Negative Affect Induces Rapid Learning of Counterfactual Representations: A Model-based Facial Expression Analysis Approach

Nathaniel Haines[1], Olga Rass[2], Yong-Wook Shin[3], Joshua W. Brown[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 (co-corresponding authors):

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

Nathaniel Haines

Department of Psychology

The Ohio State University

Columbus, OH 43210

Tel: +1-614-292-3155, Fax: +1-614-292-6798. Email: haines.175@osu.edu

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

Abstract

Counterfactual emotions including regret and disappointment play a crucial role in how people make decisions. For example, people often behave such that their decisions minimize potential regret or disappointment and therefore maximize subjective pleasure. Importantly, functional accounts of emotion suggest that the experience and future expectation of counterfactual emotions should promote goal-oriented behavioral change. Although many studies find empirical support for such functional theories, the cognitive-emotional mechanisms through which counterfactual thinking facilitates changes in behavior remain unclear. Here, we leverage computational models of risky decision-making that extend regret and disappointment theory to experience-based tasks, which we use to determine how people learn counterfactual representations of their decisions across time. Further, we use computer-vision to detect positive and negative affect (valence) intensity from participants' faces in response to feedback, which we use to determine how experienced emotion may influence cognitive mechanisms of learning, reward sensitivity, or exploration/exploitation—any of which could result in functional changes in behavior. Using hierarchical Bayesian modeling and Bayesian model comparison methods, we found that: (1) people learn to explicitly represent and subjectively weight counterfactual outcomes with increasing experience, and (2) people update their counterfactual expectations more rapidly as they experience increasingly intense negative affect. Our findings support functional accounts of regret and disappointment and demonstrate the potential for computational modeling and model-based facial expression analysis to enhance our understanding of cognition-emotion interactions.

1. Introduction

1.1 Emotions and decision-making

The idea that emotions interact with our cognitions to impact decision-making can be traced back multiple centuries, and many prominent scholars have expanded on our understanding of the interplay between emotion and cognition throughout the years (e.g., Darwin, 1872/1998; James, 1884). However, it was not until the late 20th century that decision theorists began to widely acknowledge the importance of emotion in how people make decisions (for an extended history, see Lerner, Li, Valdesolo, & Kassam, 2015). Today, we know that emotions are central to arguably every decision we make—they underlie a wealth of basic decision-making processes including outcome evaluation (Davis, Love, & Todd Maddox, 2009), valuation (Heilman, Crişan, Houser, Miclea, & Miu, 2010), and exploration/exploitation (Frey, Hertwig, & Rieskamp, 2014). Further, the neural mechanisms that subserve emotion-regulation are inseparable from those responsible for general decision-making (Mitchell, 2011), which has led to novel computational theories of emotion-regulatory processes (Etkin, Büchel, & Gross, 2015). While traditional economic perspectives view emotion-driven decision-making as a once adaptive—yet now irrational—relic of human evolution (see Keltner & Learner, 2010), these recent accounts represent a more general shift toward conceptualizing emotions as fundamental, functional components of human cognition that help us make better inferences within complex, dynamic environments (Eldar, Rutledge, Dolan, & Niv, 2016).

1.2 The functional role of counterfactual emotions

Counterfactual thinking elicits multiple emotions that are crucial to decision-making. *Regret* is an example, which is characterized by an aversive emotional state that occurs when we make a choice and then later wish to have made an alternative choice (Kahneman & Miller, 1986). In

fact, regret is one of the most frequently experienced and discussed emotions in everyday situations (Shimanoff, 1984), and 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, many people look back on regretted decisions with appreciation, which is not true of other negative emotions like anger, jealousy, anxiety, guilt, and boredom (Saffrey, Summerville, & Roese, 2008). These seemingly paradoxical findings are reconciled by functional accounts of counterfactual thinking including decision justification theory and regret regulation theory (Connolly & Zeelenberg, 2002; Zeelenberg & Pieters, 2007), both which posit that regret facilitates goal-oriented behavioral change by signaling us to avoid making unjustified choices that lead to outcomes that could have been better had we made a different choice (see also Epstude & Roese, 2008; Roese, 1994). 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 forms of negatively-valenced counterfactual thinking lead to long-term behavioral changes, suggesting that regret serves a specific function in how we decide among multiple risky choices (e.g., Zeelenberg et al., 1998).

Disappointment is another example, which occurs when comparing an obtained outcome to what could have been obtained in an alternative state of the world (Kahneman & Miller, 1986). Like regret, disappointment is experienced as aversive, yet it is evaluated favorably in retrospect

(Saffrey, Summerville, & Roese, 2008). Individual differences in how children regulate experienced disappointment predicts behavioral problems including symptoms of attention-deficit/hyperactivity disorder and oppositional/aggressive behavior (Cole, Waxler, & Smith, 1994), and adults who experience more frequent disappointing life events tend to show higher rates of depression (Goodyer, Herbert, Tamplin, Secher, & Pearson, 1997). Unlike regret, disappointment appears to play a primary role in modifying how people evaluate situational outcomes or one-off events as opposed to promoting specific changes in behavior (e.g., Markman, Gavanski, Sherman, & McMullen, 1993; Zeelenberg et al., 1998). For example, experimental evidence suggests that disappointment can lead one to feel worse about outcomes they receive when they could have received better outcomes—and alternatively, people feel better about outcomes they receive when they could have received worse outcomes (e.g., Ordóñez, Connolly, & Coughlan, 2000). Therefore, the functional value of disappointment appears to be related to coping and reacting to situational outcomes that indirectly facilitate more long-term positive gains, whereas regret is more directly related to long-term changes in behavior.

While the observational and survey-based studies described above consistently implicate counterfactual emotions in goal-oriented behavioral change, the cognitive mechanisms responsible for such changes remain underexplored. In particular, it is unclear how people learn to represent counterfactual expectations as they gain experience so that they can make choices to minimize anticipated regret and disappointment. By determining the cognitive and emotional mechanisms involved in counterfactual 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.). Similarly, identifying such mechanisms may elucidate the functional role of disappointment, thus providing an explanation for prospective links between disappointing events and behavioral problems.

1.3 Regret and disappointment theory

Theories of decision-making including regret and disappointment theory offer a formal language to understand the functional role of counterfactual emotions. Regret theory encompasses a variety of computational models that describe how people incorporate expected regret into their decisions (Bell, 1982; Loomes & Sugden, 1982b). According to regret theory, counterfactual comparisons between the outcomes of chosen and foregone actions 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 competing theories (e.g., expected utility theory) while maintaining a simple form with relatively few assumptions compared to other models (Loomes & Sugden, 1982a; 1982b). An example is the correlation effect, which manifests as a change in expected-value maximization behavior when the payoffs across choices are correlated (Diederich & Busemeyer, 1999; Grosskopf, Erev, & Yechiam, 2006). Specifically, when outcomes are negatively correlated such that a good outcome following chosen option A coincides with a bad foregone outcome for option B (and *vice-versa*), people tend to choose the less-optimal choice more frequently relative to when the outcomes are positively correlated. Further, Erev et al. (2017) showed that the correlation effect is only apparent when participants are given feedback on both the chosen and foregone choice outcomes. Importantly, strong utility models such as prospect theory (Kahneman & Tversky, 1979)—which assume that people evaluate each option independently—fail to capture the correlation effect, whereas those derived

from regret theory naturally capture such effects by assuming that people incorporate information on the value of counterfactuals into the utility of each option.

Disappointment theory was later developed and incorporated into models of regret (Bell, 1985; Loomes & Sugden, 1986). Similar to regret, disappointment is conceptualized as the difference between possible states of the world. Unlike regret, disappointment does not involve the comparison of an alternative choice to the received outcome, but instead it is a comparison of an alternative state of the chosen option to the received outcome (e.g., option A could have resulted in a positive outcomes but instead resulted in a negative outcome). By doing so, disappointment theory offers predictions regarding how experienced outcomes should relate to experienced positive/negative affect (or valence)—specifically, valence is assumed to result from differences between experienced and expected choice outcomes, such that even strongly valenced outcomes (e.g., losing a large sum of money, getting a big promotion) are not predicted to induce strong emotional states if more extreme outcomes were expected (e.g., a $5,000 salary raise may be valued negatively if a $10,000 raise was expected). Therefore, models of disappointment (and its positive counterpart elation) can capture behavior consistent with commonly found choice heuristics such as win-say-lose shift rules (e.g., Worthy, Hawthorne, & Otto, 2013), where actions are increasingly (or decreasingly) preferred for leading to better (or worse) than expected outcomes.

1.4 Learning counterfactual representations

Counterfactual models have 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 imperative to develop and test models of regret/disappointment that generalize well across both description- and experience-based paradigms. Specifically, people act *as if* they overweight low probabilities when making description-based choices yet underweight low probabilities when making experience-based decisions. Therefore, despite regret/disappointment showing robust effects across a number of different experience-based decision paradigms (e.g., Avrahami, Kareev, & Hart, 2014; Hart, Avrahami, Kareev, Todd, 2015; Kareev, Avrahami, & Fielder, 2014; Marchiori & Warglein, 2008; Rakow, Newell, Wright, 2015), the extent to which our formal models of counterfactual thinking generalize across both description- and experience-based paradigms is less clear. The naïve sampler model is an example of an experience-based model used to explain regret (Erev & Roth, 2014), which assumes a memory trace sampling process that can mimic regret effects by comparing small samples of past outcomes to current outcomes. The Best Estimate and Sampling Tools model is another example (Erev et al., 2017), which includes a recency-weighted regret minimization heuristic and has been used to explain changes in choice preferences resulting from repeated decisions from description with feedback (see also Chen, Liu, Chen, & Lee, 2011). Importantly, such models deviate strongly from traditional decision-theoretic models of risk such as prospect theory, disappointment theory, and regret theory (e.g., Delquié & Cillo, 2006; Kahneman & Tversky, 1979; Loomes & Sugden, 1982; 1986) in that they do not describe individuals' choices in terms of "subjective functions" of values and probabilities because they are not explicitly represented in the model.

Differences between computational models used to explain counterfactual thinking for description- and experience-based decision-making are important to consider for two reasons: (1) functional theories of counterfactual thinking suggest that regret plays a role in goal-oriented behavioral change, which necessitates some form of experience-based learning; and (2) instantiations of disappointment and regret theory assume that counterfactual expectations are weighted by described probabilities, yet described probabilities are not available for the majority of decisions people make that could ostensibly lead to disappointment or regret (see Griffiths & Tenenbaum, 2006; Hertwig & Erev, 2009). For example, traditional theories of regret and disappointment do not include learning mechanisms, so they cannot account for changes in either choice preferences or experienced emotion intensity over time that are suggested by functional theories of counterfactual thinking (e.g., in designs where participants make repeated decisions from description with feedback). Therefore, an important goal of the current study is to develop a model consistent with functional theories of counterfactual thinking that can be used to explain goal-oriented changes in behavior over time.

1.5 Informing computational models with experienced emotion

Although computational models show strong promise for deepening our understanding of how emotions influence 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; Erev et al., 2014; 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,

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; Mellers et al., 1997; 1999; 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 model used in Coricelli et al. (2005) did not include a learning component, so it could not be used to identify the cognitive mechanisms responsible for changes in regret/disappointment expectations across time. Additionally, no external measures of emotion were included in the model.

Decision Affect Theory (DAT) is perhaps the most comprehensive model relating objective reports of emotion to expectations of regret and disappointment (Mellers et al., 1997; Mellers et al., 1999), although it was developed for description-based tasks and therefore does not include a learning mechanism to describe changes in preferences over time. DAT assumes that individuals make decisions by maximizing their subjective expected pleasure (SEP), which is formalized as the subjective utility of a choice option plus counterfactual expectations:

$$SEP \propto \text{ (Utility)} + \text{(Disappointment or Elation)} + \text{(Regret or Rejoice)} \quad (1)$$

Here, regret and rejoice both involve upward and downward counterfactual comparisons across possible choice alternatives (i.e. "*I am upset [glad] that I chose this option over the alternative.*"), whereas disappointment and elation reflect upward and downward counterfactual comparisons across alternative states of the world (i.e. "*I am upset [glad] that my choice turned*

out this way."). Mellers et al. (1999) used Equation 1 to describe relations between individuals' self-reported emotional response to feedback from description-based decisions (e.g., *how do you feel from 50 (extremely elated) to -50 (extremely disappointed)?* after a choice returns -$8 as opposed to +$8)[1]. Importantly, they showed that self-reported emotion intensity increased in proportion to how *surprising* the outcome was (where surprise is operationalized as one minus the objective probability of the chosen and/or counterfactual outcome). The association between experienced emotion and surprisingness is consistent with more recent work showing that individuals' self-reported happiness is most strongly predicted by the prediction error induced by the outcome, such that happiness is higher when the outcome is better than expected and *vice versa* (e.g., Rutledge et al., 2014).

Altogether, previous studies have clarified relations between cognitive mechanisms underlying counterfactual thinking and experienced emotions, but they leave open the question of *how* experienced emotion interacts with such cognitive mechanisms to change behavior over time. The lack of studies using measures of emotion to explain the dynamics of counterfactual thinking is a potentially important oversight given the sophistication of regret and disappointment theory, the functional role of experienced emotion in motivating behavioral change, and the recent movements toward linking traditionally separated components of cognition and emotion (Barrett, 2009; Duncan & Barrett, 2007; Eldar et al., 2016; Etkin et al., 2015; Pessoa, 2008; Pessoa & Adolphs, 2010).

---

[1] Note that DAT is primarily a model of emotional responses to outcomes, and not of choice per se. In fact, DAT produces unstable estimates when used to model choices alone (see footnote 10 of Mellers et al., 1999, p. 342)

1.6 The current study

Here, we use: (1) risky decision-making tasks that evoke counterfactual thinking; (2) computational, emotion-driven reinforcement learning models built on regret and disappointment theory; and (3) automated computer-vision coding of facial affect intensity (e.g., Haines et al., 2019) to determine how people learn counterfactual representations of their decisions that produce changes in behavior in response to feedback gained through experience. Computer-vision allows us to rapidly measure trial-by-trial positive and negative affect in response to feedback, which fills a critical gap in previous studies of counterfactual thinking; namely, prior studies have focused almost exclusively on the cognitive—but not the affective—components of counterfactual thinking, 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 (each formalized as a computational model) regarding both how counterfactual representations develop with experience and the how experienced emotion influenced cognitive mechanisms underlying counterfactual thinking.

Specifically, we develop models within the framework of traditional regret and disappointment theories (e.g., Ahn et al., 2012; Delquié & Cillo, 2006; Loomes & Sugden, 1982; 1986), which allows us to test assumptions underlying functional theories of counterfactual thinking. We test a novel model assuming that that individuals learn a representation of counterfactual outcomes within the choice environment as they gain experience with each outcome, which provides a bridge between both: (1) traditional expected utility models and counterfactual models, and (2) description- and experience-based task formats that offer

feedback. We hypothesize that this novel model will account for experience-based changes in behavior over and above traditional models of experience-based counterfactual thinking that track a running regret/disappointment expectation for each choice option. Lastly, we hypothesize that emotional facial expressions that people produce in response to feedback will represent either: (1) learning, (2) valuation, or (3) exploration/exploitation mechanisms, any of which could provide an explanation for how experienced emotion leads to lasting changes in behavior. Below, we describe two studies designed to test our hypotheses across a broad range of risky decision-making paradigms. In the first study, we use data collected from a large number of participants undergoing various repeated description-based decisions with full feedback to develop and test our models. In the second study, we apply the best model to a smaller set of 51 participants undergoing experience-based decisions with full feedback—a portion of which had their facial expressions in response to feedback recorded and analyzed using computer-vision (n = 31).

2. Study 1 Method

2.1 Participants

Our first study included a combined total of 686 participants' data collected across both the 2015 and 2018 Choice Prediction Competitions (Erev, Ert, Plonsky, Cohen, & Cohen, 2017; Plonsky, Erev, & Ert, 2017). We excluded trials where participants made choices that did not fit our task inclusion criteria (more detail below in *4.2 Behavioral Task*), but all participants were included.

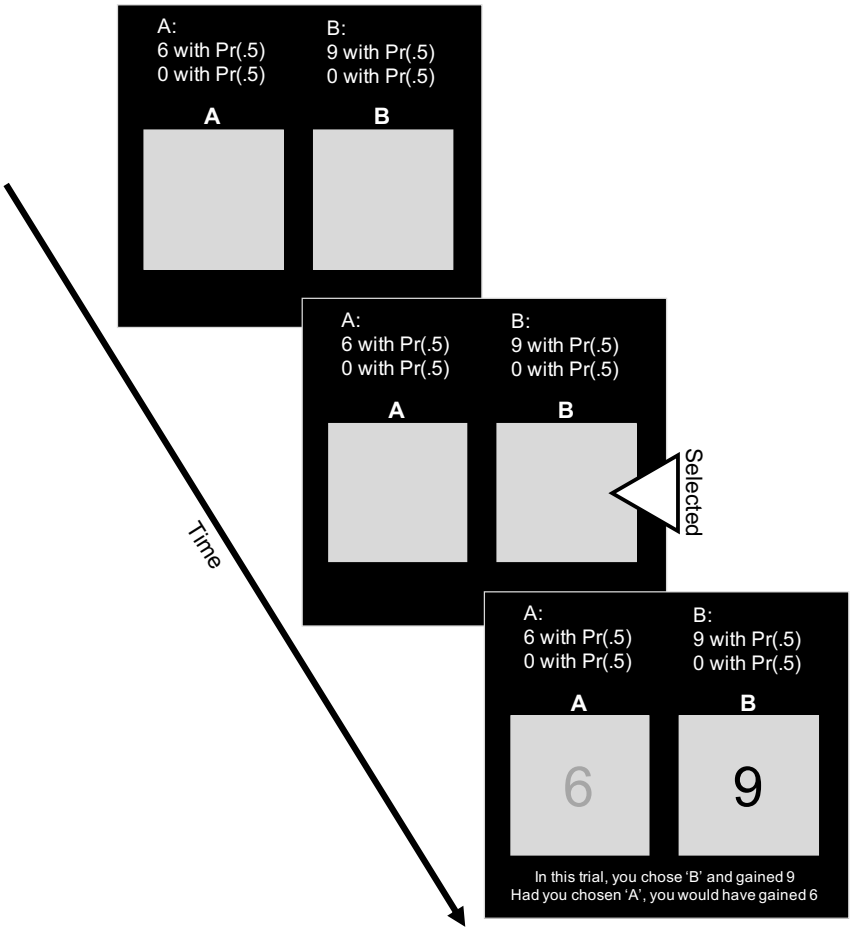
2.2 Behavioral task

In the Choice Prediction Competition data, all participants completed one of seven sets of 30 gambling games (for a total pool of 210 games) in randomized order. For the current study, we excluded games within each set containing more than two options or payoffs to restrict our focus to 2-choice paradigms. Additionally, we excluded games involving choice ambiguity, where payoffs/probabilities were only partially described. Altogether, these exclusions restricted the number of possible games to 105 (mean [SD] number of games played across participants = 14.9 [5.1]; range = 5-24) (see Supplementary Table 1 for a breakdown of included games).

Each game consisted of 25 trials, where the true payoffs and associated payoff probabilities were shown for each option before a choice was made on each trial (termed repeated description-based choices with full feedback). Each set included a mix of games using: (a) safe vs. risky options and risky versus risky options, and (b) positive and negative outcomes. Additionally, for the first 5 trials within each game, no feedback on choice outcomes was revealed; starting from the 6th trial, "full-information" feedback on both the chosen and forgone outcome was given (see Figure 1 for an example trial). Participants were instructed to make preference-based choices, and that they would be compensated using the real outcome of one randomly selected trial across trials and games.

Figure 1

Example of description-based decision with full information feedback



Note. In the description-based task, participants were presented with complete information on the outcome distribution of each choice, which remained constant within each game. For the first 5 trials, feedback (i.e. the results of the choice shown in the last panel) was not presented. Starting from the 6th trial, participants received "full information" feedback on the outcome of the chosen and foregone choice outcomes.

2.3 Description-based computational models

We fit a total of 5 different models designed to test different counterfactual learning mechanisms that participants may use to develop a representation of disappointment/regret over time. Each model assumes that participants make choices to maximize their subjective expected pleasure, which varies over time depending on the outcomes they observe with experience. The models are all equivalent in the context of pure description-based decisions with no feedback, and they also all share the same three free parameters: (1) a learning rate, (2) a utility shape parameter, and (3) a choice sensitivity (or inverse temperature) parameter. Therefore, the models only differ in how they assume that a representation of subjective expected pleasure evolves with experience over time. We describe each of the models in detail below, along with a theoretical rationale for why we included them.

Counterfactual Representation Learning Model. Of the models we tested, the Counterfactual Representation Learning model most closely resembles traditional decision-theoretic models of disappointment and regret (Loomes & Sugden, 1982;1986; Mellers et al. 1997; 1999). The subjective expected pleasure (SEP) for each choice option i on trial t is given by:

$$SEP_{it} = \sum_{j=1}^{n_i} p_{ij} m_{ijt} \tag{2}$$

where p_{ij} indicates the objective, described probability of outcome j for choice option i, $n_i$is the number of possible outcomes for choice option i, and m_{ijt} is the *modified utility*. The modified utility incorporates counterfactual information into the traditional utility function:

$$m_{ijt} = x_{ij}^{w} + D_{ijt} + R_{ijt} \tag{3}$$

Here, x_{ij} is the outcome corresponding to the probability (p_{ij}) in Equation 2, which is subjectively valued using a person-specific utility shape parameter w $(0 < w < 1.5)$.

Then, D_{ijt} is an *expected disappointment* term (or *elation*, depending on if it is negative or positive in value, respectively), which is a weighted sum of the difference between outcome j for choice option i and each possible alternative outcome *within* option i:

$$D_{ijt} = \sum_{k \neq j}^{n_i} W_{ikt}(x_{ij} - x_{ik})^{\omega} \tag{4}$$

Here, W_{ikt} is a dynamic "experience weight" for the counterfactual outcome (x_{ik}), and w is the same shape parameter from Equation 3. Note that Equation 4 deviates from classic disappointment theory by assuming that the counterfactual terms (x_{ik}) are weighted by their experience weights (W_{ikt}), as opposed to their objective probabilities (p_{ik}). Further, we assume that disappointment/elation is computed as the difference between pairs of potential outcomes within options (see Delquié & Cillo, 2006; Mellers et al., 1997; 1999), which is different from classic models of disappointment that use the difference between each possible outcome and the expected value of the option (e.g., cf. Equation 4 above to Equation 1 in Loomes & Sugden, 1986). Importantly, if the counterfactual experience weight is equivalent to its objective probability (i.e. $W_{ikt} = p_{ik}$), then $D_{ijt} = 0$ and expected disappointment drops out of the modified utility term. Similarly, because we initialize all experience weights at 0 (details described below), there is no expected disappointment/elation before participants experience feedback on the outcomes of their choices. Therefore, Equation 4 produces a dynamic disappointment expectation that varies with experience until all experience weights converge to their true values. This behavior is similar to a win-stay-lose-shift heuristic (e.g., Worthy, Hawthorne, & Otto, 2013), where options that have more recently returned their best outcome

(i.e. a "win") are increasingly preferred and vice-versa. Note that it is possible to extend the model to include different valuation mechanisms for disappointment versus elation (e.g, a different ω parameter for when the difference in Equation 4 is negative versus positive, respectively), which can lead to a non-zero disappointment expectation even when all $W_{ikt} = p_{ik}$ (see Delquié & Cillo, 2006).

Finally, R_{ijt} is an *expected regret* term (or *rejoice*, depending on if it is negative or positive in value, respectively), which is a weighted sum of the difference between each possible outcome j of option i and each possible outcome k of the alternative option i':

$$R_{ijt} = \sum_{k=1}^{n_{i'}} W_{i'kt}(x_{ij} - x_{i'k})^{\omega} \tag{5}$$

Therefore, regret is similar to disappointment, but the counterfactual expectation is computed *across* options as opposed to *within* options. In Equation 5, we again deviate from classic regret theory by assuming that the regret weight is a dynamic experience weight (W_{ikt}) as opposed to the objective probability of the counterfactual outcome (cf. Equation 5 above to Equations 2-3 in Loomes & Sugden, 1982). Similar to disappointment, the trial-dependent experience weight allows for regret to vary with experience over time, which we describe in more detail below.

Lastly, we use a logistic (i.e. softmax) choice rule to transform subjective expected pleasure (SEP_{it}) into choice probabilities for each option:

$$\mathrm{Pr}_{it} = \frac{\exp(\theta \cdot SEP_{it})}{\sum_{k=1}^{2} \exp(\theta \cdot SEP_{it})} \tag{6}$$

where θ is a person-specific and trial-independent choice sensitivity (i.e. inverse temperature) parameter determined by $\theta = 3^c - 1$ that captures how deterministically versus randomly participants make choices according to the differences in SEP across options (Yechiam & Ert,

2007). Here, the free parameter is c $(0 < c < 5)$. In Equation 6, we only sum over two options given that all gambles in our data involve only two options.

A crucial component of the model is the set of counterfactual experience weights ($\boldsymbol{W}$; see Equations 4 and 5), which have strong effects on the behavior and interpretation of the model. For example, when $\boldsymbol{W} = 0$, both the regret and disappointment terms return 0, and the modified utility reduces to a basic expected utility model similar to prospect theory (albeit without loss aversion or the probability weighting parameters in the current implementation; Kahneman & Tversky, 1979). Conversely, if $\boldsymbol{W}$ are set to their corresponding objective outcome probabilities (i.e. all $W_{ikt} = p_{ik}$), then the model is akin to traditional regret theory (disappointment drops out, as described above), which is consistent with empirical evidence that regret dominates over disappointment and experienced reward with increasing experience (e.g., Kareev et al., 2014). Additionally, we assume that participants develop a cognitive representation of regret and disappointment (and their counterparts) as they gain experience/observe the outcomes of their choices. Such learning effects can be captured by assuming that $\boldsymbol{W} = 0$ before participants receive experience-based feedback, and that $\boldsymbol{W} \rightarrow \boldsymbol{p}$ as participants gain experience. We use the following delta learning rule (a simplified version of the Rescorla-Wagner updating rule; Rescorla & Wagner, 1972) to formalize this learning process:

$$W_{ij(t+1)} = \begin{cases} W_{ijt} + \alpha(I_{ijt} - W_{ijt}), & \text{if feedback given} \\ W_{ijt}, & \text{otherwise} \end{cases} \tag{7}$$

Here, α is a learning rate determining how rapidly participants update the experience weight after observing feedback on their choice, and I_{ijt} is an indicator variable that returns 1 if an outcome was observed during feedback and 0 otherwise. We chose to weight regret by its learned 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 subjectively likely. Additionally, our previous work shows that learning rules that explicitly update outcome probabilities can capture differences between description- and experience-based tasks (Haines, Kvam, & Turner, under review)—both of which we use in the current study.

Under the updating scheme described by Equation 7, the different outcomes x_{ijt} can be thought of as features that are either observed or not after each choice, similar to category learning models (e.g., Turner, 2019). When they are observed (or not observed), their experience weights increase toward 1 (or decrease toward 0) in proportion to the learning rate and the prediction error. Therefore, experience weights converge toward the probability of observing each outcome, but they fluctuate trial-to-trial based on the recent history of outcomes. Note that although experience weights converge toward outcome probabilities, they are not proper probabilities as they do not sum to 1 (akin to the "decision weights" in prospect theory; Kahneman & Tversky, 1979). For intuition, imagine a participant making a choice between option A ("$3 with Pr(1)") and option B ("$4 with Pr(.8), 0 otherwise"). If they choose A and are subsequently given feedback that A returned $3 and B returned $4, they increase experience weights corresponding to $3 and $4 and decrease the experience weight of $0. An important implication of this updating process is that a given counterfactual experience (e.g., receiving $3 when the alternative choice results in $4) does not need to be fully experienced in order for it to be expected to occur in the future—instead, the experience weights of individual outcomes comprise what possible counterfactual outcomes are expected. In this way, the model can capture effects observed in tasks using partial feedback (i.e. where only the chosen option outcome is

revealed), which lead to slower changes in behavior relative to full information feedback tasks (e.g., Rakow et al., 2015).

In the context of the task we used, feedback is not given for the first 5 trials, so there is no updating and all experience weights therefore remain at 0. Starting from the 6th trial, the outcomes are shown for both chosen and unchosen options after each choice. Therefore, using equation 6, we assume that changes in choice behavior starting from the 6th trial result from the effects of participants learning to represent the probability of outcomes over time as they experience each outcome. This counterfactual learning process is similar in kind to reversal/fictive learning models (e.g., Gläscher, Hampton, & O'Doherty, 2009; Lohrenz, McCabe, Camerer, & Montague, 2007), although our proposed mechanism is more general due to its explicit representation of outcome probabilities which allows for a much closer correspondence to classic regret/disappointment theory (see Loomes & Sugden, 1982; 1986).

Disappointment Minimization Learning Model. The Disappointment Minimization Learning model assumes that subjective expected pleasure (SEP) for each choice option i on trial t is given by:

$$SEP_{it} = \sum_{j=1}^{n_i} p_{ij} x_{ij}^{w} + DIS_{it} \tag{7}$$

where the terms in the summation are the same as described above (see p_{ij} and x_{ij}^{ω} in Equations 2 and 3), and DIS_{it} is a trial-dependent disappointment expectation that is determined according to the following delta updating rule:

$$DIS_{i(t+1)} = \begin{cases} DIS_{it} + \alpha(\text{sign}(\text{obs}_{it} - \text{cf}_{it}) - DIS_{it}), & \text{if fb given, } j > 1 \\ DIS_{it}, & \text{otherwise} \end{cases} \tag{8}$$

Here, we formalize disappointment as the sign of the difference between the observed outcome for each option (obs_{it}) and the alternative, counterfactual outcomes (cf_{it}) that could have occurred if the choice turned out differently. For example, if option A can return either -$4 or -$12, and -$4 is observed on the current trial, disappointment for option A is equal to $\mathrm{sign}(\mathrm{obs}_{it} - \mathrm{cf}_{it}) = \mathrm{sign}([-4] - [-12]) = 1$ and DIS_{it} then increases toward 1 in proportion to the learning rate α (indicating greater expectation for *elation*) and prediction error. Similar to the experience weights in the Counterfactual Representation Learning model, we initialize all the expectations to 0 at the start of each game ($DIS_{i1} = 0$). Further, SEP values are entered into the same softmax choice rule described by Equation 6, along with the persons-specific choice sensitivity parameter (θ).

Equation 8 produces an increasing preference for options that tend to return their highest outcome (i.e. the best outcome within an option), thus minimizing the probability of experiencing disappointment. Note that updating only occurs for a given option i if it can return more than 1 potential outcome (i.e. $j > 1$) and if feedback (fb) is presented on the current trial, indicating that disappointment cannot occur when an outcome is guaranteed. This behavior is similar to other reinforcement learning models that track the expected win frequency of each option, which leads to an increased preference for options that win most frequently (see Equations 7-9 from Haines, Vassileva, & Ahn, 2018). It also produces behavior consistent with win-stay-lose-shift heuristics (e.g., Worthy, Hawthorne, & Otto, 2013).

Regret Minimization Learning Model. The Regret Minimization Learning model assumes that SEP is given by:

$$SEP_{it} = \sum_{j=1}^{n_i} p_{ij} x_{ij}^{w} + REG_{it} \tag{9}$$

where REG_{it} is a trial-dependent regret expectation that is similar DIS_{it}, but is computed *across* options as opposed to *within* options:

$$REG_{i(t+1)} = \begin{cases} REG_{it} + \alpha(\text{sign}(\text{obs}_{it} - \text{obs}_{i't}) - REG_{it}), & \text{if fb given} \\ REG_{it}, & \text{otherwise} \end{cases} \quad (10)$$

Here, we formalize regret as the sign of the difference between the observed outcomes across options. Like the models described above, we initialize all regret expectations to 0 at the start of each game ($REG_{i1} = 0$), and we use the softmax choice rule with choice sensitivity parameter (Equation 6) to transform SEP into choice probabilities. Therefore, Equation 10 produces learning dynamics that lead to an increasing preference for the option that most frequently returns the best outcome on a trial-to-trial basis, thus minimizing the probability of experiencing regret. Such behavior is similar to models including the Best Estimate and Sampling Tools model, which minimizes immediate regret through a sampling mechanism that recalls the outcomes of each option on a past trial and prefers the one with the best outcome (see Equation 1 from Erev et al., 2017). Equation 10 differs in that it learns a regret expectation (i.e. it integrates over the recent history of trials) as opposed to relying on a single previously experienced trial.

Disappointment/Regret Minimization Learning Model. The Disappointment/Regret Minimization Learning model combines the Disappointment and Regret Minimization models described above, such that SEP is given by the subjective utility of each option plus the trial-dependent DIS and REG expectations. Therefore, it is most similar to the Counterfactual Representation Learning model, in that both disappointment and regret expectations are assumed to develop with experience as outcomes are observed. However, it is different from the Counterfactual Representation Learning model in that it does not explicitly represent each

outcome probability to compute counterfactual expectations—instead, expectations are the recency-weighted average of past realized counterfactual comparisons (see Equations 8 and 10).

Expectation Maximization Learning Model. The Expectation Maximization Learning model functioned as a baseline model for the current study. We included it to determine if the counterfactual disappointment and regret learning models offered a better account of changes in participants' preferences with experience relative to simple recency effects where each option and outcome is valued independently. Specifically, SEP is given by:

$$SEP_{it} = \sum_{j=1}^{n_i} p_{ij} x_{ij}^{w} + EV_{it} \tag{11}$$

where EV_{it} is a trial-dependent expected value term that is updated using the following delta rule:

$$EV_{i(t+1)} = \begin{cases} EV_{it} + \alpha(\text{obs}_{it}^{\omega} - EV_{it}), & \text{if fb given} \\ EV_{it}, & \text{otherwise} \end{cases} \tag{12}$$

Here, EV_{it} is updated toward the utility of the observed outcome for each option i, which produces increases in preference for the option that shows a higher average subjective value across recent trials. We use the same softmax choice rule described by Equation 6 to transform SEP into choice probabilities. Therefore, the Expectation Maximization Learning model is similar to models such as the naïve sampler model and other basic reinforcement learning models (see Erev & Roth, 2014), which tend to prefer options that return the best average payoff within a small sample of recently experienced outcomes.

2.4 Model fitting procedure

We fit all models using hierarchical Bayesian analysis (HBA), with separate hierarchical models fit to each of the 7 sets of games. We chose to fit separate hierarchies to each set for

mostly computational purposes[2], but also because each set consisted of different participants and games. HBA allows for individual-level (i.e. participant-level) parameter estimation while simultaneously pooling information across participants 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), 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. 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 inverse Probit-transformed (i.e. the cumulative distribution function of the standard normal) to the constrained space to maximize

[2] A single model fit across all 686 participants would require more computer RAM that we have available for making posterior predictions (e.g., participants × games × trials × MCMC samples).

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 for each of the models was parameterized as:

$$\begin{aligned} \mu_\alpha &\sim \text{Normal}(0, 1) \\ \sigma_\alpha &\sim \text{Half-Normal}(0, 0.2) \\ \boldsymbol{\alpha}' &\sim \text{Normal}(0, 1) \\ \boldsymbol{\alpha} &= \text{Probit}^{-1}(\mu_\alpha + \sigma_\alpha \cdot \boldsymbol{\alpha}') \cdot 1.0 \end{aligned} \tag{13}$$

Here, μ_α $(-\infty < \mu_\alpha < +\infty)$ and σ_α $(0 < \sigma_\alpha < +\infty)$ are the group-level mean and standard deviation, respectively, and bold terms indicate vectors of individual-level parameters. This parameterization assumes that the prior distribution over each individual-level parameter is near-uniform, and that variance is relatively low across participants (which provides regularization of extreme individual-level deviations toward the group-level mean). Since $0 < \alpha < 1$, the scaling factor was set to 1.0 in the example above. We used the same non-centered parameterization for ω and c (including the same prior distributions), except they were scaled by 1.5 and 5.0, respectively.

We ran all models for 1,500 iterations across 4 separate sampling chains, with the first 500 samples as warm-up (analogous to burn-in in Gibbs samplers) for a total of 4,000 posterior samples for each parameter. For all models, 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.

2.5 Model comparison

We used two different methods to compare models, namely: (1) penalized model fit, which is a statistical measure of how accurately a model can predict participants' choices on the next trial given their fitted model parameters, choice history, and a penalty term for model complexity; and (2) posterior predictive simulations, which involve graphically comparing the difference between participants' entire choice histories and choice histories simulated from their fitted model parameters to determine which model best accounts for changes in behavior across time. We used multiple methods because previous studies consistently show that different model comparison methods can lead to different conclusions, and because fit statistics alone (e.g., mean squared error, AIC, BIC, etc.) do not always discern which model can best capture theoretically relevant patterns in the data (Ahn et al., 2008; 2014; Haines, Vassileva, & Ahn, 2018; Navarro, 2019; Steingroever et al., 2014; Yechiam & Ert, 2007).

2.5.1 Penalized model fit. We assessed penalized model fit 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 (LPPD) of a fitted model. To compute the LPPD of each competing model, we calculated the log likelihood of each participants' true choice on trial $t + 1$ conditional on their parameter estimates and choice history (i.e. trials $\in \{1,2, \ldots , t\}$). Therefore, LOOIC compliments the posterior predictive simulations described above, as it provides a quantitative measure of how well each model can capture participants' trial-to-trial choices while accounting for parameter uncertainty. The log likelihood was calculated for each posterior sample and summed within each participant across games (preserving all posterior samples), resulting in an $S \times N$ LPPD matrix where S and N are the numbers of posterior samples and

participants, respectively. We used the *loo* R package, which is developed by the Stan team (Vehtari et al., 2016), to compute LOOIC values from the LPPD matrix. Note that LOOIC is on the deviance scale where lower values indicate better model fit (including complexity penalization).

2.5.2 Posterior predictive simulations. Posterior predictive simulations are similar to the post-hoc absolute fit measure used in previous studies (e.g., Steingroever et al., 2014; Steingroever, Wetzels, & Wagenmakers, 2013), but they differ in that they convey uncertainty in simulated choices that is attributable to the underlying parameter estimates. Further, they differ from methods such as the absolute simulation method, which uses estimated parameters to simulate behavior on the task without conditioning on participants' actual trial-level choices (e.g., Haines, Vassileva, & Ahn, 2018; Steingroever et al., 2014; Steingroever, Wetzels, & Wagenmakers, 2013). We generated posterior predictions by first fitting each model to participants' choice data, followed by simulating participants' trial-to-trial choices using the full joint posterior distribution of their fitted model parameters and by conditioning on their actual trial-level choices. Note that we fit a different hierarchical model to each set of games, but parameters were fixed within-participants across games within each set. We then averaged across participants and compared the posterior predictive simulations and the true across-participant average choice proportions for each of the 105 games using graphical measures. Our use of posterior predictive simulations allows for us to determine if each model can capture changes in participants' observed behavior in response to the feedback that they start to experience on the 6th trial of each game, which could be obscured if we used fit statistics alone to conduct model comparison (e.g., Navarro, 2019). For example, a model could exhibit relatively good fit statistics but fail to capture learning effects observed in choice rates across trials. Posterior

predictive simulations therefore allowed us to identify specific areas of model misfit and determine whether models captured the effects of feedback on changes in behavior across trials.

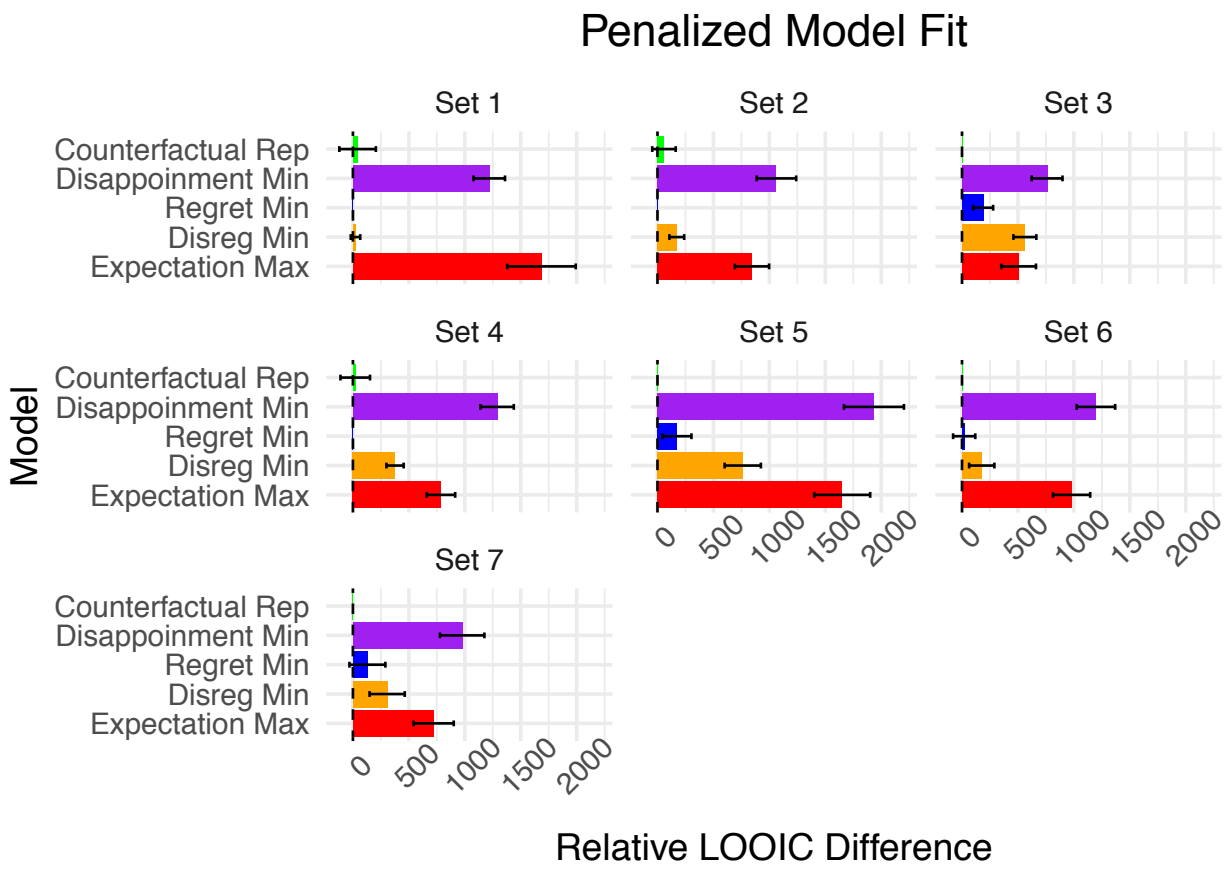
3. Study 1 Results

3.1 Model comparison: Penalized model fit.

The Counterfactual Representation and Regret Minimization Learning models showed the best penalized model fit statistics across all 7 sets, and the Counterfactual Representation Learning model performed best overall (see Figure 2). These results corroborate our posterior predictive simulations.

Figure 2

Penalized model fit in the context of description-based games



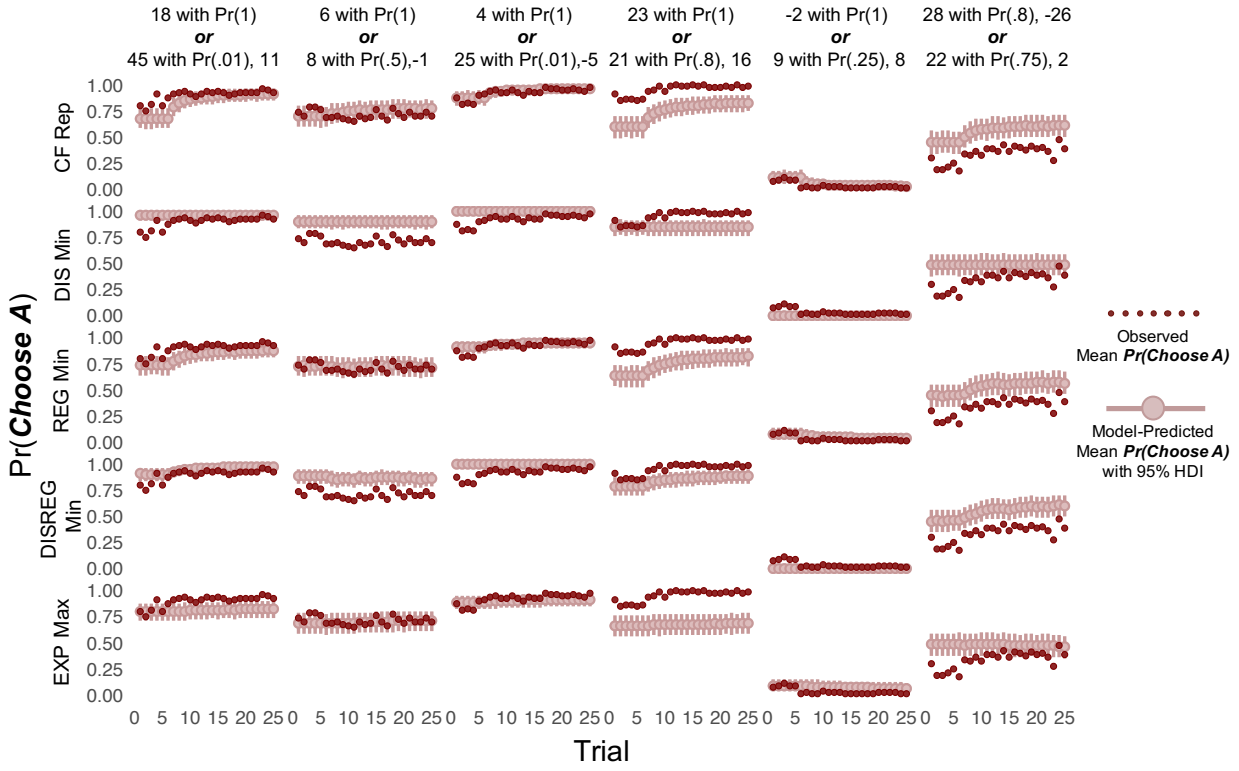
Note. Leave-one-out information criterion (LOOIC) scores relative to the best fitting model within each set of games. Lower LOOIC values indicate better model fit accounting for model complexity. Error bars represent ± 1 standard error of the difference between the best fitting model in the set and respective competing models.

3.2 Model comparison: Posterior predictive simulations.

Figure 3 shows both the true choice rates and posterior predictive simulations averaged across participants for 6 representative games within set 5. We only show these particular games for brevity, but we summarize our observations across all sets and games here (see Figure's S1-20 in the Supplementary Text for all 105 games). Overall, the Counterfactual Representation Learning, Regret Minimization Learning, and Disappointment/Regret Minimization Learning models best captured changes in participants' behavior across trials, indicating that regret expectations play a crucial role in facilitating changes in behavior in response to full information feedback. Conversely, the Disappointment Minimization and Expectation Maximization models poorly tracked the observed choice proportions across trials, and in fact often showed little to no changes in predictions in response to feedback across games despite participants showing strong changes in preferences. Of the models containing regret terms, across all games, the Counterfactual Representation Learning model more rapidly converged toward participants' observed changes in choice behavior, indicating that people may explicitly represent and update counterfactual outcome probabilities (i.e. the counterfactual "experience weight" in Equation 7) as opposed to tracking a running expectation of regret and disappointment. Explicit representation of counterfactual outcomes and probabilities is consistent with traditional models of regret and disappointment (Loomes & Sugden, 1982; 1986), along with empirical evidence that people make counterfactual comparisons between pairs of different possible outcomes (Ordóñez, Connolly, & Coughlan, 2000).

Figure 3

Posterior predictive simulations for a subset of games



Note. Posterior predictive simulations for a representative sample of games from set 5 (see Table S1 for information all sets and games, and Figure's S1-20 for posterior predictive simulations for all sets and games). Different games are shown across columns (labels describing the game), and the competing models are shown across rows. Here, dark red points indicate the mean choice proportions across participants, and lighter red points with intervals represent the mean of the posterior predictive simulations across participants with 95% highest density intervals (HDI). CF Rep = Counterfactual Representation Learning model; DIS Min = Disappointment Minimization Learning model; REG Min = Regret Minimization Learning model; DISREG Min = Disappointment/Regret Minimization Learning Model; EXP Max = Expectation Maximization Learning model.

Although the models with regret terms performed well for most games, there were some notable exceptions. Specifically, the models tended to underestimate preferences for games wherein one choice option always returned a better outcome than the alternative option. For example, Figure 2 shows a game where participants chose between option A (23 with Pr(1)) or option B (21 with Pr(.8), otherwise 0). Here, option A is always the best choice—it has the highest expected value (albeit by a small amount), and it will never result in either disappointment or regret. As expected, participants almost exclusively choose option A. However, the models all show a much less pronounced preference for option A (although it is still preferred over option B on average). We observed this pattern across many games, all sharing the feature that one option will always return a better outcome than the alternative. These patterns are consistent with the well-known *certainty effect*, wherein certain outcomes (i.e. those with Pr(1)) are given higher "decision weights" relative to uncertain outcomes (e.g., Allais, 1953). Importantly, the certainty effect can be accounted for using probability weighting functions such as those proposed by prospect theory (Kahneman & Tversky, 1979). Therefore, although we did not include a probability weighting function in our models, such an extension could potentially allow the models to better capture certainty effects (e.g., replacing p_{ij} in Equation 2 with $W(p_{ij})$, where $W(.)$ is a non-linear probability weighting function).

4. Interim Discussion

Our first study revealed that participants' changes in behavior in response to full information feedback during repeated description-based decisions are best explained by a model assuming that participants learn to represent the probability of counterfactual events (i.e. "experience weights" for disappointment, regret, and their counterparts elation and rejoice) as they gain

experience with each possible outcome. This "feature-based" account of counterfactual learning (where different numerical outcomes are conceptualized as different features that are observed or not after a choice) contrasts the competing "running expectation-based" models that we tested, all which assume that counterfactual events are integrated into a single value signal for each choice option that is learned with experience. Overall, our results are consistent with traditional disappointment and regret theory—both which assume explicit representations of the values and probabilities of counterfactual outcomes (Loomes & Sugden, 1982; 1986). More generally, our results are consistent with prior model-based findings that learning and decision-making in a number of games are driven strongly by regret, though our results go beyond earlier work in explicitly modeling the contributions of disappointment, regret, and expected value within a single model (Marchiori & Warglien, 2008).

Although findings from our first study offered insight into the cognitive mechanisms responsible for counterfactual thinking, the models did not incorporate experienced emotion, which limits our ability to test functional theories of counterfactual thinking as described in the introduction. Further, the task from study 1 gave participants complete information on the probabilities of various outcomes for each choice, yet real-world decisions are often made without given knowledge of how likely various outcomes are. Therefore, in study 2 we: (1) determine whether the models tested in study 1 can capture pure experience-based decisions (where probabilities and outcomes are not given and must be learned), and (2) identify how participants' real-time emotional experiences interact with cognitive mechanisms to facilitate goal-oriented changes in behavior.

5. Study 2 Method

5.1 Participants

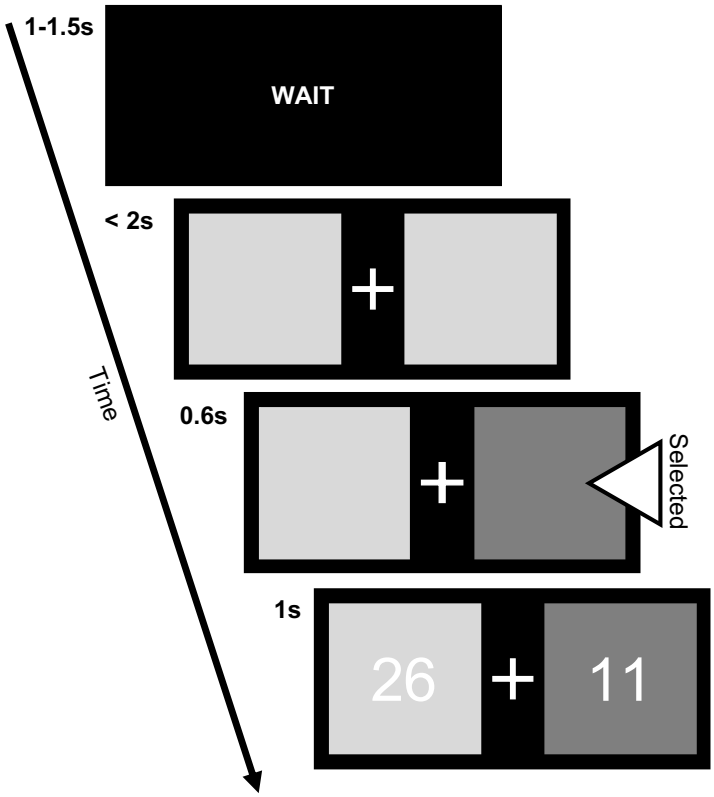
Our second study included 51 participants' data collected from two different research sites. Of these participants, 31 had facial expressions recorded, while 19 are from a previous study which did not record facial expressions (Ahn et al., 2012). We aggregated both datasets for the purposes of the current study. All participants gave informed consent prior to participating in the study. The study protocol was approved by local Institutional Review Boards at both research sites.

5.2 Behavioral task

All participants completed four separate gambling games in randomized order. Participants 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 participants were asked to choose between two options (see Figure 4). 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 participants but remained fixed within games. The probability of winning a high number of points for the risky option was fixed but unknown, and participants 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 feedback"), which allowed participants to make counterfactual comparisons between the choices they made and could have made. Participants were instructed to make choices that maximized their points. Unbeknownst to participants, the expected value for each option was identical (see Table 1 for payoff distributions for each game).

Figure 4

Behavioral task time course



Note. Unlike study 1, the task from study 2 did not describe the outcomes and associated probabilities of each choice to participants. However, the tasks are similar in that they both offer full information feedback after a choice is selected.

Table 1

Payoff structure for each game in Study 2

Game	Option A (safe)				Option B (risky)			
	A_H	(Pr_H)	A_L	(Pr_L)	B_H	(Pr_H)	B_L	(Pr_L)
1	12	(1)	–	–	56	(.2)	1	(.8)
2	11	(1)	–	–	26	(.4)	1	(.6)
3	10	(1)	–	–	16	(.6)	1	(.4)
4	9	(1)	–	–	11	(.8)	1	(.2)

Note. A_H = high outcome for option A. A_L = low outcome for option A (same notation for option B). Pr_H = high outcome for given option. Pr_L = low outcome for given option. Note that outcomes and their corresponding probabilities were not described to participants. Instead, participants had to learn from experience-based feedback.

5.3 Experience-based computational models.

We applied the same five models described in study 1 (see section 2.3) to data collected from the pure experience-based paradigm used in study 2 (e.g., Figure 4). However, because the models in section 2.3 assume that participants have access to the outcome values and probabilities of each choice option before receiving any feedback (e.g., Figure 1), we modified each of the models to reflect the absence of described outcome information in the experience-based paradigm. We describe these modifications below.

For the Counterfactual Representation Learning model, the described probabilities only appear in Equation 2, wherein they function to weight the modified utility term. Importantly, because the model explicitly represents probabilities in the form of "experience weights" (i.e. $\boldsymbol{W}$ in Equation 7), the described probability from Equation 2 can simply be replaced with the subjective experience weight to extend the model to pure experience-based decision-making:

$$SEP_{it} = \sum_{j=1}^{n_i} W_{ijt} m_{ijt} \tag{14}$$

Here, W_{ijt} is the experience weight for the corresponding outcome x_{ij} (see Equation 3). All other aspects of the model are identical to the description-based version (including the updating rule and initializing all $\boldsymbol{W} = 0$). Therefore, whereas the description-based version reduces to a standard prospect theory like model before any feedback, the experience-based version has a uniform preference for all options before feedback because all experience weights start at 0 (and SEP is thus 0 for all options).

For the remaining models, since they do not explicitly represent the probability of each outcome occurring, we dropped the terms containing the described probability (p_{ij}) from each model. Using the Disappointment Minimization Learning model as an example, we dropped the

summation term from Equation 7, such that $SEP_{it} = DIS_{it}$. We did the same for all other models. Therefore, each model assumes that SEP is given by a recency weighted average over past counterfactual experiences (except for the Expectation Maximization Learning model, which assumes that the SEP of each option is given by a recency weighted average of observed outcomes for that option).

5.4 Model-based facial expression analysis

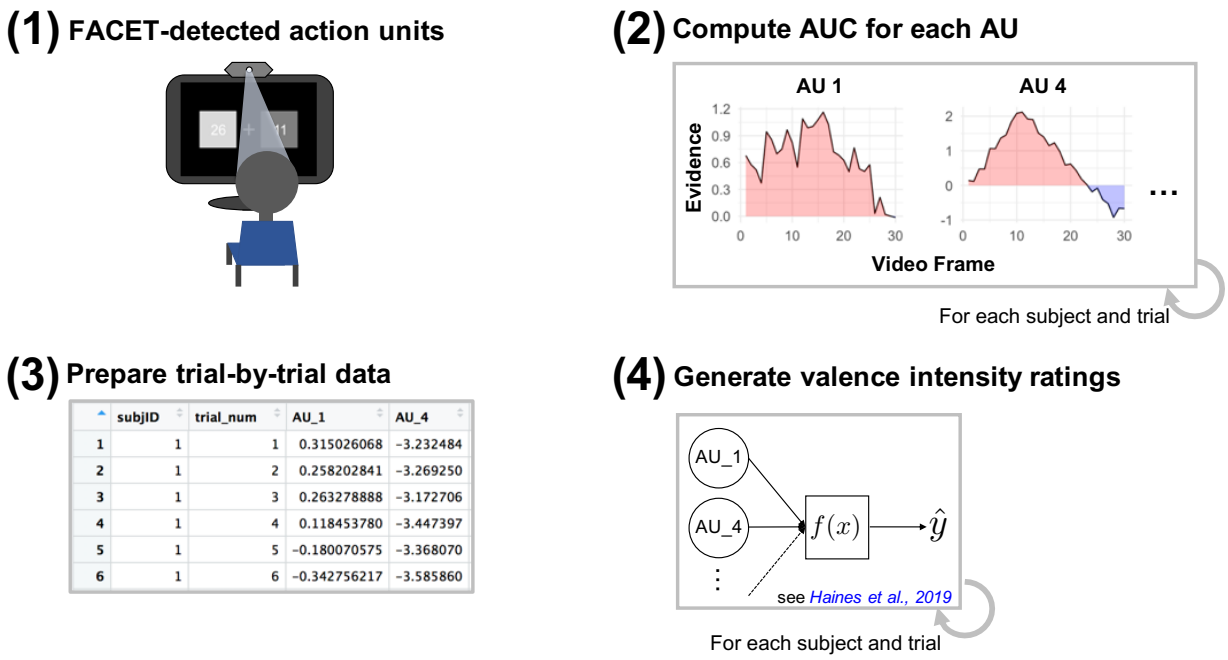
5.4.1 Automated facial expression coding. To measure the valence of participants' facial expressions during feedback, we used an automated facial expression coding (AFEC) model that we developed in a previous study (Haines et al., 2019). The AFEC model was trained to code for positive and negative affect 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 validated. In our validation study, the model showed correlations with human observer ratings of .89 and .76 for positive and negative affect intensity, respectively (for more details, see Haines et al., 2019).

Figure 5 shows the steps used to preprocess and apply the AFEC model to our participants' facial expressions in response to outcome feedback. 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 participant’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 participant (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 5

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) Participants' 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 participants' 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 participant. (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).

5.4.2 Incorporating facial expressions into computational models. To determine whether emotional facial expressions reflected choice mechanisms of learning (α), reward sensitivity (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 participant 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 participants, 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:

$$\alpha(t) = logit^{-1}(\alpha_0 + \alpha_1 \cdot FE(t)) \cdot 1.0 \tag{15}$$

Here, α_0 and α_1 indicate the baseline learning rate and effect of facial expression valence intensity on the learning rate for current trial ($\alpha(t)$), 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 $0 < \alpha(t) < 1$, 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 5.0, respectively. On trials where participants' 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:

$$\alpha(t) = logit^{-1}(\alpha_0) \cdot 1.0 \tag{16}$$

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.

5.5 Model fitting procedure

For the models incorporating facial expressions, we used an alternative parameterization for the parameter of interest to reflect the changes described in Equations 15 and 16. Using the model assuming that facial expression intensity reflected changes in the learning rate as an example, priors on the modified learning rates were set as follows:

$$\begin{aligned} \mu_{\alpha_0}, \mu_{\alpha_1} &\sim \text{Normal}(0, 1) \\ \sigma_{\alpha_0}, \sigma_{\alpha_1} &\sim \text{Half-Normal}(0, 1.35) \\ \boldsymbol{\alpha'_0}, \boldsymbol{\alpha'_1} &\sim \text{Normal}(0, 1) \\ \boldsymbol{\alpha_0} &= \mu_{\alpha_0} + \sigma_{\alpha_0} \cdot \boldsymbol{\alpha'_0} \\ \boldsymbol{\alpha_1} &= \mu_{\alpha_1} + \sigma_{\alpha_1} \cdot \boldsymbol{\alpha'_1} \end{aligned} \tag{17}$$

We used these priors because they led to near-uniform priors over the individual-level trial-by-trial parameters (i.e. $\alpha(t)$) after being determined by Equation 15. The same parameterization was used for ω and c in the models assuming relations between facial expression intensity and reward and choice sensitivity, respectively.

We ran all models for 2,500 iterations across 4 separate sampling chains, with the first 500 samples as warm-up for a total of 8,000 posterior samples for each parameter. We checked convergence to the target joint posterior distributions 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.

5.6 Model comparison

We used the same model comparison procedures as described in study 1 to determine which model performed best in study 2. However, because we only had facial expression recordings for 31 of the 51 participants, model comparison proceeded in two stages. First, we fit each model to all 51 participants and performed both LOOIC model comparison and posterior predictive simulations to determine whether the findings from study 1 generalized to study 2 wherein we used a pure experience-based task. Next, we parameterized the best performing model (the Counterfactual Representation Learning model) using the scheme described in section 5.4.2, and we then fit each model to the subset of 31 participants to identify relations between their facial expression intensity in response to feedback and cognitive mechanisms that could influence behavior. We used LOOIC to identify the best performing facial expression model, which we then used to infer how experienced emotion affected subsequent choice behavior.

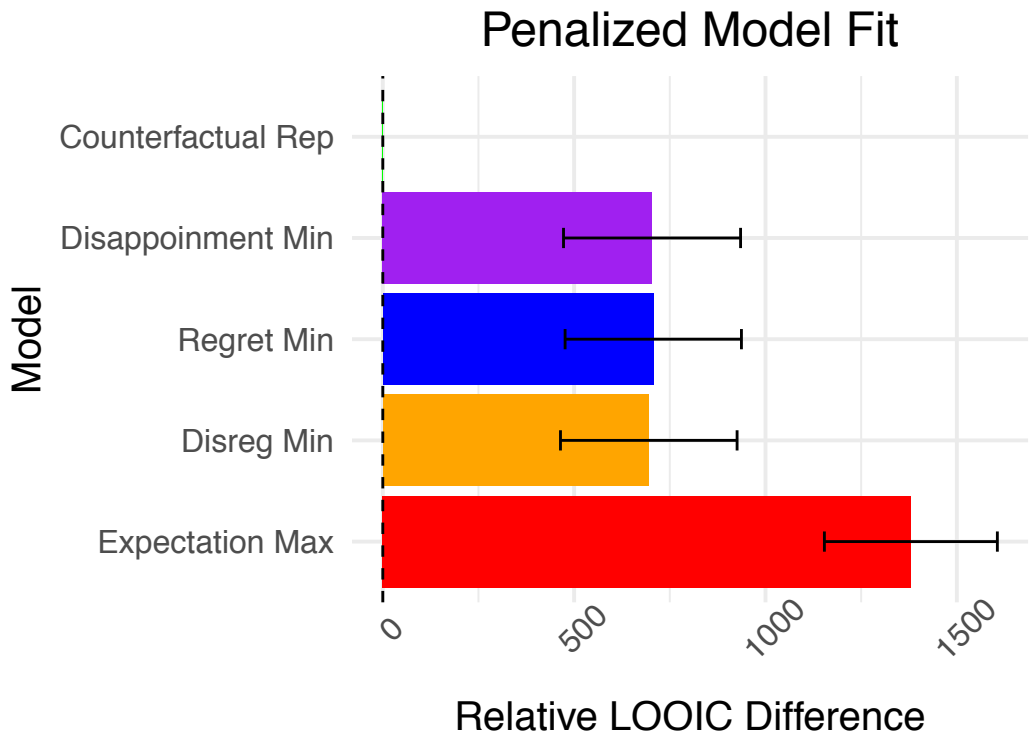
6. Study 2 Results

6.1 Model comparison: Penalized model fit.

Corroborating our findings from study 1, the Counterfactual Representation Learning model outperformed all other models when assessed using LOOIC (see Figure 6), indicating that it provided the best fit to participants' observed choice data accounting for model complexity.

Figure 6

Penalized model fit in the context of experience-based games



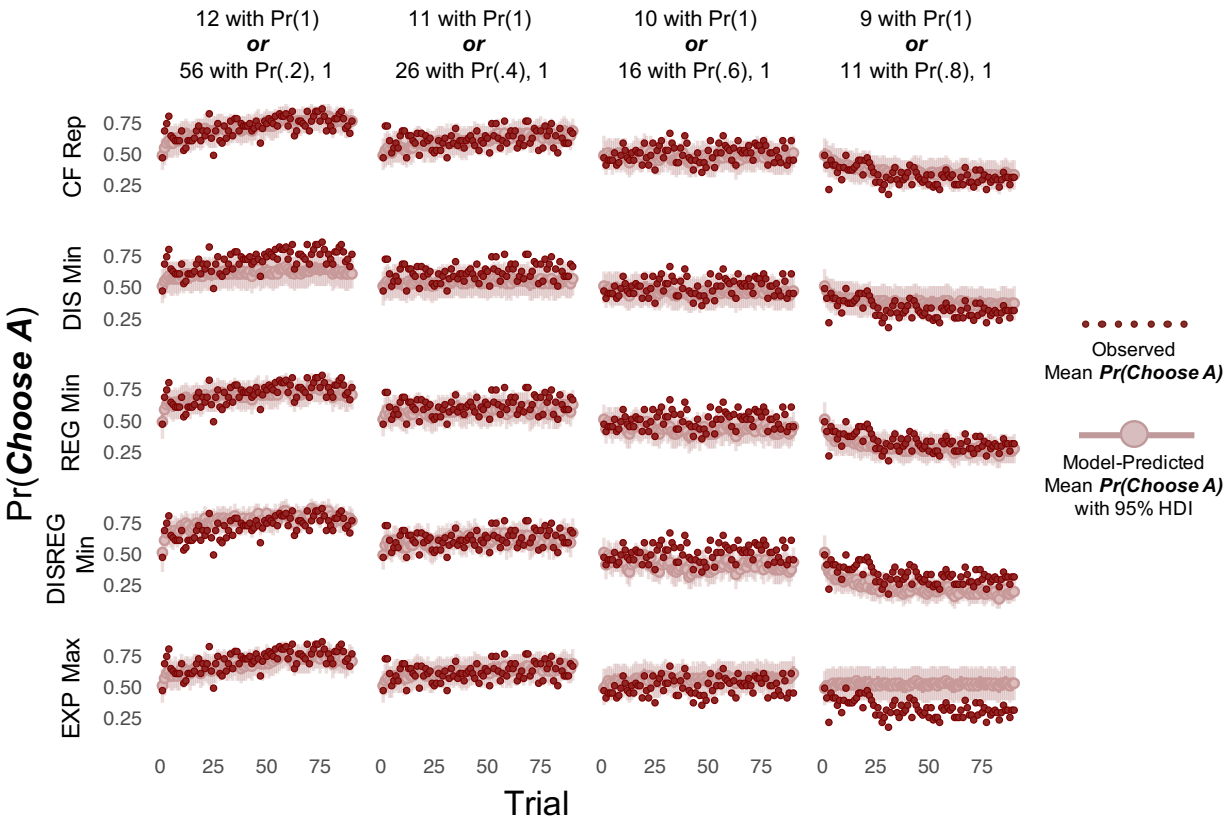
Note. Leave-one-out information criterion (LOOIC) scores relative to the best fitting model within across the 51 participants and 4 games from study 2. Lower LOOIC values indicate better model fit accounting for model complexity. Error bars represent ± 1 standard error of the difference between the best fitting model and respective competing models.

6.2 Model comparison: Posterior predictive simulations.

Figure 7 shows both the true and posterior predictive simulations across-participant (*N*=51) choice proportions for the safe versus risky options. Note that we fit each model to all 4 games simultaneously. Despite the safe and risky options having the same expected value within each game, participants 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). Similar to our results from study 1, the Counterfactual Representation and Regret Minimization Learning models best captured changes in participants' behavior across trials, suggesting that regret expectations continue to play a crucial role in changes in preference in response to feedback even when outcome distributions are unknown. Further, the Counterfactual Representation Learning model performed slightly better than the Regret Minimization model for game 3, as the Regret Minimization model tended to increasingly prefer option B across trials despite participants not showing such changes (see Figure 7). Overall, our posterior predictions for the pure experience-based task used in study 2 corroborate our findings from study 1, suggesting that the explicit representation of counterfactual values and probabilities is general across paradigms.

Figure 7

Posterior predictive simulations for all four experience-based games



Note. Posterior predictive simulations for each model across the 4 pure experience-based games in study 2. Different games are shown across columns (labels describing the game), and the competing models are shown across rows. Here, dark red points indicate the mean choice proportions across participants (*N*=51), and lighter red points with intervals represent the mean of the posterior predictive simulations across participants with 95% highest density intervals (HDI). CF Rep = Counterfactual Representation Learning model; DIS Min = Disappointment Minimization Learning model; REG Min = Regret Minimization Learning model; DISREG Min = Disappointment/Regret Minimization Learning Model; EXP Max = Expectation Maximization Learning model.

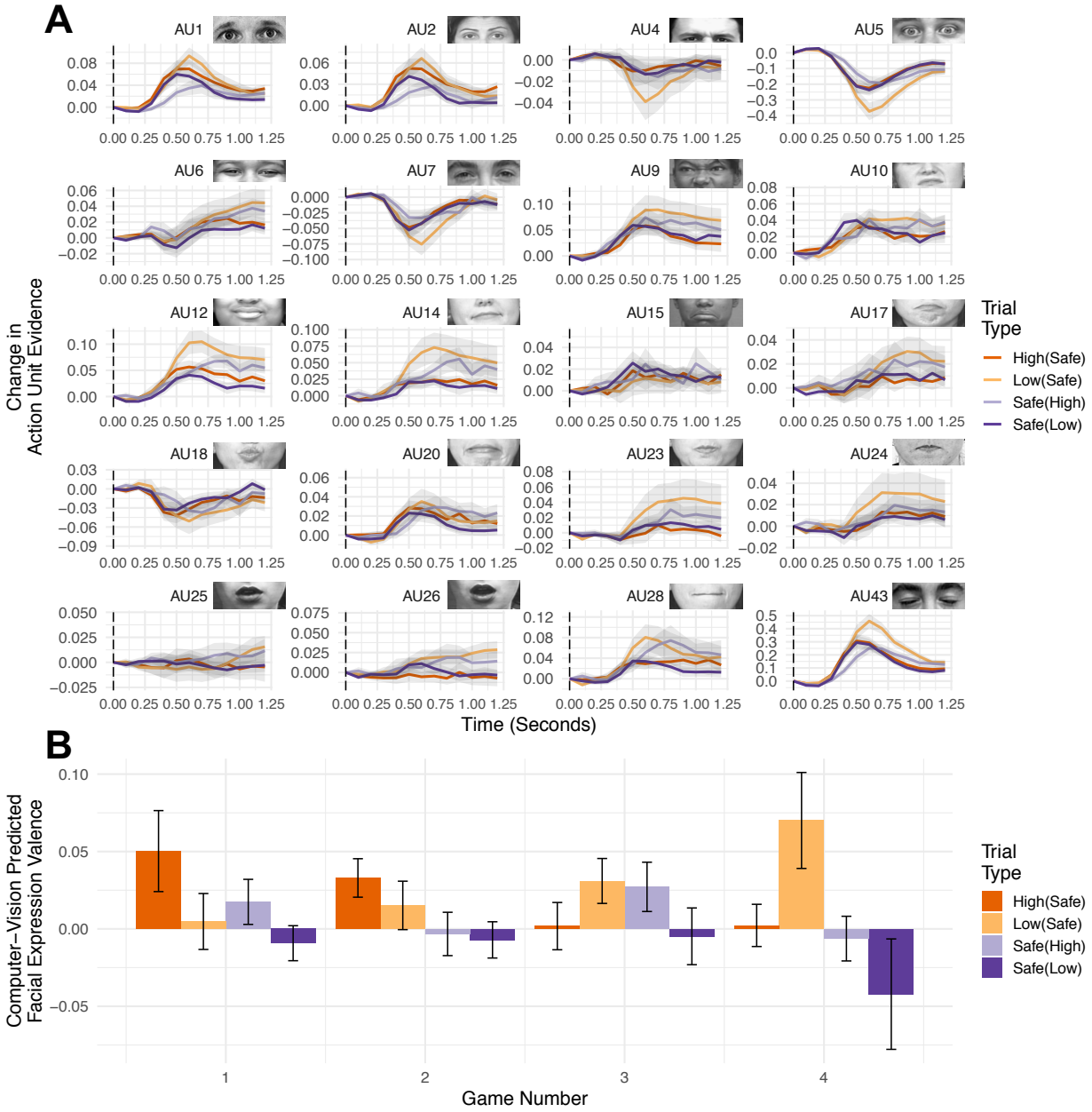
6.4 Model-based facial expression analysis.

Although our main focus was on overall changes in facial expression valence intensity and not on specific changes in facial expressions, we show a visual depiction of the changes in each of the 20 facial action units (panel A) and overall valence intensity (panel B) in response to different types of feedback in Figure 8 for completeness. In general, the most reliable changes were in facial expressions near the eyes—action units 1, 2, 9, and 43 all exhibited increases in evidence ratings (i.e. an increase in probability of being present) in response to feedback, whereas action units 5 and 7 tended to show decreases. For all these examples, changes were greatest when participants experienced both disappointment and regret on the current trial (e.g., when they chose the risky option and received the low outcome). Activation changed most strongly for AU43, which indicates that participants tended to close their eyes during feedback. This effect is consistent with studies using eye tracking, which show that people often avoid looking at both received and forgone outcomes (Ashby & Rakow, 2016).

Participants tended to have positive changes in valence in response to receiving the rare outcome upon choosing the risky option across games (i.e. the high outcome in games 1-2 and the low outcome in games 3-4; see Table 1), whereas the most consistent negative change was in game 4 when participants received the safe outcome and the rare, low outcome was foregone. These differences in descriptives across games are somewhat counterintuitive, but they are broadly consistent with regret theory, which predicts that experienced affect is proportional to the expectedness of the outcome. Because expectedness is a function of experience in our task, these descriptives are only rough proxies for the relationships we aimed to test between facial expression valence and computational model parameters. Our next analysis addresses this point, after which we interpret Figure 8B through the lens of the computational model.

Figure 8

Changes in facial expressions in response to feedback



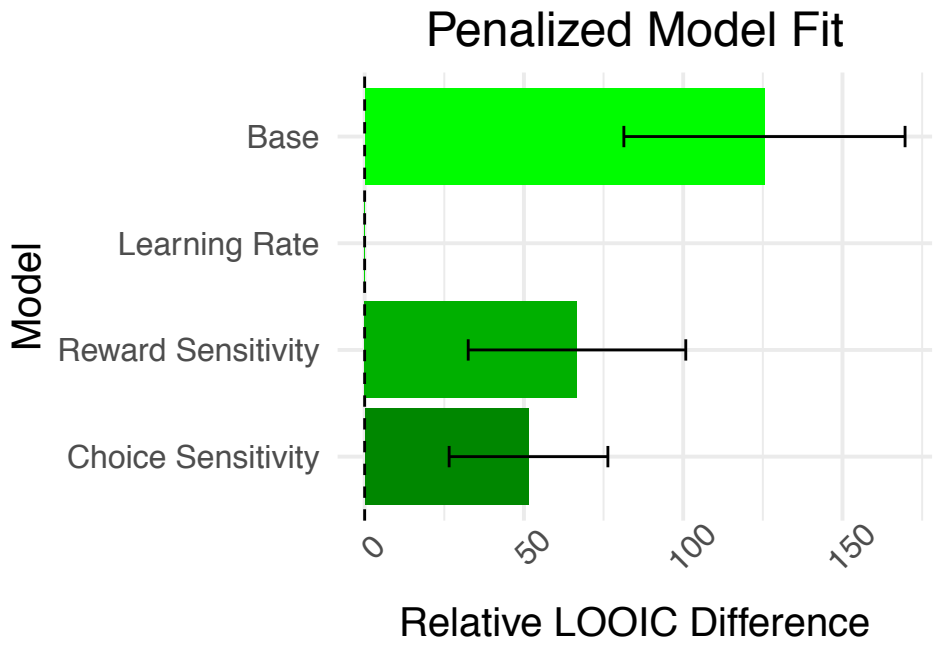
Note. (A) Changes in computer-vision predicted action unit evidence starting from the moment that feedback is received, separated across the different types of counterfactual outcomes. High(Safe) = chose risky option and received high outcome, foregone safe outcome; Low(Safe) = chose risky option and received low outcome, foregone safe outcome; etc. Change in evidence

was computed by first centering the first observation for each trial of the feedback period at 0, and then averaging across trials and participants. Shading indicates ± 1 standard error of the mean across participants. (B) The action unit evidence scores were entered into a secondary model to compute an overall valence score for each trial, and here we show the average changes in overall valence from the choice selection phase (see Figure 4) to the outcome presentation phase across games and different types of outcomes. The uncertainty intervals indicate ± 1 standard error of the mean across participants. Note that these descriptives are only rough proxies for the relationships we aimed to test between facial expression valence and computational model parameters, and we include them here for completeness.

Because the Counterfactual Representation Learning model showed better overall performance relative to competing models across a variety of games and task paradigms (e.g., description versus pure experience), we used it to further determine how facial expression valence intensity—our primary focus—influenced choice behavior through cognitive mechanisms of learning, reward sensitivity/valuation, or choice sensitivity (exploration/exploitation). Figure 9 shows that the Counterfactual Representation Learning model with a trial-by-trial valence effect on the learning rate parameter, as opposed to reward 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.

Figure 9

Penalized model fit for models facial expression models

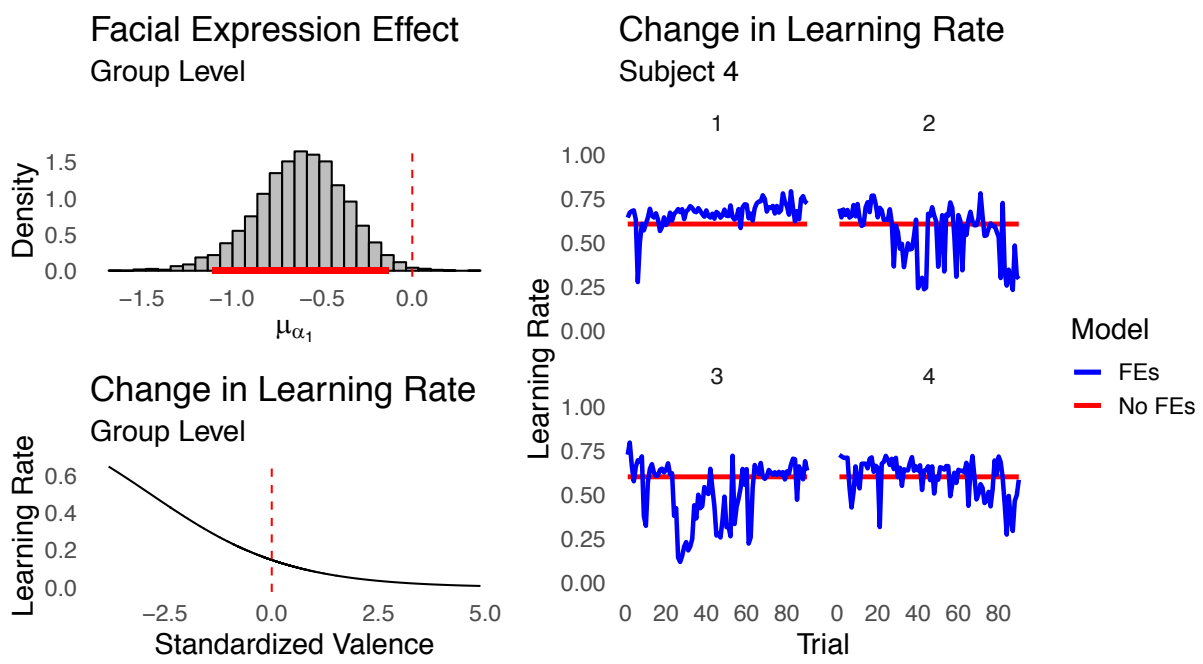


Note. Difference in LOOIC values between models including facial expressions to generate trial-by-trial model parameters. The Base model represents performance of the Counterfactual Representation Learning model for the 31 participants with facial expression recorded, without assuming facial expressions related to model parameters. Learning Rate, Reward Sensitivity, and Choice Sensitivity models assume that trial-by-trial variability in facial expression intensity is associated with changes in α, w, and c, respectively. Error bars represent ± 1 standard error of the difference between the best fitting model and respective competing models.

Figure 10 depicts the effect of facial expression valence intensity on learning. Notably, the group-level posterior distribution of valence on learning (μ_{α_1}) was almost completely negative (95% highest density interval (HDI) = [-1.11, -0.13]), with over 99% of the posterior mass below 0, indicating that participants 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. This finding is in line with functional theories of counterfactual thinking, which suggest that negative emotions facilitate goal-oriented changes in behavior (Connolly & Zeelenberg, 2002; Zeelenberg & Pieters, 2007). Additionally, it is consistent with previous studies identifying links between surprise and/or prediction error and experienced emotional valence (Mellers, 1997; 1999; Rutledge et al., 2014). Lastly, the relationship between valence intensity and learning rate facilitates interpretation of the descriptive results presented in Figure 8B. When participants made a risky choice and received an unexpected outcome (i.e. the rare outcome), they exhibited positive changes in facial expression valence on average. According to our computational model, these positive (as opposed to negative) changes indicate relatively less expectation updating and subsequently a lower probability if changing choice behavior. This overall pattern is consistent with decision justification theory (e.g., Connolly & Zeelenberg, 2002)—even when participants receive the low outcome in games 3 and 4, the risky choice is "justified" in that the low outcome is rare/unexpected.

Figure 10

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



Note. Posterior distribution (with 95% highest density interval highlighted in red) over the effect of facial expressions on the learning rate in the Counterfactual Representation Learning model (see Equations 15-16), a representation of the change in learning rate based on the group-level learning effect (μ_{α_1}) and facial expression valence ratings across participants, and an example from one participant of the learning rate with and without trial-by-trial effects of facial expression valence. Note that as facial expression valence becomes increasingly negative, the learning rate becomes increasingly rapid and vice-versa.

7. Discussion

Our findings are two-fold. First, we found that counterfactual comparisons between outcomes within (e.g., disappointment and elation) and across (e.g., regret and rejoice) choice options are weighted by "experience weights", which correspond to the subjective probability of each outcome occurring. Because each outcome probability is estimated independently, our model (the Counterfactual Representation Learning model) predicts that counterfactual expectations can influence behavior even if a counterfactual experience has not been directly experienced. For example, if I know that choosing action A can result in either $4 or $0 while action B can result in only $3, Equations 4-5 compute expectations over all possible counterfactual experiences given that I have independently observed each outcome. This account is consistent with traditional models of description-based risky decision-making (Kahneman & Tversky, 1979; Loomes & Sugden, 1982; 1986), but contrasts contemporary models of experience-based risky decision-making—most which do not assume an explicit representation of values and probabilities and instead compute expectations as averages over past experiences of disappointment and/or regret (e.g., Boorman, Behrens, & Rushworth, 2011; Erev et al., 2014; Hayden, Pearson, & Platt, 2009; Lohrenz, McCabe, Camerer, & Montague, 2007; Yechiam & Rakow, 2012).

In addition to drawing parallels to traditional models of risk, the explicit representation of outcome probabilities provides insight into functional theories of counterfactual thinking, which predict that people experience more regret when they make poor quality or unjustified decisions (e.g., Inman & Zeelenberg, 2002; Pieters & Zeelenberg, 2005). According to our model, decision quality/justifiability can be thought of in terms of a dynamic experience weight that corresponds to the subjective probability with which a disappointing or regretful outcome may occur.

Importantly, probability estimation can be affected by a number of different processes including recency of experience, similarity to other outcomes, saliency/attention, and differential sensitivity to positive versus negative prediction errors (e.g., Estes, 1976; Haines et al., under review; Lichtenstein, Slovic, Fischhoff, Layman, & Combs, 1978; Zacks & Hasher, 2002). Therefore, to the extent that such processes can be manipulated, we should expect counterfactual expectations to shift over time in predictable ways (e.g., making a regretted outcome more salient should may increase the experience weight for that outcome more so than if it is ignored, etc.). Future studies using the Counterfactual Representation Learning model may benefit from directly manipulating how outcomes are presented or how participants are instructed to evaluate outcomes to determine how to best manipulate counterfactual learning. For example, Sokol-Hessner and colleagues (2009) showed that people are less loss averse when they cognitively reframe losses as "one part of a large portfolio", rather than "an important decision in isolation". Such studies may reveal novel insights into decision-making interventions for individuals with psychiatric disorders characterized by emotion dysregulation.

Second, our results suggest that counterfactual representations are updated more rapidly when outcomes are experienced along with apparent intense negative affect (see Figure 10), which may explain why regret and disappointment can lead to negative functional outcomes yet also be looked back upon with 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, whereas appreciation may result from looking back on regretted decisions that improved later decision-making. 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), and 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). Relatedly, regret has minimal effects on individuals who are highly impulsive and lack trait anxiety (Baskin-Sommers, Stuppy-Sullivan, & Buckholtz, 2016), yet it 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. The rapid updating associated with negative emotion may be useful in fast-changing environments, but it could be detrimental when environments are stable (e.g., a consistently high learning rate may lead to an increase in perceived uncertainty despite action-outcome contingencies being consistent). 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 (e.g., Eldar, Rutledge, Dolan, & Niv, 2016).

To our knowledge, 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). Future work would benefit from combining automated facial expression coding with behavioral paradigms that collect self-reports of

emotion (e.g., Rutledge et al., 2014), which would both allow for more strenuous validity tests of automated measures and create opportunities for exploring the relationships between unobservable and observable emotional states. 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).

Although our approach is limited in that we only tested hypotheses regarding valence intensity, 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 limitations of traditional summary measures of behavioral performance for making inference on individual differences (e.g., Hedge, Powell, & Sumner, 2017), joint modeling of multiple data modalities may be a fruitful 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).

Finally, it is worth noting potential extensions to the Counterfactual Representation Learning model. As currently implemented, the probability of each possible outcome (i.e. feature) is learned independently. Although this mechanism works rather well when there are a small number of possible outcomes, it could quickly become inefficient when outcomes are continuous (e.g., if outcomes for each option were drawn from a normal distribution). To remedy this potential shortcoming, the learning mechanism could be modified to include a similarity or attention mechanism, wherein the probabilities of all possible outcomes are upweighted in proportion to their proximity to the observed outcome (e.g., see Turner, 2019). Such learning mechanisms may allow for the model to better generalize to more complex decision contexts than we explored, including multi-stage, multi-attribute, and multi-alternative contexts.

Acknowledgements

We would like to thank both Jerome R. Busemeyer and Brian F. O'Donnell for their insightful comments and suggestions on previous versions of this manuscript. The research was supported by the Basic Science Research Program through the National Research Foundation (NRF) of Korea funded by the Ministry of Science, ICT, & Future Planning (NRF-2018R1C1B3007313 and NRF-2018R1A4A1025891), the Institute for Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2019-0-01367, BabyMind), and the Creative-Pioneering Researchers Program through Seoul National University to W.-Y.A.

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

Allais, M. (1953). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica, 21*, 503. doi:10.2307/1907921

Ashby, N. J. S., & Rakow, T. (2016). Eyes on the Prize? Evidence of Diminishing Attention to Experienced and Foregone Outcomes in Repeated Experiential Choice. *Journal of Behavioral Decision Making*, *29*(2-3), 183–193. http://doi.org/10.1002/bdm.1872

Avrahami, J., Kareev, Y., & Hart, E. (2014). Taking the sting out of choice: Diversification of investments. *Judgement and Decision Making*, *9*(5), 373–386.

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

Chen, W., Liu, S.-Y., Chen, C.-H., & Lee, Y.-S. (2011). Bounded Memory, Inertia, Sampling and Weighting Model for Market Entry Games. *Games*, *2*(1), 187–199. http://doi.org/10.3390/g2010187

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

Cole, P. M., Zahn-Waxler, C., & Smith, K. D. (1994). Expressive control during a disappointment: Variations related to preschoolers' behavior problems. *Developmental Psychology*, *30*(6), 835–846. http://doi.org/10.1037/0012-1649.30.6.835

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

Darwin, C. (1998). *The expression of the emotions in man and animals*. New York, NY: Oxford University Press (Original work published 1872). xxxvi, 472 p. pp.

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

Delquié, P., & Cillo, A. (2006). Disappointment without prior expectation: a unifying perspective on decision under risk. *Journal of Risk and Uncertainty*, *33*(3), 197–215. http://doi.org/10.1007/s11166-006-0499-4

Diederich, A., & Busemeyer, J. R. (1999). Conflict and the Stochastic-Dominance Principle of Decision Making. *Psychological Science*, *10*, 353-359. doi:10.1111/1467-9280.00167

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 Anger Induction Positron Emission Tomography Study in Patients With Major Depressive Disorder 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

Erev, I., Ert, E., Plonsky, O., Cohen, D., & Cohen, O. (2017). From anomalies to forecasts: Toward a descriptive model of decisions under risk, under ambiguity, and from experience. *Psychological Review*, *124*(4), 369–409. http://doi.org/10.1037/rev0000062

Erev, I., & Roth, A. E. (2014). Maximization, learning, and economic behavior. *Proceedings of the National Academy of Sciences*, *111*, 10818–10825. http://doi.org/10.1073/pnas.1402846111

Estes, W. K. (1976). The cognitive side of probability learning. *Psychological Review, 83*, 37-64. doi:10.1037/0033-295X.83.1.37

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

Gläscher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a Role for Ventromedial Prefrontal Cortex in Encoding Action-Based Value Signals During Reward-Related Decision Making. *Cerebral Cortex*, *19*(2), 483–495. http://doi.org/10.1093/cercor/bhn098

Goodyer, I. M., Herbert, J., Tamplin, A., Secher, S. M., & Pearson, J. (1997). Short-Term Outcome of Major Depression: II. Life Events, Family Dysfunction, and Friendship Difficulties as Predictors of Persistent Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(4), 474–480. http://doi.org/10.1097/00004583-199704000-00009

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

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science*, *17*, 767-773. doi:10.1111/j.1467-9280.2006.01780.x

Grosskopf, B., Erev, I., & Yechiam, E. (2006). Foregone with the Wind: Indirect Payoff Information and its Implications for Choice. *International Journal of Game Theory, 34*, 285-302. doi:10.1007/s00182-006-0015-8

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., Kvam, P., Turner, B. M. (under review). Explaining the description-experience gap in risky decision-making: Asymmetric learning of positive and negative prediction errors as a causal mechanism.

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

Hart, E., Avrahami, J., Kareev, Y., & Todd, P. M. (2015). Investing Even in Uneven Contests: Effects of Asymmetry on Investment in Contests. *Journal of Behavioral Decision Making*, *28*(4), 395–409. http://doi.org/10.1002/bdm.1861

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

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*, 517-523. doi:10.1016/j.tics.2009.09.004

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

James, W. (1884). What is an emotion? *Mind, 9*, 188-205.

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

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. http://doi.org/10.2307/1914185

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.

Kareev, Y., Avrahami, J., & Fiedler, K. (2014). Strategic interactions, affective reactions, and fast adaptations. *Journal of Experimental Psychology: General*, *143*(3), 1112–1126. http://doi.org/10.1037/a0034641

Keltner D., & Lerner J. S. (2010). Emotion. In The handbook of social psychology, ed. DT Gilbert, ST Fiske, G Lindzey, pp. 317-52. New York, NY: Wiley

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*.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology*, *66*(1), 799–823. http://doi.org/10.1146/annurev-psych-010213-115043

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

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 551-578. doi:10.1037/0278-7393.4.6.551

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

Marchiori, D., & Warglien, M. (2008). Predicting Human Interactive Learning by Regret-Driven Neural Networks. *Science*, *319*(5866), 1111–1113. http://doi.org/10.1126/science.1151185

Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1993). The Mental Simulation of Better and Worse Possible Worlds. *Journal of Experimental Social Psychology*, *29*(1), 87–109. http://doi.org/10.1006/jesp.1993.1005

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

Mitchell, D. G. V. (2011). The nexus between decision making and emotion regulation: A review of convergent neurocognitive substrates. *Behavioural Brain Research*, *217*(1), 215–231. http://doi.org/10.1016/j.bbr.2010.10.030

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

Navarro, D. J. (2018). Between the Devil and the Deep Blue Sea: Tensions Between Scientific Judgement and Statistical Model Selection. *Computational Brain & Behavior*, *2*(1), 28–34. http://doi.org/10.1007/s42113-018-0019-z

Ordóñez, L. D., Connolly, T., & Coughlan, R. (2000). Multiple reference points in satisfaction and fairness assessment. *Journal of Behavioral Decision Making*, *13*(3), 329–344. http://doi.org/10.1002/1099-0771(200007/09)13:3<329::AID-BDM356>3.0.CO;2-Q

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

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature reviews neuroscience*, *9*(2), 148-158. doi:10.1038/nrn2317

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance. *Nature reviews neuroscience*, *11*(11), 773-782. doi:10.1038/nrn2920

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

Plonsky, O., Erev, I., & Ert, E. (2017). Calibration Data for Choice Prediction Competition 2018 (CPC18) [Data set]. Zenodo. http://doi.org/10.5281/zenodo.845873

Rakow, T., Ben R Newell, & Wright, L. (2015). Forgone but not forgotten: the effects of partial and full feedback in “harsh” and ‘kind’ environments. *Psychonomic Bulletin & Review*, *22*(6), 1807–1813. http://doi.org/10.3758/s13423-015-0848-x

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.

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

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., Phelps, E. A., & Smith, E. E. (2009). Thinking like a Trader Selectively Reduces Individuals' Loss Aversion. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(13), 5035–5040. http://doi.org/10.2307/40455144

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. (2019). Toward a common representational framework for adaptation. *Psychological Review*, *126*(5), 660–692. http://doi.org/10.1037/rev0000148

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

Worthy, D. A., Hawthorne, M. J., & Otto, A. R. (2013). Heterogeneity of strategy use in the Iowa gambling task: A comparison of win-stay/lose-shift and reinforcement learning models. *Psychonomic Bulletin & Review*, *20*(2), 364–371. http://doi.org/10.3758/s13423-012-0324-9

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., & 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

Zacks, R. T., & Hasher, L. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier & T. Betsch (Eds.), *ETC. Frequency Processing and Cognition* (pp. 21-36). New York, NY, US: Oxford University Press. doi:10.1093/acprof:oso/9780198508632.003.0002

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