ to represent the emotional response (R) to received (A) and foregone (B) payoffs.

The resulting emotional expectations are entered into equation 3 to generate a probability of selecting the safe and risky options.

2.3.3 Expectedness-weighted Regret (Modified). The Modified regret model is equivalent to the Original version in all but one respect: the Modified model weights regret terms by their expectedness rather than by their surprisingness (Ahn et al., 2012). We accomplish this by reversing the following emotional response terms (bolded) from equation 5:

$$\begin{aligned} R_{M(H)}(t+1) &= M^{\omega} - |M-H|^{\omega} \cdot \mathbf{Pr_H}(t+1), \\ R_{L(M)}(t+1) &= L^{\omega} - |L-M|^{\omega} \cdot \mathbf{Pr_L}(t+1) - |L-H|^{\omega} \cdot \mathrm{Pr}_H(t+1) \end{aligned} \tag{7}$$

We weighted regret by its expectedness because observational and survey-based evidence consistently shows that people experience more regret when they make low-quality or unjustified decisions (e.g., Inman & Zeelenberg, 2002; Pieters & Zeelenberg, 2005). Intuitively, equation 7 captures this effect by making the regret term more extreme as the regretful outcome becomes more likely. The resulting emotional response terms are converted to choice probabilities in the same way as the Original model (see equations 6 and 3).

2.3.4 Unweighted. The Unweighted model is identical to the Original version, except the decision weights for all emotion terms are set to 1 (i.e. $W = 1$; see equation 4). We tested an unweighted model to determine whether or not the assumption of surprisingness- and expectedness-weighted emotional expectations in the Original and Modified models were supported by the data.

2.3.5 Expectedness-weighting for all terms (All Expected). The All Expected model is identical to the Original version, except the decision weights for all emotion terms are reversed to indicate expectedness rather than surprise. We tested this model to determine if expectedness-weighting was specific to regret or common to all emotional terms.

2.4 Model-based Facial Expression Analysis

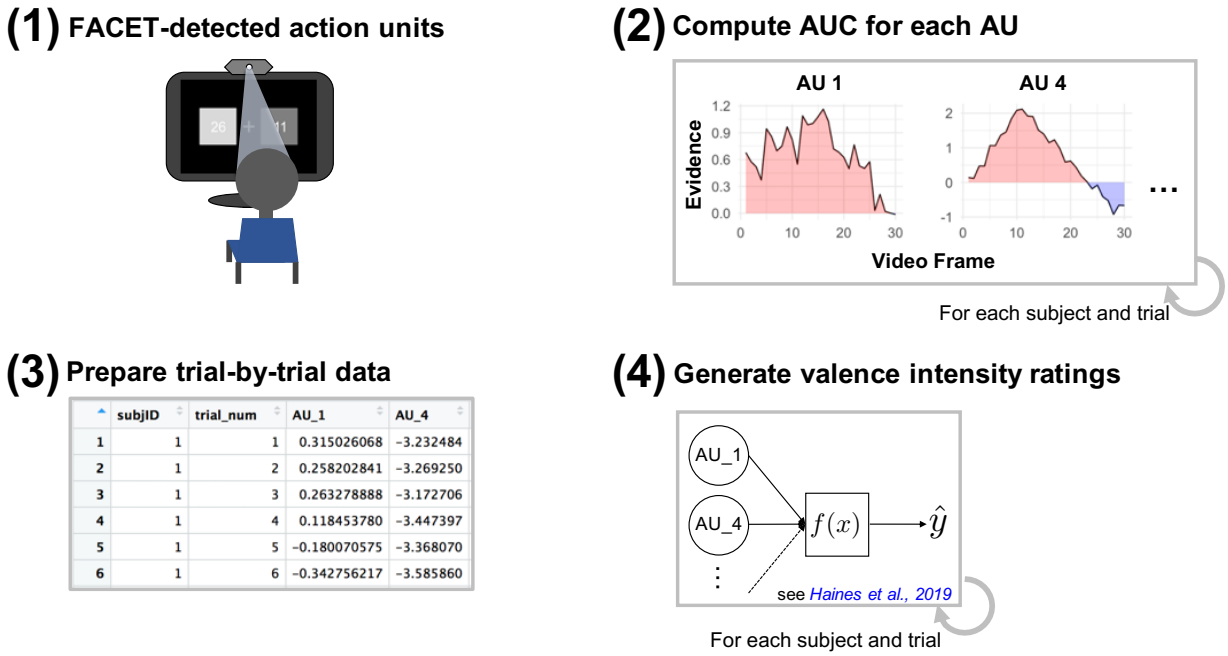
2.4.1 Automated Coding. To measure the valence of subjects' facial expressions during feedback, we used an automated facial expression coding (AFEC) model that we developed in a previous study (Haines et al., 2019). The AFEC model was trained to code for positive and negative affect intensity on a scale from 1 (*no affect*) to 7 (*extreme affect*), where positive and negative affect are coded separately rather than on a polarized positive–negative valence

continuum. The AFEC model first uses FACET—a computer vision software (iMotions, 2018)—to detect the presence of 20 different facial action units (Ekman, Friesen, & Hager, 2002), which are then translated to affect intensity ratings using a machine learning model that we previously developed (Haines et al., 2019). In our validation study, the model showed correlations with human ratings of .89 and .76 for positive and negative affect intensity, respectively (for more details, see Haines et al., 2019).

Figure 2 shows the steps used to preprocess and apply the AFEC model to our subjects' facial expressions. First, we used FACET to detect the presence of 20 different facial action units (AUs) during the feedback phase of the task. FACET-detected AUs are derived from the anatomically-based Facial Action Coding System (FACS; Ekman et al., 2002), which is arguably the most comprehensive and widely-used facial coding systems available today. FACET outputs a time-series of values for each AU at a rate of 30 Hz, where values represent the probability (i.e. "evidence") that a given AU is present in each sample. Second, we computed the area-under-the-curve (AUC) of each AU time-series and divided (a.k.a. normalized) the resulting value by the total length of time that a face was detected throughout the 1 second feedback phase in the task (per trial). Normalization ensures that clips of varying quality (e.g., 70% versus 90% face detection accuracy) do not affect the magnitude of the AUC values, which is important for the machine learning step. We excluded any trials where a subject's face was detected for less than 10% of the total 30 samples in the given trial (~3% of trials excluded in total). After iterating step 2 for each trial and subject (step 3), we entered the resulting values as predictors in the AFEC model described above to generate intensity ratings for positive and negative affect (Haines et al., 2019). We used the positive and negative affect ratings as input into the computational models as described below.

Figure 2

Steps for preparing facial expression data



	subjID	trial_num	AU_1	AU_4
1	1	1	0.315026068	-3.232484
2	1	2	0.258202841	-3.269250
3	1	3	0.263278888	-3.172706
4	1	4	0.118453780	-3.447397
5	1	5	-0.180070575	-3.368070
6	1	6	-0.342756217	-3.585860

Notes. (1) Subjects' facial expressions were recorded during the outcome phase of the counterfactual task. We used FACET to capture the probability of each of 20 facial Action Units (AUs) being present in subjects' facial expressions in response to the outcome. (2) Evidence for each AU (i.e. probability of AU being present) over the 1 s outcome phase was converted to a single score by taking the area under the curve (AUC) of each evidence time series. The AUC values were normalized by the clip length. (3) Steps (1) and (2) were repeated for each trial within each game and subject. (4) We used two Random Forest models developed in a previous study to generated separate ratings for positive and negative affect intensity based on the AUC scores computed in step (3).

2.4.2 Facial Expression Models. To determine whether emotional facial expressions reflected choice mechanisms of learning (α), valuation (ω), or exploration/exploitation (c), we developed 3 competing models that used the positive and negative affect intensity scores to modulate trial-by-trial model parameters ($\alpha,\ \omega,\ c$). To do so, we computed an overall valence score for each trial by taking the difference in positive and negative affect ratings for the given subject and trial. We chose this approach over modeling positive and negative affect as separate

dimensions to reduce the number of possible models for model comparison purposes. We then standardized (i.e. z-scored) the valence ratings across subjects, games, and trials before using them as input to the model. We parameterized each model such that the respective model parameter for trial t was a linear combination of a baseline parameter and a parameter determining the effect of emotion valence intensity. For example, the learning rate parameter was determined by:

$$\alpha(t+1) = logit^{-1}(\alpha_0 + \alpha_1 \cdot FE(t)) \cdot 1.0 \tag{8}$$

Here, α_0 and α_1 indicate the baseline learning rate and effect of facial expression valence intensity on the learning rate for the next trial ($\alpha(t+1)$), respectively, and $FE(t)$ is the standardized facial expression valence rating on trial t. Note that the inverse logit function transforms the output so that $\alpha(t+1) \in (0, 1)$, which are the appropriate lower and upper bounds, so we scaled the output by 1.0 (i.e. no scaling). We used the same parameterization for ω and c, except the scaling factors were 1.5 and 5, respectively. On trials where subjects' faces were detected for less than 10% of the feedback stage, we assumed that parameters were not affected by facial expressions. Using the learning rate model as an example, if facial expression data on trial t was discarded, then:

$$\alpha(t+1) = logit^{-1}(\alpha_0) \tag{9}$$

In summary, each of the models makes an explicit assumption about which of the three decision mechanisms (i.e. learning, valuation, or exploration/exploitation) is affected by moment-to-moment emotional valence intensity, which allowed us to take a model-based approach to test our competing hypotheses.

To further identify which specific emotional response terms (i.e. regret, rejoice, disappointment, and elation) were most influenced by momentary emotional valence, we took the best fitting of the above models and parameterized it so that only the specific emotional response was influenced by emotional valence. Specifically, with the Modified regret model where emotional valence determined the learning rate parameter, we developed 4 variations where each assumed that the effect of emotion on learning was specific to either: regret, rejoice, disappointment, or elation expectations. For example, there are two emotional responses that

include terms for regret, namely $R_{M(H)}$ and $R_{L(M)}$. In the regret-specific learning model, we assumed that the probability of a high payoff on the risky option ($\Pr_H$) was learned separately (i.e. $\Pr_{Hreg}$) for regret-specific terms in $R_{M(H)}$ and $R_{L(M)}$. $\Pr_{Hreg}$ was updated according to:

$$\Pr_{Hreg}(t+1) = \Pr_H(t) + \alpha(t+1) \cdot [Y(t) - \Pr_H(t)] \tag{10}$$

Here, $\Pr_H$ is determined by equation 2, except the learning rate is set to $\alpha = logit^{-1}(\alpha_0)$. Conversely, the learning rate for $\Pr_{Hreg}$ is determined by momentary emotional valence such that $\alpha(t+1) = logit^{-1}(\alpha_0 + \alpha_1 \cdot FE(t))$. Therefore, the regret-specific learning model assumes that $\Pr_H$ is learned according to a standard delta updating rule (i.e. equation 2), but the learned expectedness of regret responses for each trial are further influenced by emotional valence (cf. bolded terms in equation 7)[3]:

$$\begin{aligned} R_{M(H)}(t+1) &= M^{\omega} - |M-H|^{\omega} \cdot \mathbf{Pr_{Hreg}}(t+1), \\ R_{L(M)}(t+1) &= L^{\omega} - |L-M|^{\omega} \cdot \mathbf{Pr_{Lreg}}(t+1) - |L-H|^{\omega} \cdot \Pr_H(t+1) \end{aligned} \tag{11}$$

Models specific to rejoice, disappointment, and elation terms followed the same logic, where $\Pr_{Hreg}$ and $\Pr_{Lreg}$ replaced $\Pr_H$ and $\Pr_L$ for the respective terms. In summary, the emotion-specific learning models allowed us to determine if momentary emotion valence had specific learning effects on different types of counterfactual comparisons.

2.5 Hierarchical Bayesian Analysis

We fit all models using hierarchical Bayesian analysis (HBA). HBA allows for individual-level (i.e. subject-level) parameter estimation while simultaneously pooling information across subjects to increase certainty in individual-level estimates. Further, HBA has previously been shown to provide better parameter recovery than traditional methods such as individual-level maximum likelihood estimation (MLE) (e.g., Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011),

---

[3] $\Pr_{Lreg}(t+1) = 1 - \Pr_{Hreg}(t+1)$

suggesting that individual-level HBA estimates can be interpreted with more confidence compared to traditional MLE estimates. We used Stan (version 2.15.1) to implement HBA. Stan is a probabilistic programming language that employs the No-U-Turn Hamiltonian Monte Carlo (HMC) sampler, which is a variant of Markov Chain Monte Carlo (MCMC), to efficiently sample from the joint posterior distribution across all specified model parameters (Carpenter et al., 2017).

We used a standard convention to parameterize the prior distributions for all model parameters (Ahn, Haines, & Zhang, 2017). Specifically, we assumed that each set of individual-level parameters was drawn from a group-level distribution. We assumed normal group-level distributions, where prior means and standard deviations were set to normal distributions.[4] We used non-centered parameterizations to decrease the dependence between group-level mean and standard deviation parameters (Betancourt & Girolami, 2013). Bounded parameters (e.g., $\alpha \in (0, 1)$) were estimated in an unconstrained space and then Probit-transformed (i.e. inverse cumulative distribution function of the standard normal) to the constrained space to maximize MCMC sampler efficiency (Ahn et al., 2014; 2017; Wetzels, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2010). Once transformed to the constrained space, parameters with upper bounds greater than 1 were scaled accordingly. For example, the learning rate parameter from the Fictive, Original, and Modified models was parameterized as:

$$\begin{aligned} \mu_\alpha &\sim \text{Normal}(0, 1) \\ \sigma_\alpha &\sim \text{Normal}(0, 0.2) \\ \boldsymbol{\alpha}' &\sim \text{Normal}(0, 1) \\ \boldsymbol{\alpha} &= \text{Probit}(\mu_\alpha + \sigma_\alpha \cdot \boldsymbol{\alpha}') \end{aligned} \tag{12}$$

Here, μ_α and σ_α are the group-level mean and standard deviation, respectively, and bold terms indicate vectors of individual-level parameters. This parameterization ensures that the prior

[4] Note that group-level hierarchies were assigned to each game separately to test parameter consistency across games, but the same hierarchy was assigned to all games to assess simulation performance and post-hoc model fits (section 2.6 describes this in more detail).

distribution over each individual-level parameter is near-uniform. For the models incorporating facial expressions, we used an alternative parameterization. Using the learning rate from the Modified learning model as an example, priors on the modified learning rates were set as follows:

$$\begin{aligned} \mu_{\alpha_0,\alpha_1} &\sim \text{Normal}(0,1) \\ \sigma_{\alpha_0,\alpha_1} &\sim \text{Normal}(0,1.35) \\ \boldsymbol{\alpha_0}, \boldsymbol{\alpha_1} &\sim \text{Normal}(0,1) \end{aligned} \tag{13}$$

We used these priors because they led to near-uniform priors over the individual-level trial-by-parameters (i.e. $\alpha(t+1)$) after being determined by equation 8. The same parameterization was used for ω and c in the Modified valuation and exploration/exploitation models.

We ran all models for 4,000 iterations across 4 separate sampling chains, with the first 1,000 samples as warm-up (analogous to burn-in in Gibbs samplers) for a total of 12,000 posterior samples for each parameter. We checked convergence to the target joint posterior distribution by visually inspecting trace-plots and ensuring that all Gelman-Rubin (a.k.a. $\hat{R}$) statistics were below 1.1, which suggests that the variance between chains is lower than the variance within chains (Gelman & Rubin, 1992). R and Stan codes for the computational models will be made available on the hBayesDM package (Ahn et al., 2017) upon publication.

2.6 Model comparison

We used three different methods to compare models, namely: (1) parameter consistency across the 4 games, which is a measure of how generalizable model parameters are across different variations of the same task (Yechiam & Busemeyer, 2008); (2) long-term prediction accuracy, which is a measure of the difference between subjects' entire choice histories and choice histories simulated from their fitted model parameters (Steingroever, Wetzels, & Wagenmakers, 2014); and (3) short-term prediction accuracy, which is a measure of how accurately a model can predict subjects' choices on the next trial given their fitted model parameters and choice history (Ahn, Busemeyer, Wagenmakers, & Stout, 2008). We used multiple methods because previous studies consistently show that different model comparison methods can lead to different conclusions (Ahn et al., 2008; 2014; Haines, Vassileva, & Ahn, 2018; Steingroever et al., 2014; Yechiam & Ert, 2007).

2.6.1 Parameter consistency. Parameter consistency, which is a measure that indicates how representative model parameters are of subjects' underlying cognitive processes, refers to how similar parameters are between the same model fit to different variations of the same task (e.g., Rieskamp, Busemeyer, & Laine, 2003; Yechiam & Busemeyer, 2008). In our study, the 4 games have different payoff distributions, yet we assume that subjects should learn, value, and explore/exploit options in the same way across tasks. Therefore, an ideal model should estimate parameters that are similar across the 4 games when fit to each game separately (i.e. the group-level hierarchies described in section 2.5 were applied separately to each game). After fitting each model/game separately, we used a graphical check to determine parameter consistency across games.

2.6.2 Long-term prediction accuracy. We assessed long-term prediction accuracy using the simulation method (Ahn et al., 2008; Steingroever et al., 2014). The simulation method proceeds by first fitting each model to subjects' choice data, followed by extracting fitted model parameters and using them to simulate subjects' choice preferences given the appropriate task structure. Note that we fit a single model—with one group hierarchy (see section 2.5) for each parameter—to all 4 games to test if a single set of model parameters for each subject could accurately capture choice behavior across games. We then compared the simulated and true choice preferences using graphical and quantitative measures. Here, we apply a fully-Bayesian simulation method that has been validated in previous studies (Haines, Vassileva, & Ahn, 2018; Steingroever et al., 2014; Steingroever, Wetzels, & Wagenmakers, 2013). Briefly, we took 100 random draws (with replacement) from each subject's joint posterior distribution and simulated their choice behavior for each of the 4 games. For each game, we stored each subject's simulated probability of selecting the risky option on each trial and for each posterior sample, averaged across the 100 samples within each subject, and then averaged across subjects to determine the average probability of selecting the risky choice across subjects. We then computed the mean squared deviation (MSD) between true and simulated choices as follows:

$$\text{MSD} = \frac{1}{4 \cdot n} \sum_{t=1}^{n} \sum_{g=1}^{4} (\bar{D}_{true_g}(t) - \bar{D}_{sim_g}(t))^2 \tag{14}$$

Here, n is the number of trials in each game (i.e. 90), t is the trial number, g is the game number, and $\overline{D}_{true}$ and $\overline{D}_{sim}$ are the true and simulated choice proportions after averaging across subjects,

respectively. Note that we smoothed the true and simulated choice proportions after computing the across-subject averages with a moving average filter of width 5 before computing the MSD, which is customary in prior studies using the MSD metric (Ahn et al., 2008; Haines, Vassileva, & Ahn, 2019).

2.6.3 Short-term prediction accuracy. We assessed short-term prediction accuracy using the leave-one-out information criterion (LOOIC; Vehtari, Gelman, & Gabry, 2016), which is a fully-Bayesian analogue of traditional, commonly used information criteria (e.g., AIC and BIC). LOOIC approximates true leave-one-out prediction accuracy, and it can be computed using the log pointwise posterior predictive density (lpd) of a fitted model. Similar to long-term prediction accuracy, we fit a single hierarchical model across all 4 games to determine which model could best represent choice behavior across games using a single set of parameters for each subject. To compute the lpd of each competing model, we calculated the log likelihood of each subjects' true choice on trial $t + 1$ conditional on their parameter estimates and choice history (i.e. trials $\in \{1,2, \ldots , t\}$). The log likelihood was calculated for each posterior sample and summed within each subject across games (preserving all posterior samples), resulting in an $N \times S$ lpd matrix where N and S are the numbers of subjects and posterior samples, respectively. We used the *loo* R package, which is developed by the Stan team (Vehtari et al., 2016), to compute LOOIC values from the lpd matrix. Note that LOOIC is on the deviance scale where lower values indicate better model fit (including complexity penalization).

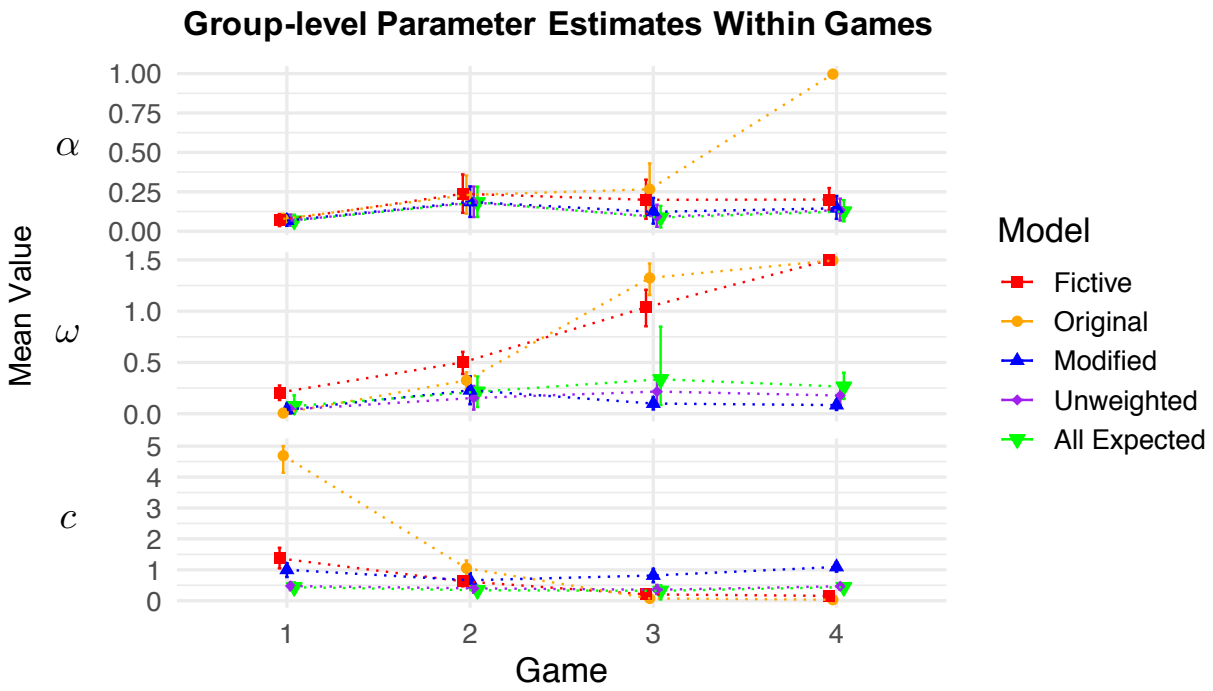
3. Results

3.1 Model Comparison: Parameter consistency.

Fig. 3 shows the group-level means (N=50) for the learning rate, valuation, and choice sensitivity (exploration/exploitation) parameters for each of the models and fit separately to the 4 games. Clearly, the Modified and Unweighted models provide the most consistent, precise parameter estimates across games, whereas both the Fictive and Original models have parameters that vary widely from one game to another. The All Expected model provides consistent parameter estimates, yet some are not precise (i.e. ω on game 3). These results suggest that the Modified and Unweighted models provide parameter estimates that more reliably represent subjects' underlying cognitive processes.

Figure 3

Group-level Parameter Estimates within Games



Note. Means and 95% highest density intervals (HDIs) are shown for each group-level parameter across the 4 games used in the current study. See Table 1 for descriptions of each game.

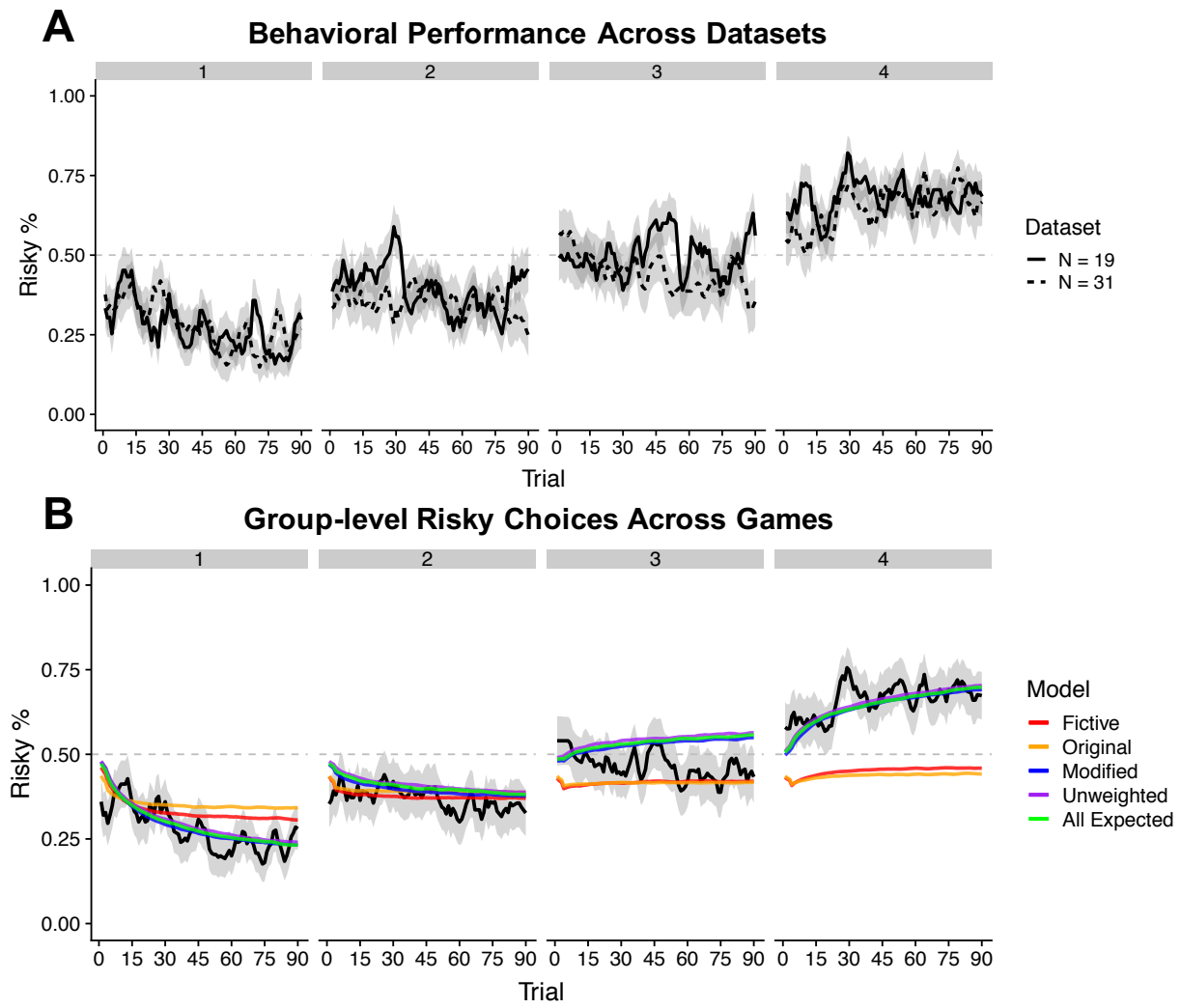
3.2 Model Comparison: Long-term prediction accuracy.

Fig. 4A shows that subjects performed similarly across both datasets, which justifies aggregating the data to increase the sample size for model comparison purposes. Fig 4B shows both the true and simulated across-subject (*N*=50) choice proportions for the safe versus risky options. Note that we fit each model to all 4 games simultaneously to assess long-term prediction accuracy. Despite the safe and risky options having the same expected value within each game, subjects showed a clear preference for the risky option in games where the high payoff/extreme outcome was more likely to occur—this pattern of behavior is consistent with previous studies showing that people tend to underweight rare events and/or over-value extreme outcomes when making decisions from experience (e.g., Barron & Erev, 2003; Hertwig et al., 2004; Ludvig & Spetch, 2011; Ludvig, Madan, & Spetch, 2013). Further, the graphical simulation results suggest that the Modified, Unweighted, and All Expected models are able to capture the evolution of subjects' choice preferences over time better than the Fictive and Original models. The MSDs for the Fictive, Original, Modified, Unweighted, and All Expected models were 0.0141, 0.0166,

0.00295, 0.00381, and 0.00332, respectively, indicating that the Modified model produced simulated choice preferences with the least deviation from subjects' true choices.

Figure 4

Group-level Behavioral Choice Proportions



Note. (A) Aggregated choice behavior across subjects, within datasets and games. The solid lines represent the mean choice proportions and shading represents ± 2 standard errors around the mean. A moving average filter of width 5 was used to smooth the behavioral data. Note that performance was similar across both datasets, so we aggregated the datasets to assess long- and short-term prediction accuracy with a larger sample (*N*=50). (B) True behavioral (black line) and model simulated (colored lines) proportion of risky choices across datasets and subjects within each game. Shading around the behavioral data represents ± 2 standard errors. A moving average

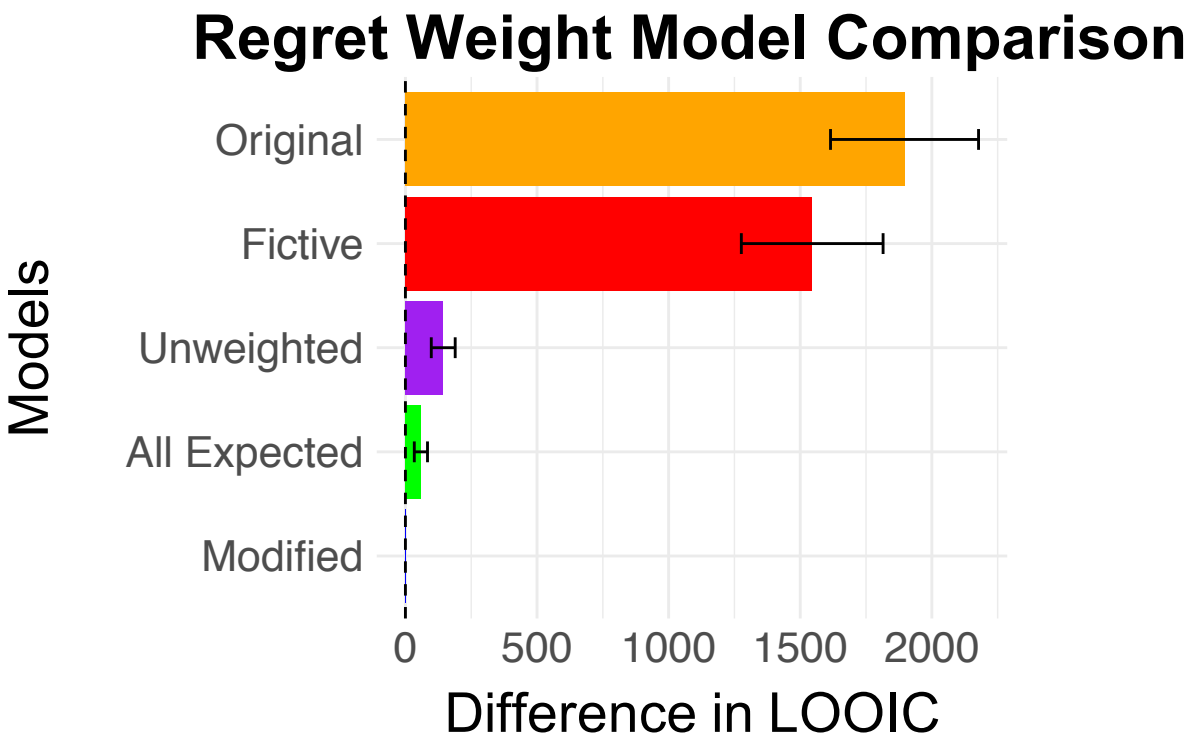
filter of width 5 was used to smooth both the behavioral and simulated data. Note that unlike in Figure 3, a single set of parameters was estimated for each model across games to test how well a single set of parameters could generate the varying choice patterns across games.

3.3 Model Comparison: Short-term prediction accuracy.

The Modified model showed the best short-term prediction accuracy, followed by the All Expected, Unweighted, Fictive, and then Original models (see Fig. 5). These results are consistent with a previous analysis of 19 of the 50 subjects in the current dataset (Ahn et al., 2012), despite using a different model fitting procedure (MCMC versus MLE) and different fit statistics (LOOIC versus BIC). Notably, the short- and long-term prediction accuracy and parameter consistency tests all identified the Modified model as either the best-fitting model or one of the best fitting models.

Figure 5

Model comparison



Note. Leave-one-out information criterion (LOOIC) scores relative to the best fitting model. Lower LOOIC values indicate better model fit accounting for model complexity. Error bars

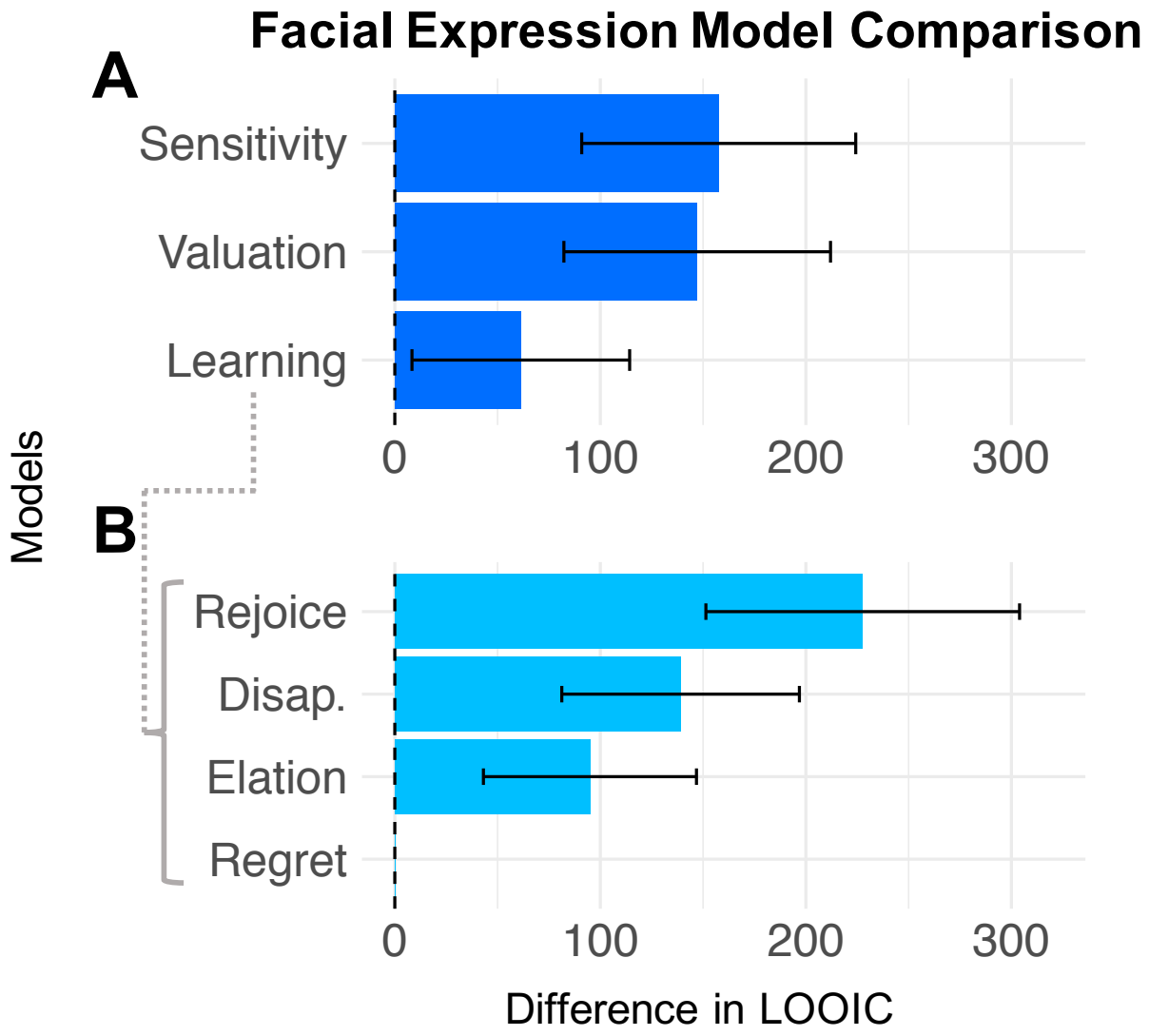
represent ± 1 standard error of the difference between the best fitting model (i.e. the Modified model) and respective competing models.

3.4 Model-based facial expression analysis.

Note that all facial expression analyses were conducted on the 31 subjects with facial expressions recorded. Because the Modified regret model showed better performance than the competing models across all three model comparison methods, we used it to further determine the effect of emotion valence intensity—measured through automated facial expression coding—on either learning, valuation, or choice sensitivity (exploration/exploitation). Additionally, we do not have a generative model of facial expressions to use for assessing long-term prediction accuracy (i.e. simulation), so we limited model comparison to short-term prediction accuracy (i.e. LOOIC) for all facial expression models. Fig. 6 shows that the Modified model with a trial-by-trial effect on the learning rate parameter, as opposed to valuation or choice sensitivity parameters, provided the best fit, suggesting that the facial expressions that people make in response to feedback are indicative of a dynamic learning mechanism. Notably, the group-level posterior distribution of valence on learning (μ_{α_1}) was mostly negative (95% highest density interval (HDI) = [-1.05, 0.13]), with 93% of the posterior mass below 0, indicating that subjects tend to update their expectations more rapidly as they express more intense negative facial expressions and more slowly as they express more positive facial expressions. Further, our data were best described by a model assuming that only regret expectations, as opposed to rejoice, disappointment, or elation expectations (or all 4), are linked to rapid emotion-induced updating (see Fig. 6 for comparisons and Table 1 for absolute LOOIC values). In the Modified model with a trial-by-trial effect on the learning rate for regret expectations only (see equations 10–11), the effect of emotion valence was also mostly negative (95% HDI = [-1.19, 0.10]), with 95% of the posterior mass below 0 (for a graphical depiction of the effect, see Fig. 7). These results suggest that the rapid learning following negative emotional expressions is most specific to regret expectations.

Figure 6

Comparison of choice mechanism and emotion-specific learning models



Note. Difference in LOOIC values between models including facial expressions to generate trial-by-trial model parameters. (A) Sensitivity, Valuation, and Learning models were used to determine which model parameter that facial expressions best captured, and we subsequently (B) developed variants of the best fitting model (the Modified regret with trial-by-trial learning rate) to test if emotion-induced learning was specific to regret compared to other emotional expectations. Error bars represent ± 1 standard error of the difference between the best fitting model and respective competing models. Note that differences are all relative to the Regret model in the bottom panel.

Table 1

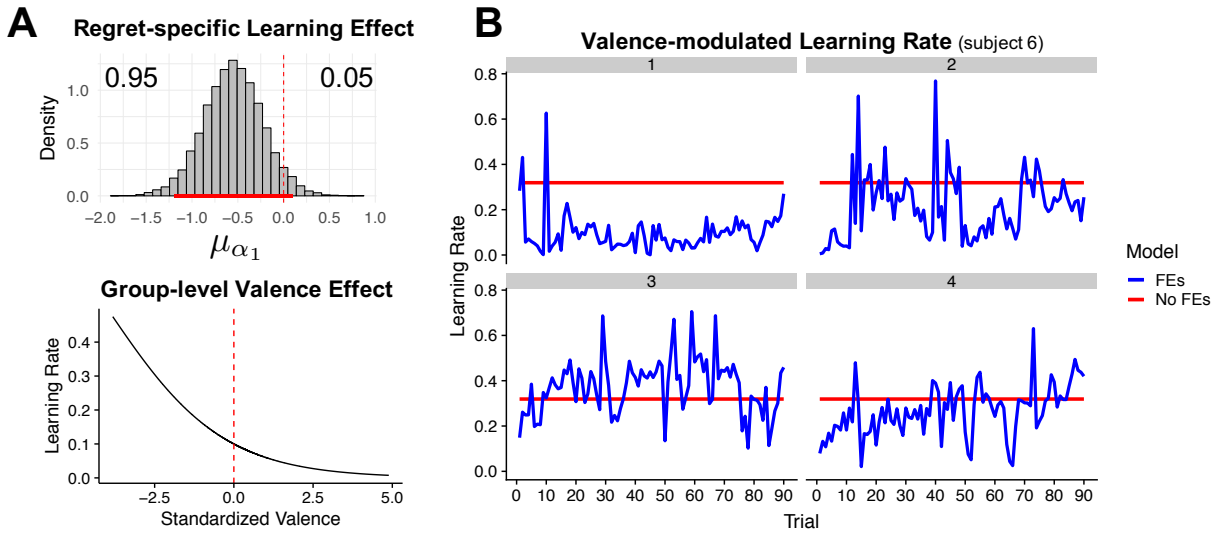
Absolute LOOIC values for competing facial expression models

	Model	Absolute LOOIC	
		Mean	SE
Cognitive Mechanisms	Sensitivity	12541	270
	Valuation	12534	271
	Learning	12444	277
Emotion-specific Learning	Rejoice	12619	266
	Disappointment	12538	275
	Elation	12490	255
	Regret	12387	276

Note. Absolute LOOIC values (and standard errors) for each of the models compared in Fig 6. Lower absolute LOOIC values indicate better model fits while penalizing for model complexity. Emotion-specific Learning models are testing for specific effects of trial-by-trial learning on Rejoice, Disappointment, Elation, or Regret.

Figure 7

Group- and individual-level effects of emotional valence intensity on regret learning



Note. (A) Posterior distribution over the effect of facial expressions on the regret-specific learning rate (see eq. 8–10) and a representation of the range of learning rates based on the group-level learning effect and valence ratings across subjects. As facial expression valence becomes increasingly negative, the learning rate becomes increasingly rapid and vice-versa. (B) An example from one subject of the learning rate with and without trial-by-trial effects of facial expression valence.

4. Discussion

Our findings are two-fold. First, we found that regret is weighted by its expectedness, rather than its surprisingness, for experience-based decisions. Specifically, subjects expected increasingly intense regret as they made decisions that were increasingly likely to lead to a regretful outcome. These results are in line with functional accounts of regret, which show that people experience more regret when they make poor quality or unjustified decisions (e.g., Inman & Zeelenberg, 2002; Pieters & Zeelenberg, 2005). Consistent with previous studies using models derived from DAT, subjects' choices were best described when other emotional expectancies (i.e. rejoice, disappointment, and elation) were weighted by their surprisingness (Mellers et al., 1997; 1999). Second, we found that people's emotional valence intensity during feedback—measured through their facial expressions—modulated their subsequent learning, such that increasingly negative (positive) valence intensity led to more rapid (slow) learning. While previous studies have determined that regret plays a role in behavioral change, our findings shed light on the cognitive mechanisms through which negative affect interacts with cognitive components of regret to produce changes in behavior.

Altogether, our results suggest that regret leads to behavioral change when the counterfactual, cognitive component is accompanied by intense negative affect (see Fig. 7), which may explain why regret can sometimes lead to negative functional outcomes and other times lead to positive experiences such as appreciation (Kocovski et al., 2005; Lecci et al., 1994; Monroe et al., 2005; Saffrey et al., 2008). For example, negative functional outcomes may be the result of dysfunctional interactions between the cognitive and emotional components of regret that typically facilitate learning and behavioral change. Indeed, the orbitofrontal cortex (OFC) and amygdala interact to produce both regret-averse decision making and extinction learning in healthy adults (Coricelli et al., 2005; Finger, Mitchell, Jones, & Blair, 2008); altered OFC-amygdala functional/structural connectivity is associated with a number of psychiatric disorders

(e.g., Dougherty et al., 2004; Greening & Mitchell, 2015; Hahn et al., 2011; Passamonti et al., 2012). Further, regret has minimal effects on individuals who are highly impulsive and lack trait anxiety (Baskin-Sommers, Stuppy-Sullivan, & Buckholtz, 2016), yet regret is overabundant in those with high trait anxiety (e.g, Roese et al., 2009), implicating an important role for regret-driven decision making in psychological disorders characterized by emotion dysregulation. Future studies may use regret-inducing tasks to further explore how the interactions between affective and cognitive components of regret present in individuals with different personality traits and/or psychiatric disorders (see Etkin, Büchel, & Gross, 2015). More broadly, our results are consistent with recent shifts toward conceptualizing emotions as fundamental, adaptive components of human cognition that help us make optimal inferences within constantly changing environments (Eldar, Rutledge, Dolan, & Niv, 2016).

Notably, this is the first study of its kind to include dynamic facial expressions as direct input into a cognitive model, although similar model-based approaches are becoming increasingly common in cognitive neuroscience (Turner, Forstmann, Love, Palmeri, & Van Maanen, 2017). Further, work using automated facial expression coding is gaining traction in social and behavioral sciences due to its efficiency relative to human coders (e.g., Cheong, Brooks, & Chang, 2017; Haines et al., 2019). The advantage of using facial expressions, as opposed to other measurement modalities such as EEG or psychophysiological measures (e.g., heart rate variability, skin conductance, facial electromyography, etc.), is that visual facial features consistently provide the single most optimal measure of emotional valence intensity, whereas other modalities are better suited for arousal (e.g., Chao, Tao, Yang, Li, & Wen, 2015; Kanluan, Grimm, & Kroschel, 2008). While our approach is limited in that we only measured valence, future studies may incorporate multiple measurement modalities to better capture different dimensions of emotion. In fact, physiological measures such as skin conductance response, eye-tracking, EEG, and fMRI have previously been used to inform cognitive models (e.g., Cavanagh, Eisenberg, Guitart-Masip, Huys, & Frank, 2013; Frank et al., 2015; Krajbich, Armel, & Rangel, 2010; Jian Li et al., 2011). Moreover, work on joint modeling suggests that model parameters can be estimated more precisely as more measurement modalities are included (Turner, Rodriguez, Norcia, McClure, & Steyvers, 2016). Given the known limitations of behavioral data collected from many popular decision making tasks (e.g., Hedge, Powell, & Sumner, 2017), leveraging multiple data modalities may be an optimal way forward in developing and testing increasingly complex cognitive models. For example, future studies may use the joint modeling

approach to explore the social dynamics of decision making, wherein regret expectations and facial expressions—among other response modalities—play a crucial role in negotiation and trust (Larrick & Boles, 1995; Martinez & Zeelenberg, 2014; Reed, DeScioli, & Pinker, 2014; Reed, Zeglen, & Schmidt, 2012).

Acknowledgement

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT, & Future Planning (2018R1C1B3007313). Some of this research was presented at the 34th Annual Meeting of the Cognitive Science Society 2012, Sapporo, Japan.

Declarations of interest

None

References

Ahn, W.-Y., Rass, O., Shin, Y.-W., Busemeyer, J., Brown, J., & O'Donnell, B. (2012). Emotion-based reinforcement learning. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 124–129).

Ahn, W.-Y., Busemeyer, J., Wagenmakers, E.-J., & Stout, J. (2008). Comparison of Decision Learning Models Using the Generalization Criterion Method. *Cognitive Science*, *32*(8), 1376–1402. http://doi.org/10.1080/03640210802352992

Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Computational Psychiatry*, *1*, 24–57. http://doi.org/10.1162/CPSY_a_00002

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), 95–110. http://doi.org/10.1037/a0020684

Ahn, W.-Y., Vasilev, G., Lee, S.-H., Busemeyer, J. R., Kruschke, J. K., Bechara, A., & Vassileva, J. (2014). Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. *Frontiers in Psychology*, *5*, 1376. http://doi.org/10.3389/fpsyg.2014.00849

Barrett, L. F. (2009). The Future of Psychology: Connecting Mind to Brain. *Perspectives on Psychological Science*, *4*(4), 326–339. http://doi.org/10.1111/j.1745-6924.2009.01134.x

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*(3), 215–233. http://doi.org/10.1002/bdm.443

Baskin-Sommers, A., Stuppy-Sullivan, A. M., & Buckholtz, J. W. (2016). Psychopathic individuals exhibit but do not avoid regret during counterfactual decision making. *Proceedings of the National Academy of Sciences*, *113*(50), 14438–14443. http://doi.org/10.1073/pnas.1609985113

Bell, D. E. (1982). Regret in Decision Making under Uncertainty. *Operations Research*, *30*(5), 961–981. http://doi.org/10.1287/opre.30.5.961

Bell, D. E. (1985). Disappointment in Decision Making Under Uncertainty. *Operations Research*, *33*(1), 1–27. http://doi.org/10.1287/opre.33.1.1

Betancourt, M. J., & Girolami, M. (2013, December 3). Hamiltonian Monte Carlo for Hierarchical Models. *arXiv.org*.

Boorman, E. D., Behrens, T. E., & Rushworth, M. F. (2011). Counterfactual Choice and Learning in a Neural Network Centered on Human Lateral Frontopolar Cortex. *PLoS Biol*, *9*(6), e1001093. http://doi.org/10.1371/journal.pbio.1001093

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., et al. (2017). Stan : A Probabilistic Programming Language. *Journal of Statistical Software*, *76*(1), 1–32. http://doi.org/10.18637/jss.v076.i01

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal Theta Overrides Pavlovian Learning Biases. *Journal of Neuroscience*, *33*(19), 8541–8548. http://doi.org/10.1523/JNEUROSCI.5754-12.2013

Chao, L., Tao, J., Yang, M., Li, Y., & Wen, Z. (2015). Long Short Term Memory Recurrent Neural Network based Multimodal Dimensional Emotion Recognition. *the 5th International Workshop* (pp. 65–72). New York, New York, USA: ACM. http://doi.org/10.1145/2808196.2811634

Cheong, J. H., Brooks, S., & Chang, L. J. (2017, November 1). FaceSync: Open source framework for recording facial expressions with head-mounted cameras. https://doi.org/10.31234/osf.io/p5293

Connolly, T., & Zeelenberg, M. (2002). Regret in Decision Making. *Current Directions in Psychological Science*, *11*(6), 212–216. http://doi.org/10.1111/1467-8721.00203

Coricelli, G., Critchley, H. D., Joffily, M., O'Doherty, J. P., Sirigu, A., & Dolan, R. J. (2005). Regret and its avoidance: a neuroimaging study of choice behavior. *Nature Neuroscience*, *8*(9), 1255–1262. http://doi.org/10.1038/nn1514

Davis, T., Love, B. C., & Todd Maddox, W. (2009). Anticipatory emotions in decision tasks: Covert markers of value or attentional processes? *Cognition*, *112*(1), 195–200. http://doi.org/10.1016/j.cognition.2009.04.002

Dougherty, D. D., Rauch, S. L., Deckersbach, T., Marci, C., Loh, R., Shin, L. M., et al. (2004). Ventromedial Prefrontal Cortex and Amygdala Dysfunction During an AngerInduction Positron Emission Tomography Study in Patients With Major DepressiveDisorder With Anger Attacks. *Archives of General Psychiatry*, *61*(8), 795–804. http://doi.org/10.1001/archpsyc.61.8.795

Duncan, S., & Barrett, L. F. (2007). Affect is a form of cognition: A neurobiological analysis. *Cognition & Emotion*, *21*(6), 1184–1211. http://doi.org/10.1080/02699930701437931

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial Action Coding System (2nd ed.). Salt Lake City: A Human Face.

Eldar, E., Rutledge, R. B., Dolan, R. J., & Niv, Y. (2016). Mood as Representation of Momentum. *Trends in Cognitive Sciences*, *20*(1), 15–24. http://doi.org/10.1016/j.tics.2015.07.010

Epstude, K., & Roese, N. J. (2008). The Functional Theory of Counterfactual Thinking. *Personality and Social Psychology Review*, *12*(2), 168–192. http://doi.org/10.1177/1088868308316091

Etkin, A., Büchel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews Neuroscience*, *16*(11), 693–700. http://doi.org/10.1038/nrn4044

Finger, E. C., Mitchell, D. G. V., Jones, M., & Blair, R. J. R. (2008). Dissociable roles of medial orbitofrontal cortex in human operant extinction learning. *Neuroimage*, *43*(4), 748–755.

Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., & Badre, D. (2015). fMRI and EEG Predictors of Dynamic Decision Parameters during Human Reinforcement Learning. *Journal of Neuroscience*, *35*(2), 485–494. http://doi.org/10.1523/JNEUROSCI.2036-14.2015

Frey, R., Hertwig, R., & Rieskamp, J. (2014). Fear shapes information acquisition in decisions from experience. *Cognition*, *132*(1), 90–99. http://doi.org/10.1016/j.cognition.2014.03.009

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*. http://doi.org/10.2307/2246093

Greening, S. G., & Mitchell, D. G. V. (2015). A network of amygdala connections predict individual differences in trait anxiety. *Human Brain Mapping*, *36*(12), 4819–4830. http://doi.org/10.1002/hbm.22952

Hahn, A., Stein, P., Windischberger, C., Weissenbacher, A., Spindelegger, C., Moser, E., et al. (2011). Reduced resting-state functional connectivity between amygdala and orbitofrontal cortex in social anxiety disorder. *Neuroimage*, *56*(3), 881–889. http://doi.org/10.1016/j.neuroimage.2011.02.064

Haines, N., Southward, M. W., Cheavens, J. S., Beauchaine, T., & Ahn, W.-Y. (2019). Using computer-vision and machine learning to automate facial coding of positive and negative affect intensity. *PLoS ONE*, *14*(2), e0211735. http://doi.org/10.1371/journal.pone.0211735

Haines, N., Vassileva, J., & Ahn, W.-Y. (2018). The Outcome-Representation Learning Model: A Novel Reinforcement Learning Model of the Iowa Gambling Task. *Cognitive Science*, *47*(4), 1–28. http://doi.org/10.1111/cogs.12688

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive Reward Signals in the Anterior Cingulate Cortex. *Science*, *324*(5929), 948–950. http://doi.org/10.1126/science.1168488

Hedge, C., Powell, G., & Sumner, P. (2017). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *103*(3), 1–21. http://doi.org/10.3758/s13428-017-0935-1

Heilman, R. M., Crişan, L. G., Houser, D., Miclea, M., & Miu, A. C. (2010). Emotion regulation and decision making under risk and uncertainty. *Emotion*.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science*, *15*(8), 534–539. http://doi.org/10.1111/j.0956-7976.2004.00715.x

iMotions (2018). *Facial expression analysis: The definitive guide.* Retrieved from https://imotions.com/facialexpression-guide-ebook/ on 3/17/2018.

Inman, J. J., & Zeelenberg, M. (2002). Regret in Repeat Purchase versus Switching Decisions: The Attenuating Role of Decision Justifiability. *Journal of Consumer Research*, *29*(1), 116–128. http://doi.org/10.1086/339925

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*(2), 136–153. http://doi.org/10.1037/0033-295X.93.2.136

Kanluan, I., Grimm, M., & Kroschel, K. (2008). Audio-visual emotion recognition using an emotion space concept. Presented at the 16th European Signal Processing Conference, Lausanna, Switzerland.

Keltner, D., & Gross, J. J. (1999). Functional Accounts of Emotions. *Cognition & Emotion*, *13*(5), 467–480. http://doi.org/10.1080/026999399379140

Kocovski, N. L., Endler, N. S., Rector, N. A., & Flett, G. L. (2005). Ruminative coping and post-event processing in social anxiety. *Behaviour Research and Therapy*, *43*(8), 971–984. http://doi.org/10.1016/j.brat.2004.06.015

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. http://doi.org/10.1038/nn.2635

Larrick, R. P., & Boles, T. L. (1995). Avoiding Regret in Decisions with Feedback: A Negotiation Example. *Organizational Behavior and Human Decision Processes*, *63*(1), 87–97. http://doi.org/10.1006/obhd.1995.1064

Lecci, L., Okun, M. A., Karoly, P., 1994. (1994). Life regrets and current goals as predictors of psychological adjustment. *Journal of Experimental Psychology: General*.

Li, Jian, Schiller, D., Schoenbaum, G., Phelps, E. A., & Daw, N. D. (2011). Differential roles of human striatum and amygdala in associative learning. *Nature Neuroscience*, *14*(10), 1250–1252. http://doi.org/10.1038/nn.2904

Lohrenz, T., McCabe, K., Camerer, C. F., & Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proceedings of the National Academy of Sciences*, *104*(22), 9493–9498. http://doi.org/10.1073/pnas.0608842104

Loomes, G., & Sugden, R. (1982a). A rationale for preference reversal. *The American Economic Review*, *73*, 428–432.

Loomes, G., & Sugden, R. (1982b). Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. *The Economic Journal*, *92*(368), 805.

Loomes, G., & Sugden, R. (1986). Disappointment and Dynamic Consistency in Choice under Uncertainty. *The Review of Economic Studies*, *53*(2), 271–282. http://doi.org/10.2307/2297651

Ludvig, E. A., & Spetch, M. L. (2011). Of Black Swans and Tossed Coins: Is the Description-Experience Gap in Risky Choice Limited to Rare Events? *PLoS ONE*, *6*(6), e20262. http://doi.org/10.1371/journal.pone.0020262

Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2013). Extreme Outcomes Sway Risky Decisions from Experience. *Journal of Behavioral Decision Making*, *27*(2), 146–156. http://doi.org/10.1002/bdm.1792

Martinez, L. F., & Zeelenberg, M. (2014). Trust me (or not): Regret and disappointment in experimental economic games. *Decision*, *2*(2), 118–126. http://doi.org/10.1037/dec0000025

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision Affect Theory: Emotional Reactions to the Outcomes of Risky Options. *Psychological Science*, *8*(6), 423–429. http://doi.org/10.1111/j.1467-9280.1997.tb00455.x

Mellers, B., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology: General*, *128*(3), 332–345. http://doi.org/10.1037/0096-3445.128.3.332

Monroe, M. R., Skowronski, J. J., Macdonald, W., & Wood, S. E. (2005). The Mildly Depressed Experience More Post–Decisional Regret Than the Non–Depressed. *Journal of Social and Clinical Psychology*, *24*(5), 665–690. http://doi.org/10.1521/jscp.2005.24.5.665

Passamonti, L., Fairchild, G., Fornito, A., Goodyer, I. M., Nimmo-Smith, I., Hagan, C. C., & Calder, A. J. (2012). Abnormal Anatomical Connectivity between the Amygdala and Orbitofrontal Cortex in Conduct Disorder. *PLoS ONE*, *7*(11), e48789. http://doi.org/10.1371/journal.pone.0048789

Pieters, R., & Zeelenberg, M. (2005). On bad decisions and deciding badly: When intention–behavior inconsistency is regrettable. *Organizational Behavior and Human Decision Processes*, *97*(1), 18–30. http://doi.org/10.1016/j.obhdp.2005.01.003

Reed, L. I., DeScioli, P., & Pinker, S. A. (2014). The Commitment Function of Angry Facial Expressions. *Psychological Science*, *25*(8), 1511–1517. http://doi.org/10.1177/0956797614531027

Reed, L. I., Zeglen, K. N., & Schmidt, K. L. (2012). Facial expressions as honest signals of cooperative intent in a one-shot anonymous Prisoner's Dilemma game. *Evolution and Human Behavior*, *33*(3), 200–209. http://doi.org/10.1016/j.evolhumbehav.2011.09.003

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. … conditioning II: Current research and theory.

Rieskamp, J., Busemeyer, J. R., & Laine, T. (2003). How Do People Learn to Allocate Resources? Comparing Two Learning Theories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1066–1081. http://doi.org/10.1037/0278-7393.29.6.1066

Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, *66*(5), 805–818. http://doi.org/10.1037/0022-3514.66.5.805

Roese, N. J., & Summerville, A. (2005). What We Regret Most... and Why. *Personality and Social Psychology Bulletin*, *31*(9), 1273–1285. http://doi.org/10.1177/0146167205274693

Roese, N. J., Epstude, K., Fessel, F., Morrison, M., Smallman, R., Summerville, A., et al. (2009). Repetitive Regret, Depression, and Anxiety: Findings from a Nationally Representative Survey. *Journal of Social and Clinical …*, *28*(6), 671–688. http://doi.org/10.1521/jscp.2009.28.6.671

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences*, *111*(33), 12252–12257. http://doi.org/10.1073/pnas.1407535111

Saffrey, C., Summerville, A., & Roese, N. J. (2008). Praise for regret: People value regret above other negative emotions. *Motivation and Emotion*, *32*(1), 46–54. http://doi.org/10.1007/s11031-008-9082-4

Sharpe, M. J., & Schoenbaum, G. (2016). Back to basics: Making predictions in the orbitofrontal–amygdala circuit. *Neurobiology of Learning and Memory*, *131*, 201–206. http://doi.org/10.1016/j.nlm.2016.04.009

Shimanoff, S. B. (1984). Commonly named emotions in everyday conversations. *Perceptual and Motor Skills, 58*, 514. doi:10.2466/pms.1984.58.2.514

Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (2013). A Comparison of Reinforcement Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning. *The Journal of Problem Solving*, *5*(2). http://doi.org/10.7771/1932-6246.1150

Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (2014). Absolute performance of reinforcement-learning models for the Iowa Gambling Task. *Decision*, *1*(3), 161–183. http://doi.org/10.1037/dec0000005

Turner, B. M., Forstmann, B. U., Love, B. C., Palmeri, T. J., & Van Maanen, L. (2017). Approaches to analysis in model-based cognitive neuroscience. *Journal of Mathematical Psychology*, *76*, 65–79.

Turner, B. M., Rodriguez, C. A., Norcia, T. M., McClure, S. M., & Steyvers, M. (2016). Why more is better: Simultaneous modeling of EEG, fMRI, and behavioral data. *Neuroimage*, *128*, 96–115. http://doi.org/10.1016/j.neuroimage.2015.12.030

Ungemach, C., Chater, N., & Stewart, N. (2009). Are Probabilities Overweighted or Underweighted When Rare Outcomes Are Experienced (Rarely)? *Psychological Science*, *20*(4), 473–479. http://doi.org/10.1111/j.1467-9280.2009.02319.x

Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 1–20. http://doi.org/10.1007/s11222-016-9696-4

Wetzels, R., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2010). Bayesian parameter estimation in the Expectancy Valence model of the Iowa gambling task. *Journal of Mathematical Psychology*, *54*(1), 14–27. http://doi.org/10.1016/j.jmp.2008.12.001

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, *144*(2), 140–176. http://doi.org/10.1037/bul0000115

Yechiam, E., & Busemeyer, J. R. (2008). Evaluating generalizability and parameter consistency in learning models. *Games and Economic Behavior*, *63*(1), 370–394. http://doi.org/10.1016/j.geb.2007.08.011

Yechiam, E., & Ert, E. (2007). Evaluating the reliance on past choices in adaptive learning models. *Journal of Mathematical Psychology*, *51*(2), 75–84. http://doi.org/10.1016/j.jmp.2006.11.002

Yechiam, E., & Rakow, T. (2012). The Effect of Foregone Outcomes on Choices From Experience. *Experimental Psychology*, *59*(2), 55–67. http://doi.org/10.1027/1618-3169/a000126

Zeelenberg, M., & PIETERS, R. (2007). A Theory of Regret Regulation 1.0. *Journal of Consumer Psychology*, *17*(1), 3–18. http://doi.org/10.1207/s15327663jcp1701_3

Zeelenberg, M., van Dijk, W. W., van der Pligt, J., Manstead, A. S. R., van Empelen, P., & Reinderman, D. (1998). Emotional Reactions to the Outcomes of Decisions: The Role of Counterfactual Thought in the Experience of Regret and Disappointment. *Organizational Behavior and Human Decision Processes*, *75*(2), 117–141. http://doi.org/10.1006/obhd.1998.2784