A decision-making model with anticipation of surprise for explaining ‘irrational’ economic behaviors

Ho Ka Chan¹ and Taro Toyoizumi¹,²

¹ Laboratory for Neural Computation and Adaptation, RIKEN Center for Brain Science
² Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo

hoka.chan@riken.jp (HKC); taro.toyoizumi@riken.jp (TT)

Many experimental observations have shown that the expected utility theory is violated when people make decisions under risk. Here, we present a decision-making model inspired by the prediction of error signals reported in the brain. In the model, we choose the expected value across all outcomes of an action to be a reference point which people use to gauge the value of different outcomes. Action is chosen based on a nonlinear average of ‘anticipated surprise’, defined by the difference between individual outcomes and the abovementioned reference point. The model does not depend on non-linear weighting of the probabilities of outcomes. It is also straightforward to extend the model to multi-step decision-making scenarios, in which new reference points are created as people update their expectation when they evaluate the outcomes associated with an action in a cascading manner. The creation of these new reference points could be due to partial revelation of outcomes, ambiguity, or segregation of probable and improbable outcomes. Several economic paradoxes and gambling behaviors can be explained by the model. Our model might help bridge the gap between theories on decision-making in quantitative economy and neuroscience.

Introduction

The expected utility theory (EUT) (Friedman & Savage, 1948) is one of the most well-established models for analyzing decision-making under risk. Nevertheless, throughout the past decades, several experimental work, e.g. Allais (1953), Kahneman & Tversky (1979), have revealed instances where people’s choice cannot be explained by EUT. Some of those examples with a higher profile are coined the term ‘paradox’. Subsequently, new theories aiming to resolve these paradoxes, e.g. the prospect theory (PT) (Kahneman & Tversky, 1979), the extended Cumulative Prospect Theory (CPT) for actions with multiple outcomes (Tversky & Kahneman, 1992), and the regret theory (RT) (Bleichrodt & Wakker, 2015; Loomes & Sugden, 1982), have been proposed. Despite the effort, the mechanisms behind decision-making which accounts for behaviors not conforming to EUT is still an open area of research.

A well-known solution to this problem, proposed by PT and CPT, is to adopt a non-linear probability weighting function. The rationale for that is that overweighting of low
probability events could be a plausible explanation for risk seeking or risk aversion behavior in the face of such events. However, it is often not easy to decouple the effect of subjective probability perception and valuation of an outcome (Gonzalez & Wu, 1999). Moreover, studies have shown that the form of mapping between objective and subjective probability, or even the very existence of subjective probability perception, could be highly context dependent (Gallistel, Krishan, Miller, & Latham, 2014; Hertwig, Barron, Weber, & Erev, 2004; Wu, Delgado, & Maloney, 2009; Zhang & Maloney, 2012). Hence, creating a model with both probability weighting and valuation of monetary outcome could be prone to overfitting to a particular decision-making task.

Another important issue is what the reference points which people use to gauge the attractiveness of outcomes are. In PT and CPT, the reference point is decided on a case-by-case basis, through deciphering the linguistic and contextual nuances of the problem as it was presented to the decision maker. Adding on top of it are editing rules (Kahneman & Tversky, 1979) which modifies some of the outcomes. The reference point may also be changed as a result of the abovementioned editing. However, these editing rules can sometimes be contradictory among themselves, and the application of the said rules is subject to interpretation (Birnbaum, 2008). With these issues, a complex decision can be formulated by many possible ways, which leads to different predictions of people’s behaviors.

Here, we propose a model that does not rely on probability weighting and has an unambiguous reference point. The model is inspired by some neuroscience findings and ideas from economics. First, people may either consciously or unconsciously anticipate possible outcomes of an action before executing it. It was reported that when planning an action, neurons in the brain exhibit activities that resemble possible task-state sequences (Schuck & Niv, 2019). Second, such anticipation can influence decision-making. Many models in economics explain people’s choices by applying different concepts of ‘anticipation’, e.g. by considering anticipated regret for not choosing the other options (Bleichrodt & Wakker, 2015; Loewenstein & Sugden, 1982); by using the accumulation of anticipated elation or disappointment to explain the preference for immediate or delayed consumption (Loewenstein, 1987); and by separating anticipatory emotions into a general one and an option-specific one to explain time inconsistency in preference (Caplin & Leahy, 2001; Koszegi, 2010). Third, dopamine neurons in the brain that encode reward prediction errors, i.e., the deviation of reward outcome from its prediction (or, in layman’s term, ‘surprise’), are responsible not only for learning but also for motivating actions (Wise, 2004). Combining these ideas, we propose that ‘anticipated surprise’ due to probabilistic outcome is an essential component in decision-making. More specifically, we postulate that in decision-making involving multiple options, people preemptively compute how much each possible outcome of an option deviates from a reference point corresponding to their expectation, a natural choice of which would be the expected value across all outcomes for that option. This deviation, which we call the anticipated surprise, is then non-linearly scaled by a ‘surprise function’ and weighted by the objective probability of the realization of an outcome. In cases where the expected values of all options are the same, the option that maximizes such anticipated surprise is chosen.
When facing a complex decision, it is commonly believed that people, rather than considering all possible outcomes simultaneously, may break down the problem into modules and consider the said modules individually and sequentially (Anderson, 2002; Anderson et al., 2011; Doya, Samejima, Katagiri, & Kawato, 2002). To take this into account, we extend our model by introducing multi-step sequential branching. This may happen when some outcomes are much more/less likely than others, or when information about the outcomes are partially revealed in a sequential manner. We propose that in sequential branching, the above-mentioned reference point is updated from step to step. In each step, it corresponds to the expected value of the branches of the next steps. The difference between the expected value of the current and the previous step also contributes to the overall anticipated surprise.

In this paper, we first introduce our model in its simplest form without sequential branching in Section 1. In Section 2, we outline the properties of the surprise function which allows the reproduction of the results in Kahneman & Tversky (1979; Problem 3, 4, 7, 8), namely risk seeking for lotteries and risk averse otherwise in the gain domain, and the reflection effect. In Section 3, we extend the model to consider sequential branching and show how we can apply it to different scenarios to explain well known behaviors, e.g. ambiguity aversion, which is related to the Ellsberg paradox, behaviors in blackjack gambling, and event splitting, which is related to Allais paradox. In Section 4, we compare our model to other prominent decision-making models, suggest possible neural mechanisms related to our model, and briefly discuss possible ways to extend our model to consider cases where options have different expected values.

Section 1: Reproducing the patterns of prospect theory

1.1 From expected utility theory to prospect theory

In Kahneman’s 1979 paper (Problem 3, 4, 7, 8, and their primed version) in which PT is first proposed, a type of gambling problems, as shown in Table 1, was extensively discussed.

| Option 1 | Option 2 |
|----------|----------|
| Reward   | Probability | Reward   | Probability |
| Outcome 1 | \( \bar{x}/p \) | \( p \) | \( \bar{x}/p' \) | \( p' \) |
| Outcome 2 | 0 | 1 - \( p \) | 0 | 1 - \( p' \) |

Table 1: Outline for the gambling problems in Kahneman’s 1979 paper (Problem 3, 4, 7, 8, and their primed version). Without loss of generality, we assumed that \( p > p' \).

In these problems, people have to weigh between an option that gives them a comparatively small amount of money with a large probability (Option 1) and one that gives them a larger amount of money with a lower probability (Option 2). In Table 1, the rewards are scaled such that the expected value \( \bar{x} \) is the same for both options. In the actual examples in Kahneman’s paper, even though this is not always strictly the case, the expected values for both options have always been kept close. We assume that \( p > p' \).
Experimentally, it has been found that, given the reward is positive (i.e. in the gain domain), most people choose Option 1 (i.e. risk averse) when \( p' \) is large and Option 2 (i.e. risk seeking) when \( p' \) is small. The preference reverses if the reward is negative (i.e. in the loss domain), which is coined the term ‘reflection effect’. This result is inconsistent with the popular EUT (Friedman & Savage, 1948). In EUT, people choose the option that maximizes the expected utility \( U = E(u(x)) \), where \( u \) is the utility function, typically assumed to be increasing and concave. Kahneman’s result revealed that if we pick a large \( p' \), Option 1 is chosen, suggesting that \( U_1 > U_2 \), where \( U_a \) is the expected utility for Option \( a \). If we now scale the probability for both options with a sufficiently small factor \( \epsilon \), Option 2 will instead be chosen in the new problem, suggesting that \( U_1' = U_2' \), where \( U_{a'} \) is the expected utility for Option \( a \) in the new problem. However, by definition \( U_{a'} = \epsilon U_a \). This leads to a contradiction.

One obvious fix to the abovementioned inconsistency is to remove the linear relationship between \( U \) and the probability \( p \) that leads to \( U_{a'} = \epsilon U_a \). This is what PT assumes. In PT, the prospect function \( V \) is given by \( V = \sum \pi(p(x_j)) v(x_j) \), where \( x_j \) denotes possible outcomes in an option. The prominent difference with EUT in the gain domain is the application of non-linearity in the probability \( p \) by the introduction of the subjective probability function \( \pi \). Specifically, it overweights small probability (\( \pi(p) > p \) for small \( p \)) and vise versa (\( \pi(p) < p \) for large \( p \)). The function \( v \) is given by

\[
v(x) = \begin{cases} 
  f(x) & \text{for } x \geq 0 \\
  -k f(|x|) & \text{for } x < 0
\end{cases}
\]

using an increasing concave function \( f(x) \) defined for \( x \geq 0 \) and a constant \( k > 1 \). The qualitative difference between \( v \) and \( u \) is in the loss domain, where \( v \) is a convex function and is amplified by \( k \). The convexity of \( v \) is to account for the reflection effect described above. The inclusion of the amplifying factor \( k \) is to account for the risk aversiveness of the decision maker as shown in a well-known experimental observation (Rabin, 2000; Tom, Fox, Trepel, & Poldrack, 2007) that people resist taking 50-50 gamble for even money (e.g. preferring nothing to having 50% chance of gaining $10 and 50% chance of losing $10).

### 1.2 The anticipated surprise model

We aim to account for the above-mentioned behavioral observations using a simpler model. The model is based on anticipated surprise. We define ‘surprise’ \( \pi \) by the difference between an individual outcome and the expected value across all outcomes, i.e. \( \pi = x - E(x) \). Then, the surprise value \( \Delta \) of a choice is given by

\[
\Delta = E(\pi(z)),
\]

where \( \pi \) takes the same form as \( v \) in eq. (1) except that an increasing convex function \( f \) is used instead. We set \( k > 1 \) for the risk aversiveness factor in the definition of \( \Delta \) as in PT.

The major difference between PT and our model is that we do not make any transformation of \( p \). Instead, we use the expected value across all outcomes as the reference point to gauge the value of each outcome. By contrast, PT uses the status quo as the reference point in
most cases. In addition, we transform the surprise by a convex function in the gain domain as opposed to PT, which uses a concave function. The opposite is true for the loss domain.

Our model can reproduce all response patterns observed in the type of problems described in Table 1 for any arbitrary convex function $f$. The details of the proof for a general surprise function $\Delta$ is shown in Appendix 1. In Figure 1, we show how $\Delta$ varies with $p$ using $f(z) = z^{1.5}$ as an example. For the gain domain, $\Delta$ decreases from 0 at $p = 1$ to negative values as $p$ decreases, indicating people’s general preference for certainty as opposed to gambles. Nevertheless, as $p$ becomes small, $\Delta$ increases and becomes positive, indicating people’s preference for lotteries. When $k = 1$, the switch of preference takes place at $p = 0.5$, which is much larger than the experimental observations in Kahneman’s 1979 paper and other works (Somasundaram & Diecidue, 2017) suggest (Figure 1, red solid). By choosing a larger value of $k$, the point of switch can be lowered to a more realistic value of $p$ (Figure 1, red dashed).

![Figure 1: The surprise function $\Delta$ as a function of $p$. In this example, $f(z) = z^{1.5}$. Red (Blue) lines are the results in the gain (loss) domain. For solid lines, $k = 1$. For dashed lines, $k = 2.5$.](image)

In our model, the reflection effect is observed for all values of $p$ when $k = 1$ (See Figure 1, blue solid and Appendix 1). When $k > 1$, the model can still reproduce the important results of risk aversiveness (risk seeking) at high $p$ and risk seeking (risk aversiveness) at low $p$ in the gain domain (in the loss domain, respectively) (Figure 1). However, the reflection effect becomes imperfect, such that the model predicts risk aversiveness for both the gain and loss domains when $p$ takes intermediate values (around $p = 0.5$ in Figure 1; see Section 5.1 for discussion on the excessive risk aversiveness in the loss domain when $k > 1$).
Section 2: Sequential anticipation

2.1 Editing rules of Prospect Theory

In PT, the value function $v$ is computed relative to a reference point. It is natural to pick the status quo, i.e. the position before a decision problem is presented to the subject, as the reference point. However, it is easy to reframe a problem in such a way that, even though all the possible outcomes as compared to the status quo and the associated probability remain the same, people would tend to choose a different option from the original problem. This is depicted in problem 11 and 12 in Kahneman & Tversky (1979). To explain this phenomenon, PT introduces several heuristic editing rules which alter the structure of the outcomes of a problem, effectively changing the reference point. These editing rules do not have objective application criteria and are sometimes self-contradictory. Therefore, in many cases, it is possible to apply them in a post-hoc manner in a bid to explain experimental results (Birnbaum, 2008).

2.2 Intermediate states and update of reference points

Here, we further extend the model by allowing the creation of ‘intermediate states’ with new reference points amid a decision-making problem. Imagine your favorite sports team is playing another weaker team. Since your team is stronger, you probably expect your team to win before the match. Now, let us assume that your team is behind in the score midway through the match. At this point, you may already feel disappointed because the probability of your team not winning is higher than it was before the match. If your team indeed loses at the end of the match, you may not be very disappointed at that point because you have already updated your expectation while your team was behind. On the other hand, if your team makes an unlikely comeback win, you may feel pleasantly surprised.

The above example is the case of having ‘intermediate states’ due to revelation of new information about the outcomes. To account for these intermediate states, we assume that people mentally simulate possible sequences of events comprising the above-mentioned intermediate states during the decision-making process. Upon every simulated transition to an intermediate state, surprise is anticipated because the probability of attaining each outcome changes. We will quantify the surprise that emerges from the change in the expected value of the outcome in the next section. In addition, it is possible that the decision maker does not expect to receive extra information about the outcomes, but these intermediate states are still created during mental simulation simply because of the sequential nature of people’s anticipation process or how people group events together. We will discuss these cases with examples in Section 3. As we will show, such intermediate states have a major impact on the anticipated surprise in outcomes that the model computes.
2.3 Our model for sequential anticipation

Let us consider the decision-making problem illustrated in Figure 2.

![Decision-making problem with sub-branches](image)

Figure 2: A decision-making problem with sub-branches.

In this two-step problem, there are an intermediate step and an outcome step. In the intermediate step, there are \( m \) intermediate states, each happening with a probability \( p_i \), \( i = 1, 2, \ldots, m \) as depicted by the blue lines. From each intermediate state, it branches out to \( n_i \) possible final outcomes \( x_i^j \), \( j = 1, 2, 3, \ldots, n_i \) with probability \( p_i^j \) as depicted by the red lines.

The surprise value \( \Delta \) is given by:

\[
\Delta = \sum_{i=1}^{m} p_i \delta(E_i - E_0) + \sum_{i=1}^{m} p_i \sum_{j=1}^{n_i} p_i^j \delta(x_i^j - E_i),
\]

(3)

where \( E_0 = \sum_i p_i \sum_j p_i^j x_i^j \) is the expected value of all possible outcomes and \( E_i = \sum_j p_i^j x_i^j \) is the expected value for the outcomes in the \( i^{th} \) sub-branch.

Note that the 1st term in eq. (3) corresponds to the surprise in the intermediate step and the 2nd term corresponds to the surprise in the outcome step. In the model, we assume they contribute additively to the aggregated surprise value.

More generally, \( \delta \) can be computed in a cascading manner. Considering a decision-making process with \( T \) steps (including the outcome step), we define \( y_k(t) \) to be one of the possible
state trajectories in our mental branching process, where \( t = 0,1, \ldots, T \) refers to
the hierarchy of the branch and \( k = 1,2, \ldots, K \) refers to the index of the trajectory, with \( K \) being
the total number of trajectories. The surprise value of the \( k \)th trajectory is given by

\[
\Delta_k = \sum_{t=1}^{T} p(y_k(t)) \delta(E_k(t) - E_k(t-1)),
\]

where \( p(y_k(t)) = \prod_{t'=1}^{t} p(y_k(t')|y_k(t'-1)) \) is the probability of observing the \( k \)th
trajectory up to the \( t \)th step with \( p(y_k(t')|y_k(t'-1)) \) referring to the transition probability
of entering the state \( y_k(t') \) given that the previous state is \( y_k(t'-1) \). \( E_k(t) \) is the expected
value of the outcomes computed at the state \( y_k(t) \). Note that \( E_k(T) \) is the actual outcome
at the end of trajectory \( k \).

The aggregated surprise value is then obtained by linear summation of the surprise value of
all trajectories

\[
\Delta = \sum_{k=1}^{K} \Delta_k
\]

Please note that we are not advocating that the branching process would continue infinitely
for any decision problem, because of the obvious cognitive burden it presents. How long
would the process go on, whether there would be discount for remote branches, and how
the sum in eq. (5) can potentially be approximated by subsampling trajectories are beyond
the scope of this work.

The difference between our proposed model and the editing rules in PT is that in our model,
\( \Delta \) depends on both the surprise from the intermediate steps and the outcome step. In other
words, every change in the expected value of the outcomes along the anticipation trajectory
plays a part in the evaluation of an option. By contrast, while PT can handle some multi-
stage decisions, the prospect value \( V \) (See Section 1.1) is evaluated exclusively in the
outcome step using an updated reference point (Kahneman & Tversky, 1979). This means
that how the reference point is altered during the preceding stages is considered irrelevant.
We also note that there is no clear prescription of how to set the reference point in PT as
this is often done by invoking editing rules subjectively (Birnbaum, 2008). Furthermore,
since PT uses non-linear probability weighting, it is non-trivial to formulate Markovian
transitions of intermediate states under its framework (See Section 4.1 for more discussion
about potential issues of adopting the equivalence of sequential anticipation in PT).

In the following section, we will describe 3 types of situations where sequential branching
makes sense and can explain behaviors that are otherwise hard to understand. They are 1.
partial and sequential revelation of outcomes; 2. ambiguity; and 3. restructuring of the
problem into probable and improbable outcomes.
Section 3: Applications of the sequential branching model

3.1 Partial revelation of outcomes: blackjack gambling

In blackjack gambling, there are two phases of play: the player phase and the dealer phase. At the start of the player phase, the player makes their decision based on the available information on their and the dealer’s hand. At the end of the player phase, the player may receive new information on their hand depending on their earlier decision. While the final outcome may still be uncertain, the probability of the player winning may have changed, which would lead to an update of the expectation of the player. Here, we will illustrate how our sequence branching model captures the effect of such an intermediate state and lead to correct prediction on prominent gambling behavior. Please refer to Appendix 3 for the relevant details of the blackjack rules.

Bennis (2004) outlined 3 scenarios in which experienced gamblers face two options with very similar expected return and overwhelmingly choose the slightly worse option. They are 1. standing instead of hitting when the player’s cards totaling 16 points and the dealer’s totaling 10 points (16 vs 10 situations), 2. taking insurance side bets when they have blackjack and 3. taking insurance side bets when they have ‘good hands’, i.e. card compositions that imply a large probability of winning. For simplicity, we make the approximation that the expected returns of both options are exactly equivalent in all the scenarios.

3.1.1: blackjack gambling: 16 vs 10 situations

In this problem, the player may choose to hit (i.e. taking an extra card) or to stand (not taking an extra card). Without sequential branching or other forms of restructuring of the problem, the options of hitting and standing are indistinguishable, since for both options, the possible outcomes are equivalent (winning: $x; losing: $−x, where $x > 0$ is the bet size). With the additional constraint that the expected return for both options is the same, it implies that the probability of obtaining those outcomes is also the same.

So, how can we understand the observed preference for standing through sequential branching? Note that if the player chooses to stand, the player phase is immediately over. The outcome solely depends on what happens in the dealer phase. On the other hand, if the player chooses to hit, the expected returns changes depending on the results of the player phase. If the player busts, the game ends immediately (the player loses). If the player does not bust, the final outcome is still undecided as the dealer has yet to play, but the expected return has now risen because the possibility that the player loses due to busting has been eliminated. The branching schemes for standing and hitting are illustrated in Figure 3.
Figure 3: Branching schemes when the player decides to stand (left) and hit (right). Here $p_0$ is the probability that the dealer goes bust, $p_1$ is the probability that the player does not go bust after taking a card, $p_2$ ($0 < p_2 < 1$) is the probability that the player wins the hand (either because their hand is or the dealer goes bust) provided that they do not go bust. The constraint that the expected value for both options is the same means that $p_0 = p_1 p_2$. We further assume that when the player hits, only one more card will be taken, since taking more cards would lead to significantly worse expected return.

The surprise values for standing ($\Delta_{\text{stand}}$) and hitting ($\Delta_{\text{hit}}$) are computed in Appendix 2. It turns out that $\Delta_{\text{stand}} > \Delta_{\text{hit}}$ (Please see Appendix 2 for details.) This result matches the observed behaviors of gamblers.

This example illustrates how branching can influence the preference of people. In the case of hitting, if the player does not bust, their chance of winning increases, meaning that the intermediate state has a higher expected value than the initial state. Thus, the anticipated surprise for winning is split into two small components: one associated with the transition from the initial state to the intermediate state, and the other with the transition from the intermediate state to the outcome state. Because of the convexity of the surprise function, the aggravation of these small components of surprise is less than the big surprise acquired when the player anticipates standing and winning. By similar reasoning, the negative surprise acquired when the player anticipates hitting and losing is more severe than standing and losing.

3.1.2: blackjack gambling: taking side bets when the player has a good hand

In this problem, the player may choose to either take or not take an insurance side bets (i.e. betting that the dealer has blackjack. Please refer to Appendix 2 for details of the rules).

3.1.2.1: Special case: when the player has blackjack

When the player has blackjack, taking the side bet leads to a certain reward of the bet size, while not taking the side bet leads to a reward of 1.5 times the bet size with the probability $\frac{2}{3}$, and 0 otherwise. Note that this is exactly the same decision-making problem as the one
depicted in Table 1 with \( p = 1, \ p' = \frac{2}{3} \). We have already established that the certain option, i.e. taking the side bet, is preferred. Sequential branching is irrelevant in this case.

4.1.2.2: When the player has a general good hand

The study by Bennis (2004) took place in the US. In most casinos in the US, the dealer peeks for blackjack, which means that before the player phase, the dealer checks if their hand is blackjack. If so, the round will promptly end. If not, the expected return for the player increases, since the original expectation has factored in the possibility that the dealer has blackjack and that has been eliminated. This creates a similar branching scheme as for the 16 vs 10 situations with players hitting, as depicted in Figure 4.

![Figure 4: Branching schemes when the player decides to take side bets (left) and not to take side bets (right). Here \( p_1 \) is the probability that the dealer has blackjack, \( p_2 \) is the probability that the player wins the hand (either because their hand is larger or the dealer goes bust) provided that the dealer does not have blackjack. Since the risk aversion factor \( k \) only appears when negative surprise is incurred, the expressions for \( \Delta_{\text{bet}} \) and \( \Delta_{\text{no bet}} \) depend on the value of \( p_2 \). We will leave the details to Appendix 2. Here, we state the results that \( \Delta_{\text{bet}} > \Delta_{\text{no bet}} \) if \( p_2 > \frac{3}{8} \) (Note that the converse is not true. The result is inconclusive if \( p_2 \leq \frac{3}{8} \)). Since a large \( p_2 \) corresponds to a ‘good hand’, the prediction of the model matches the observed behaviors that people choose to take the side bet when they have a good hand.

To understand the results, first note that the red branches for both options are effectively equivalent, since the deviations from the updated expected values for the outcomes are the same. It makes sense since the side bet is already lost when it is revealed that the dealer’s hand is not blackjack, while the red branch corresponds to the situation after the above-mentioned revelation. For the blue branches, taking the side bet leads to relatively ‘certain’ outcomes. With good hands, the player’s chance of winning is large enough such that the
decision-making problem is outside the regime of lottery. In such cases, as we have studied in Section 1, risk aversion dominates, and certain outcomes are preferred.

3.1.4: final remarks

One may argue that these decision-making problems are so complex that the gamblers misinterpret the probability of winning underlying each option. However, information on the ‘optimal strategy’, i.e. the best option to take in order to maximize the expected return, for each initial card composition is well known and widely available in public. Bennis (2004) showed that experienced gamblers are indeed familiar with such optimal strategy, suggesting that gamblers are mostly likely aware that the options they choose are slightly inferior (again, in terms of expected return).

3.2 Ambiguity: Ellsberg paradox (2-urn problem)

In the above examples, the intermediate states (e.g., if the dealer has blackjack or not) are revealed during the game. Here, we propose that this identifiability of intermediate states is not necessary for sequential branching to happen. The intermediate states could remain unidentified and just be a mental product of people conceptualizing a problem. To illustrate this, let us consider the 2-urn version of the Ellsberg’s paradox (Ellsberg, 1961). In the experimental set-up, the subjects are told that there are two urns: one containing $n$ red balls and $n$ black balls; the other containing an undisclosed number of red and black balls totaling $2n$. The subjects will then pick a ball in one of the urns. They will win $1 if the ball is red (or black, the color does not matter in the observed behavior) and win nothing otherwise. It has been shown that most subjects prefer to pick balls from the urn with known ball composition. This phenomenon is known as ambiguity aversion.

Let us call picking balls from the urn with known composition ‘Option 1’ and picking balls from the other urn ‘Option 2’. For Option 1, it gives 50% chance of a reward of $1 and 50% of a reward of $0. The surprise value $\Delta_1$ is given by:

$$
\Delta_1 = \frac{1}{2} \delta \left( \frac{1}{2} \right) + \frac{1}{2} \delta \left( -\frac{1}{2} \right) 
$$

(6a)

For Option 2, it has been proposed that subjects may view it as a compound lottery, i.e. they first envision the various possible compositions in the urn (Krähmer & Stone, 2006; Segal, 1987). In the context of our model, these possible urn compositions constitute the intermediate states. The intermediate states can be parameterized by the fraction $m$ ($m = 0, \frac{1}{2n}, \frac{2}{2n}, \ldots, 1$) of balls that are red. Without further information, one may assume that the probability distribution underlying the ball compositions is symmetric about even number of red and black balls, i.e. the probability for entering the intermediate states, $p$, is constraint by $p(m) = p(1 - m)$ for all $0 \leq m < \frac{1}{2}$. For each intermediate state $m$, there are two final outcomes ($1$ with probability $m$ or $0$ with probability $1 - m$). Under this formalism, the surprise value $\Delta_2$ for Option 2 is given by

$$
\Delta_2 = \sum_m p(m) \left[ \delta \left( m - \frac{1}{2} \right) + m \delta \left( 1 - m \right) + (1 - m) \delta \left( -m \right) \right] 
$$

(6b)
Then, the difference is computed as

$$\Delta_1 - \Delta_2 = (k - 1) \sum_{m < 1/2} p(m) \left[ f \left( \frac{1}{2} - m \right) + mf(1 - m) + (1 - m)f(m) - f \left( \frac{1}{2} \right) \right] \quad (7)$$

In eq. (7), any symmetric pair of intermediate states $m$ and $1 - m$ gives the same magnitude of surprise with opposite signs, and therefore, the positive perfector $k - 1$ appears after combining their influence (See Appendix 3 for the mathematical details). $\Delta_1$ can be larger or smaller than $\Delta_2$ depending on the choice of $f$, suggesting that both ambiguity seeking and ambiguity aversion is in principle possible. However, we can show that ambiguity aversion is guaranteed if $f$ is not very convex (e.g., $f''(m) \leq \frac{3}{2} f'(m)$ for $0 < m < 1/2$) or if $f$ is highly convex (e.g., $\mu_0 f (1 - \mu_1) > f \left( \frac{1}{2} \right)$, with $\mu_c = \frac{\sum_{m < 1/2} m^c p(m)}{\sum_{m < 1/2} p(m)}$, which is between 0 and 1/2; see Appendix 3). Figure 5 plots $\Delta_2 - \Delta_1$ using power functions, $f(m) = m^r$, assuming a uniformly distributed $p(m)$. Consistent with our analysis, ambiguity aversion is observed both at small and large $r$, corresponding to the regime where $f$ is mildly and strongly convex. In this case, ambiguity aversion dominates over a large parameter space, which echoes the general preference for Option 1 as observed in experiments.

![Figure 5: $\Delta_1 - \Delta_2$ vs $r$ for $f(m) = m^r$ and $p(m) = \frac{1}{2n+1}$. $\Delta_1 - \Delta_2 > 0$ (ambiguity aversion) except for the small regime $2 < r < 2.5. n = 50, k = 2.$](image_url)

To understand the results, imagine that during the thought process in evaluating the value of picking the ambiguous urn, we anticipate that the ambiguous urn may contain more balls with the prize-winning color than our initial expectation. We would be pleasantly surprised by this potential scenario, culminating in a positive surprise on expectation. Nevertheless, in the 2nd branch, with the expectation now risen, we would have a negative overall anticipated surprise from expectation because of the possibility of not winning, which would lead to an outcome much worse than the updated expectation (The 2nd branch is exactly the
problems we discussed in Section 1. In this scenario, \( p \) is large). In the end, it is a trade-off between the surprise generated from the first and that from the second transition. The outcome of this trade-off determines one’s affinity for ambiguity.

As we mentioned we do not have to assume that the subjects are actually informed of the urn composition and/or the underlying probability distribution of the composition for the effects of sequential branching to kick in. How any potential difference in behavior when they are told of the urn composition and/or their distribution, and when they are not, can be theorized, while interesting, is not within the scope of this work (For experimental work, please refer to e.g. Halevy (2007)).

One may argue that when \( n \) is large, it is infeasible to mentally consider the huge number of branches as depicted in Figure 2. In practice, people may evaluate only a limited number of intermediate states such as a single pair of symmetric states, \( m \) and \( 1 - m \). As shown above as well as in Appendix 3, ambiguity aversion holds robustly in this case (regardless of the form of \( p(m) \)) if \( f \) is not very convex (or if \( f \) is highly convex). Similarly, people may use coarse-grained intermediate states to approximate the surprise value.

3.3 Segregation of probable and improbable outcomes

Now we turn to another type of scenario where sequential branching could possibly occur. Here, intermediate states do not correspond to physical states but reflect mental representations that group multiple improbable outcomes together instead. In real life, there are many events that we may encounter but with very low probability, like earthquake, traffic accident, winning a jackpot. While we do not completely ignore these events, they are not processed in conjunction with other more probable events. For example, we may buy insurance or draw separate contingency plans to mitigate the effects for several disastrous events at once. After that, we may not consider these events during our daily activity. In the context of our model, branches comprising of all these rare events are created. To illustrate this, let us consider the Allais paradox (Allais, 1953). The decision-making problems in Allais paradox are shown in Table 2.

(a) Problem 1

|                | Option 1 | Probability | Option 2 | Probability |
|----------------|----------|-------------|----------|-------------|
| Outcome 1      | 0        | 0.89        | 0        | 0.9         |
| Outcome 2      | 1        | 0.11        | 0        | 0.1         |

(b) Problem 2

|                | Option 1 | Probability | Option 2 | Probability |
|----------------|----------|-------------|----------|-------------|
| Outcome 1      | 1        | 1           | 0        | 0.01        |
| Outcome 2      |          | 1           | 0        | 0.89        |
| Outcome 3      |          | 0           | 0        | 0.1         |
Table 2: The two decision-making problems in Allais paradox. The asterisk depicts the option preferred by most people.

In Problem 1, Option 2 is preferred by most people. This can easily be understood by the significantly larger reward size for Option 2 as compared to the marginally smaller probability in obtaining the reward. Problem 2 is obtained by replacing the probability 0.89 portion of the reward ‘0’ by reward ‘1’ for both options. Despite this equal treatment of both options, Option 1 is favored in Problem 2 instead. The two options have slightly different expected returns, but we assume that it plays a relatively minor role in decision-making and therefore ignore its effect (but see Discussion for a potential extension to model problems with options having different expected values). First, we will show how sequential branching, by splitting the improbable outcomes, ‘0’ and ‘5’ in Problem 2, into sub-branches, makes Option 2 less appealing than the case without. The branching schemes for Option 2 when the improbable outcomes are grouped together, and when they are not, are depicted in Figure 6.

![Figure 6](image-url)

Figure 6: (a) The branching schemes for Option 2 of Problem 2 in Table 2 when the improbable outcomes are grouped together and when they are not. (b) The difference in surprise values between Option 1 and Option 2 for problem 1 (left), Problem 2 without...
grouping for improbable outcomes (middle), Problem 2 with group for improbable outcomes (right). Yellow (Blue) color corresponds to the regime where Option 1 (2) is preferred. The dark red line is the boundary where the options are equally preferred, i.e. $\Delta_{option\ 1} = \Delta_{option\ 2}$. We set $f(x) = x^r$.

The surprise value with (and without) grouping $\Delta_{group}$ (and $\Delta_{no\ group}$) are given by:

\begin{align*}
\Delta_{group} &= -0.89kf(E_0 - 1) + 0.1(f(E_1 - E_0) + f(5 - E_1)) + 0.01(f(E_1 - E_0) - kf(E_1)) \\
\Delta_{no\ group} &= -0.89kf(E_0 - 1) + 0.1f(5 - E_0) - 0.01kf(E_0)
\end{align*}

(8a)

(8b)

with the expected values at the initial state $E_0 = 0.89 \times 1 + 0.1 \times 5 = 1.39$ and intermediate state $E_1 = (0.1/0.11) \times 5 = 4.55$. In eqs. (7a) and (7b), the first, second, and third terms in the right-hand side represent the contributions from obtaining the outcomes ‘1’, ‘5’, and ‘0’, respectively. As the result of grouping, the positive surprise of getting the outcome ‘5’ is reduced (since $f(E_1 - E_0) + f(5 - E_1) < f(5 - E_0)$ for any convex and increasing $f$) and the negative surprise of getting the outcome ‘0’ is magnified (since $kf(E_1) > kf(E_0) + f(E_1 - E_0)$ for any convex and increasing $f$ and $k \geq 1$). This means that with grouping, Option 2 is more unfavorable than the case without, implying that the parameter space where Option 2 is worse than Option 1 is larger when branching is considered.

Now, by using a common class of convex functions $f(x) = x^r$, we compare the surprise values of the two options for both problems numerically. As shown in Figure 6b, for Problem 1, Option 2 is preferred for most regime (only except when the convexity of $f$ is weak and negative surprise is exacerbated by large $k$) as predicted. For Problem 2, without considering grouping, the preference remains unchanged in contrary to the experimental observation. However, when grouping is considered, in line with our analysis, the surprise value for Option 2 reduces (See also Figure S1 in Appendix), such that Option 1 is now preferred for a sizable parameter space. Our model is consistent with the experimental observation either when $k$ is not too large ($k < 3$) and $f$ is moderately convex ($1 < r < 2$), or when $k$ is very large and $f$ is very convex.

The grouping of improbable outcomes can also be meaningful if there are multiple events that give the same outcome. In EUT and PT, such events are effectively combined into a single event. In other words, these models predict that experimentally, combining same-outcome events makes no difference in the observed behaviors. However, this is not necessarily true. This is best illustrated by an experiment in (Birnbaum, 2008, Table 1) as shown in Table 3.

| (a) Problem 1 | Option 1 | Option 2 |
|---------------|----------|----------|
| **Color of ball drawn** | **Reward** | **Probability** | **Color of ball drawn** | **Reward** | **Probability** |
| Red          | 100      | 0.85     | Black         | 100      | 0.85     |
| White        | 50       | 0.1      | Yellow        | 100      | 0.1      |
| Blue         | 50       | 0.05     | Purple        | 7        | 0.05     |
**Problem 2**

| Color of ball drawn | Reward | Probability | Color of ball drawn | Reward | Probability |
|---------------------|--------|-------------|---------------------|--------|-------------|
| Black               | 100    | 0.85        | Red                 | 100    | 0.95        |
| Yellow              | 50     | 0.15        | White               | 7      | 0.05        |

Table 3: The two decision-making problems in (Birnbaum, 2008, Table 1). The asterisk depicts the option preferred by most people.

Although the two problems would be identical if the same-outcome events were combined, most people chose different options for Problem 1 and Problem 2. This motivates us to again investigate the branching schemes that group the improbable events without combining the same-outcome events, as shown in Figure 7a. Again, we neglect the effects of different expected returns for the two options (c.f. Discussion).
Figure 7: (a) The branching scheme for Problem 1 (top) and Problem 2 (bottom) in Table 3. (b) The difference in surprise values between Option 1 and Option 2 for problem 1 (left), problem 2 (right). As in Figure 6b, yellow (Blue) color corresponds to the regime where Option 1 (2) is preferred. The dark red line is the boundary where the options are equally preferred, i.e. $\Delta_{\text{option } 1} = \Delta_{\text{option } 2}$. We set $f(x) = x^r$.

For Option 1 in Problem 1, the events that lead to the same outcome are both improbable, and therefore get grouped into the same branch. In this case, they can be combined since the branch they are in leads to a certain outcome, i.e. no ‘surprise’. In other words, Option 1 of Problem 1 and 2 are equivalent.

For Option 2 in Problem 1, one event with the reward of ‘100’ is probable while the other is improbable, causing them to be grouped into separate branches. In this case, the branches cannot be combined (Note the similarity to Figure 6a).

The surprise values for Option 2 of Problem 1 ($\Delta_{\text{group}}$) and Problem 2 ($\Delta_{\text{no group}}$) are given by

$$\Delta_{\text{group}} = 0.85f(100 - E_0) - 0.15kf(E_0 - E_1) + 0.1f(100 - E_1) - 0.05kf(E_1 - 7)$$  \(9a\)

$$\Delta_{\text{no group}} = 0.95f(100 - E_0) - 0.05kf(E_0 - 7)$$  \(9b\)

with the expected values at the initial state $E_0 = 0.95*100 + 0.05*7 = 95.35$ and intermediate state $E_1 = 0.1*100 + 0.05*7 = 69$. For our model to explain the experimental results, $\Delta_{\text{group}} > \Delta_{\text{no group}}$ should hold. This is true if $k = 1$ due to the convexity of the surprise function $f$ in the gain domain, similar to the case in the Allais paradox. More generally, when $k > 1$, $\Delta_{\text{group}} > \Delta_{\text{no group}}$ requires that the effects of $f$’s convexity to overwhelm that of the risk aversion factor $k$. We illustrate this by comparing the surprise value of the two options for both problems numerically. As the consequence of event grouping, the surprise value for Option 2 increases as long as $r$ is not too small (See Figure S1 in Appendix), such that the model can reproduce the experimentally observed preference when $k$ is not too large ($k < 3$) and $f$ is moderately convex ($1 < r < 3$) (Figure 7b). Note that this regime roughly overlaps with the regime where our model’s prediction is consistent with experimental observation for the Allais paradox we previously discussed at $2 < k < 3$ and $1.2 < r < 1.8$. 
In both the Allais paradox and the Birnbaum problems, the preferred options can be altered by grouping improbable events together and postponing the detailed anticipation of each individual event to a later branching step. As a result, the associated surprise is broken down into a general one about the group and a specific one about individual events in the group. Since the surprise function is non-linear, it results in changes in the total aggregated value of surprise.

Section 4: Discussion

In this work, we introduce a decision-making model based on anticipated surprise that tackles the problems in which the predictions from EUT are inconsistent with experimental observations. The model hinges on 3 main assumptions: 1. The reference point being the expected value of all outcomes; 2. The surprise function is convex in the gain domain (and concave in the loss domain) and 3. risk aversion (\( k > 1 \)). Our model can explain Problem 3, 4, 7, 8 in Kahneman & Tversky (1979) without having to introduce extra concepts, like subjective probability weighting as PT does.

We then introduce the idea of updating the reference points within a single problem through the process of sequential anticipation. We outlined 3 scenarios in which sequential anticipation is applicable: 1. revelation of additional information about the outcome; 2. ambiguity and compound lottery; 3. Segregation of probable and improbable outcomes. The sequential anticipation model can predict behaviors that cannot be explained by both EUT and PT, especially those involving options that are considered equivalent by both models.

Our numerical results and mathematical analysis show that in order to replicate the experiments, the surprise function needs to be moderately convex and \( k \) needs to be larger than 1 but not too large. This is generally in line with the assumption of the model we made when we considered the single step problems in Section 1.

In addition, our sequential anticipation formalism allows us to have new perspectives on the axioms of EUT and the editing rules of PT, e.g. change of reference points, combination of events that lead to the same outcomes, providing guidelines of when they are valid and when they are not, as demonstrated in our examples in Section 3.

4.1 Difference between our model and the Prospect Theory

PT, as well as the slightly modified CPT, uses a concave (convex) value function for the gain (loss) domain. This could explain the risk adverse (seeking) behaviors in the gain (loss) domain for gambling with a decent probability of getting positive reward but could not explain the opposite behavior when the probability of getting positive reward is small (lottery/insurance). To this end, non-linear probability weighting is used. As discussed in Introduction, subjective probability perception is highly context-dependent (Gallistel et al., 2014; Hertwig et al., 2004; Wu et al., 2009; Zhang & Maloney, 2012), meaning that there is a risk that the non-linear probability weighting is overfit to the specific problems studied.
In addition, these models introduce a number of editing rules which modify some of the outcomes by splitting and/or disregarding them. While they have provided guidelines of how to apply these rules depending on the structure of the problem, there are situations, e.g. in Allais paradox, where multiple editing rules are eligible according to those guidelines. The choice of which rules to apply and/or the order of their application is subjective and could lead to different predictions of behaviors (Birnbaum, 2008).

A related issue involves the reference point. PT mostly uses the status quo as the reference point. While the status quo constitutes a reasonable initial reference point, it may change during the decision-making process for reasons we discussed. PT sometimes uses other reference points based on contextual information and the abovementioned editing rules (See e.g. Problem 11-12 in Kahneman & Tversky (1979)), but again there are only rough guidelines regarding how to implement them. The subjectiveness involved in the determination of the reference points and editing rules, and the ensued possibility for multiple predictions of behavior for the same problem calls for further elucidation on the model. On the other hand, the proposed multi-step branching framework this work presents has less ambiguity about reference points. As a result, the evaluation of each option and hence the prediction on people’s behavior is more objective.

A possible weakness of our model as compared to PT is that it predicts excessive risk averse behavior in the loss domain. For example, it leads to the prediction that people would take the more certain options for both the gain and loss regime at around \( p = 0.5 \) (Figure 1) when \( k \neq 1 \), which could be at odds with some experimental results (Kahneman & Tversky, 1979). Nevertheless, generally speaking, preferences of people tend to be ambiguous at intermediate values of \( p \) (Ruggeri et al., 2020; Somasundaram & Diecidue, 2017). So, qualitatively speaking, the partial lack of reflection effect predicted by our model in this regime may not be a fundamental weakness of the model. Another related problem is the ‘Asian disease problem’ (Tversky & Kahneman, 1981). The experiment showed that people essentially change their choices from being risk averse to being risk seeking when a constant reward was subtracted from all possible outcomes. Our model cannot reproduce this experimental result because anticipated surprise is computed only in relative to the expected value. To address this, a simple way to modify our model is to make \( k \) (the risk averse constant) a variable which is decreasing with the expected value of the outcome. As the expected value becomes negative, \( k \) becomes small, and hence the decision maker becomes more risk seeking.

4.2 Difference between our model and the Regret Theory

With the convex surprise function and the lack of probability weighting, one may find resemblance between our model and the Regret Theory (RT) (Bleichrodt & Wakker, 2015; Loomes & Sugden, 1982). In RT, decision-making is based on ‘anticipated regret’, computed by the difference in value between the outcome of an action and the outcome should other actions be chosen. RT can also reproduce Problem 3, 4, 7, 8 in Kahneman & Tversky (1979) (See e.g. eq. (7) in Loomes & Sugden (1982)). Nevertheless, there are fundamental differences between RT and our model. Unlike the regret theory, the evaluation of an option is irrelevant to the evaluation of other options in our model. Here, we would like to emphasize that ‘regret’ and ‘surprise’ are not competitive in nature, and they may provide
us with complementary views on how a decision is made. It is plausible that for some problems, both ‘anticipated regret’ and ‘anticipated surprise’ play a part in decision-making. However, we also note that it may be difficult to evaluate anticipated regret when the outcome structures of the options available are very different, since in such situations, there is no apparent way to conceive how the outcomes among these options would correspond to each other. This problem is aggravated when we consider sequential anticipation, where the number of plausible ways to relate intermediate states of different options explode as the number of branches increases. In such cases, regret may be too fuzzy to be perceived even after a decision has been made, let alone anticipating it beforehand.

4.3 Decision-making among options with different expected values

In this work, we are focusing on decisions with similar expected values. In real life, most decision-making problems involve options with significantly different expected values. It is therefore of interest to extend our model to also cover these cases. A possible way is to consider the change from the status quo to the initial expected value of an option to be part of the sequential anticipation process, i.e. to affix initial expected value \( E(t = 0) = 0 \) as the status quo in eq. (4) such that the next step corresponds to the initial state of the decision problem with \( E(t = 1) \) being the expected value of all outcomes. This extension naturally interpolates between the basic decision-making scenario solely based on expected outcomes that EUT targets and the scenario solely based on risk that our current work targets. One caveat of this approach is that the influence of surprise could still dominate over the influence of the expected values in the mixed domain (in which possible outcomes include both gain and loss) if the effect of an outcome-outlier is overly magnified by a highly convex \( f \) and/or large \( k \). However, this is unlikely if \( f \) is only mildly convex and \( k \) is not too large, which is consistent with the parameter regime suggested by the examples examined in this work. This proposed transition from the status quo to the initial state based on the expected values of an option is also biologically plausible. Dopamine neurons are essential for action execution and encoding reward prediction errors (DeWitt, 2014; Schultz, 2010). In conditioning experiments, these neurons give a first response when a reward prediction signal is given, which reflects the difference between the expected reward associated with the signal and the status quo, and then give a second response when the reward is given, which reflects the difference between the given and the expected reward (Pan, Schmidt, Wickens, & Hyland, 2005; Schultz, 2010). In the context of our model, a similar sequence could be anticipated during planning with the first component of the anticipated surprise being driven by a (mentally) simulation of picking an option and the second component being driven by a simulation of outcome presentation.

4.4 Possible neural implementation of anticipated surprise

Decision-making is often studied in neuroscience. Our decision-making model is based on two concepts: ‘surprise’ and ‘anticipation’. Both concepts have strong neural bases. For the former, as discussed in the previous section, the activity of dopamine neurons encodes the reward prediction errors, i.e. the deviation of the actual reward from the expectation of the experimental subject (DeWitt, 2014; Schultz, 2010, 2017). Stauffer, Lak, & Schultz (2014)
showed that the above-mentioned expectation can be well approximated by the expected value of the reward.

For the latter, it has been shown that shortly before an animal executes an action, hippocampus place cells exhibit a sequence of firing patterns that are highly predictive of the sequence of action the subject takes later (Ólafsdóttir, Bush, & Barry, 2018; Pfeiffer & Foster, 2013). This suggests that these anticipative patterns may be relevant to planning and deciding what actions to take by the subject. These neural activities related to anticipation is observed in multiple brain areas, including the hippocampus, the neocortex, and the dopaminergic midbrain. Integrated information from these areas is predictive of the decisions made by people (Iigaya et al., 2020).

Another question is how the evaluation in different sequential branches depicted in our model could be efficiently computed in neural systems, especially for complex problems with an elaborated branching structure. A recent work (Dabney et al., 2020) presented findings that are suggestive of parallel computing by dopamine neurons. In their experiment, subjects are given a cue representing probabilistic reward, and it was shown that a subset of dopamine neurons exhibit different reversal points characterizing the reward-activity curve. Also, some neurons only respond strongly to positive surprise and some only to negative surprise. The variety in their response profile allows them to simultaneously encode the surprise associated with probabilistic outcomes. Parallel computing during decision-making could have practical importance as suggested by experimental and theoretical studies. Parallel computing implemented in brain areas relevant for planning and action selection, like the basal ganglia (Alexander & Crutcher, 1990), enhances the efficiency of searches through branching trees (Kriener, Chaudhuri, & Fiete, 2020), which is especially important in situations where decisions have to be made under time pressure (Wan et al., 2011).

Our theory of how risk-preference is computed from anticipated surprise may provide a novel direction for exploring the neuroscience bases of decision-making in economics that involves branching outcomes.

4.5 Effect of learning on decision-making under anticipated surprise

The problems we studied are hypothetical in the sense that probability and reward size are explicitly given (Blackjack could be considered an exception, though the probability of winning for each card combination and for each option to take is widely available). In real life, the probability and the reward size may have to be learnt through trial-and-error. There is ample evidence that learning and past experiences affect decision-making. Momennejad, Otto, Daw, & Norman (2018) shows that offline replay, corresponding to learning and memory consolidation, alters the subsequent decision made during tasks. Chen, Kim, Nofsinger, & Rui (2007), and Hoffmann & Post (2017) shows that a trader’s return in the recent past affects their risk attitude and expected future return, even though it is often unclear whether past return is indicative of future return due to the volatility of the financial market.

How this learning is implemented in conjunction with decision-making under risk is outside of the scope of the model in this work. Simplistically, one could use one of Bayesian
reinforcement learning frameworks to train an “anticipated surprise model” for risky decision-making. Nevertheless, it is unlikely that learning and decision-making can be completely decoupled. Based on the similarity of online replay (during task performance) and offline replay (outside task performance), it is strongly suggested that that similar neural activity patterns are involved in planning and learning (Ólafsdóttir et al., 2018). Moreover, anticipated surprise likely changes along the learning stages. For example, it has been shown (Burger, Hendriks, Pleeging, & van Ours, 2020) that both the happiness for winning a small price and sadness for losing in a lottery is smaller for people who have bought a ticket than those who have not bought a ticket but rather given one. Since inevitably, there were people who regularly buy tickets, and therefore frequently experiencing winning and losing, in the group of ticket buyers, the results could possibly be interpreted as desensitization of emotions by repeated exposure to certain events. Indeed, it has been shown that responses from dopamine neurons, which modulates emotional processing, habituate with repeated presentations of the same stimuli (Ardiel et al., 2016; Menegas, Babayan, Uchida, & Watabe-Uchida, 2017). Given the role of dopamine neurons in relation to decision-making and encoding reward prediction errors, it is not unreasonable to speculate that such repeated stimulation may have a similar desensitizing effect on anticipated surprise. The interplay between learning and decision-making is an interesting future topic to study.

Reference

Alexander, G. E., & Crutcher, M. D. (1990). Functional architectures of basal ganglia circuits. *Trends in Neurosciences, 13*(7), 266–271.

Allais, M. (1953). Le Comportement de l’Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l’Ecole Americaine. *Econometrica, 21*, 503–546.

Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science, 26*(1), 85–112. https://doi.org/10.1016/S0364-0213(01)00062-3

Anderson, J. R., Bothell, D., Fincham, J. M., Anderson, A. R., Poole, B., & Qin, Y. (2011). Brain regions engaged by part-and whole-task performance in a video game: A model-based test of the decomposition hypothesis. *Journal of Cognitive Neuroscience, 23*(12), 3983–3997. https://doi.org/10.1162/jocn_a_00033

Ardiel, E. L., Giles, A. C., Yu, A. J., Lindsay, T. H., Lockery, S. R., & Rankin, C. H. (2016). Dopamine receptor DOP-4 modulates habituation to repetitive photoactivation of a C. elegans polymodal nociceptor. *Learning and Memory, 23*(10), 495–503. https://doi.org/10.1101/lm.041830.116

Bennis, W. (2004). Blackjack playing strategies and beliefs: A view from the field. *Journal of Gambling Issues, (10).* https://doi.org/10.4309/jgi.2004.10.1

Birnbaum, M. H. (2008). New Paradoxes of Risky Decision Making. *Psychological Review, 115*(2), 463–501. https://doi.org/10.1037/0033-295X.115.2.463

Bleichrodt, H., & Wakker, P. P. (2015). Regret theory: A bold alternative to the alternatives. *Economic Journal, 125*(583), 493–532. https://doi.org/10.1111/ejoc.12200

Burger, M. J., Hendriks, M., Pleeging, E., & van Ours, J. C. (2020). The joy of lottery play: evidence from a field experiment. *Experimental Economics, 23*(4), 1235–1256.
Under Uncertainty. The Economic Journal, 92(368), 805–824. https://doi.org/10.2307/2232669

Menegas, W., Babayan, B. M., Uchida, N., & Watabe-Uchida, M. (2017). Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. Elife, 6, 1–26. https://doi.org/10.7554/eLife.21886

Momennejad, I., Otto, A. R., Daw, N. D., & Norman, K. A. (2018). Offline replay supports planning in human reinforcement learning. Elife, 7, 1–25. https://doi.org/10.7554/eLife.32548

Ólafsdóttir, H. F., Bush, D., & Barry, C. (2018). The Role of Hippocampal Replay in Memory and Planning. Current Biology, 28(1), R37–R50. https://doi.org/10.1016/j.cub.2017.10.073

Pan, W. X., Schmidt, R., Wickens, J. R., & Hyland, B. I. (2005). Dopamine cells respond to predicted events during classical conditioning: Evidence for eligibility traces in the reward-learning network. Journal of Neuroscience, 25(26), 6235–6242. https://doi.org/10.1523/JNEUROSCI.1478-05.2005

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place cell sequences depict future paths to remembered goals. Nature, 497(7447), 74–79. https://doi.org/10.1038/nature12112

Rabin, M. (2000). Risk Aversion and Expected-Utility Theory: A Calibration Theorem. Econometrica, 68(5), 1281–1292.

Ruggeri, K., Alí, S., Berge, M. L., Bertoldo, G., Bjørndal, L. D., Cortijos-Bernabeu, A., … Folke, T. (2020). Replicating patterns of prospect theory for decision under risk. Nature Human Behaviour, 4(6), 622–633. https://doi.org/10.1038/s41562-020-0886-x

Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. Science, 364(6447). https://doi.org/10.1126/science.aaw5181

Schultz, W. (2017). Reward prediction error. Current Biology, 27(10), R369–R371. https://doi.org/10.1016/j.cub.2017.02.064

Segal, U. (1987). The Allais Paradox and Risk Aversion: An Anticipated Utility Approach. International Economic Review, 28(1), 175–201.

Somasundaram, J., & Diecidue, E. (2017). Regret theory and risk attitudes. Journal of Risk and Uncertainty, 55(2–3), 147–175. https://doi.org/10.1007/s11166-017-9268-9

Stauffer, W. R., Lak, A., & Schultz, W. (2014). Dopamine reward prediction error responses reflect marginal utility. Current Biology, 24(21), 2491–2500. https://doi.org/10.1016/j.cub.2014.08.064

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. Science, 211(4481), 453–458. https://doi.org/10.3366/jbctv.2017.0387

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty, 5(4), 297–323. https://doi.org/10.1016/0340-1650-2006-6-331

Wan, X., Nakatani, H., Ueno, K., Asamizuya, T., Cheng, K., & Tanaka, K. (2011). The Neural Basis of Intuitive Best Next-Move Generation in Board Game Experts. Science,
Wise, R. A. (2004). Dopamine, learning and motivation. Nature Reviews Neuroscience, 5(6), 483–494. https://doi.org/10.1038/nrn1406

Wu, S. W., Delgado, M. R., & Maloney, L. T. (2009). Economic decision-making compared with an equivalent motor task. Proceedings of the National Academy of Sciences of the United States of America, 106(15), 6088–6093. https://doi.org/10.1073/pnas.0900102106

Zhang, H., & Maloney, L. T. (2012). Ubiquitous log odds: A common representation of probability and frequency distortion in perception, action, and cognition. Frontiers in Neuroscience, 6(JAN), 1–14. https://doi.org/10.3389/fnins.2012.00001
Appendix 1: Reproducing the patterns of prospect theory in a model with surprise functions of general form

Here we are showing that the predictions from our model conform to the experimental observation in Kahneman & Tversky (1979) for the type of problems described in Table 1. For the gain domain, it is useful to study how the surprise function varies with respect to \( p \) in the setting described in Table 2.

| Outcome | Reward | Probability |
|---------|--------|-------------|
| Outcome 1 | \( \frac{1}{p} \) | \( p \) |
| Outcome 2 | 0 | \( 1 - p \) |

Table A1: The setting for studying the property of the surprise function in the gain domain. Note that we set \( \bar{x} = 1 \) to simplify analysis.

In this setting, the surprise value \( \Delta \), as a function \( p \), is given by:

\[
\Delta(p) = pf\left(\frac{1}{p} - 1\right) + (1 - p)f(-1) = pf\left(\frac{1}{p} - 1\right) - (1 - p)kf(1)
\]

(A1)

From (A1), we can obtain the following important results:

\[
\Delta(p = 1) = 0
\]

\[
\lim_{\varepsilon \to 0} \delta(p = \varepsilon) = pf\left(\frac{1}{p}\right) = \infty
\]

\[
\frac{\partial \delta}{\partial p} \bigg|_{p=1} = k_1f(1) - k_2f'(0) \geq k_2(f(1) - f'(0)) > 0
\]

\[
\frac{\partial^2 \delta}{\partial p^2} = \frac{1}{p^3}k_2f''\left(\frac{1}{p} - 1\right) > 0
\]

These results mean that \( \Delta(p) \) has a U-shape, being positive at small \( p \), decreasing and reaching negative value at some intermediate values of \( p \), and then increasing and reaches the value of 0 at \( p = 1 \). See, for example, figure 1.

The value of \( p \) where \( \Delta(p) = 0 \), aside from the trivial solution \( p = 1 \), corresponds to the probability where the subject switch from the gambling option to the certain option. For the special case where \( k = 1 \), it can be easily shown from (A1) that \( \Delta(p = 0.5) = 0 \). For the realistic case where \( k > 1 \), note that \( \delta \) is decreasing with \( k \), which means that \( \Delta(p = 0.5) < 0 \) and that \( \Delta \) crosses 0 at a smaller value of \( p \).

In the loss domain, the relevant setting is shown in Table A2.

| Outcome | Reward | Probability |
|---------|--------|-------------|
| Outcome 1 | \(-1/p\) | \( p \) |
| Outcome 2 | 0 | \( 1 - p \) |

Table A2: The setting for studying the property of the surprise function in the loss domain
In this setting, the surprise value $\Delta_-$ is given by:

$$\Delta_-(p) = pkf \left( \frac{-1}{p} - (-1) \right) + (1 - p)f(-(-1)) = -pkf \left( \frac{1}{p} - 1 \right) + (1 - p)f(1) \quad \text{(A2)}$$

We are particularly interested in knowing whether reflection effect holds in our model. For a pair options with probability $p_1$ and $p_2$ respectively, reflection effect holds when $\Delta(p_1) > \Delta(p_2)$ and $\Delta_-(p_1) < \Delta_-(p_2)$, or when $\Delta(p_1) < \Delta(p_2)$ and $\Delta_-(p_1) > \Delta_-(p_2)$. For $k = 1$, $\Delta_-(p) = -\Delta(p)$, which implies that the reflection effect can be observed at all values of $p$, since $\Delta(p_1) > \Delta(p_2)$ implies $\Delta_-(p_1) < \Delta_-(p_2)$ for all $p_1, p_2$, and vice versa. For $k > 1$, reflection effect is not always observed. For example, both $\Delta(0.5)$ and $\Delta_-(0.5)$ are negative and $\Delta(1) = \Delta_-(1) = 0$. So, we have $\Delta(1) > \Delta(0.5)$ and $\Delta_-(1) > \Delta_-(0.5)$, violating the reflection effect. In general, the reflection effect can be observed for extreme values of $p_1$ and $p_2$, while it becomes ambiguous when they take intermediate values.

**Appendix 2: Blackjack gambling**

The rule of Blackjack

Here we are only describing the rules of the blackjack that are relevant to this work. Also please note that there are no official rules for blackjack as casinos are free to introduce their own house rules. However, the one we describe here is one of the best known, widely used in casinos and work in gambling analysis (Shackleford, 2019a).

Blackjack uses standard 52-card decks. In a casino setting where many games are played, multiple decks of cards are used and played cards are introduced back into the deck often. This is to minimize the change of the winning odds as a result of changes in the composition of the deck as cards are exhausted. The goal of the player is to obtain a hand of cards with higher value than that of the dealer. The value of the hand depends on a point system. Points are calculated by summing up the numbers on all cards in the hand. Face cards (i.e. Js, Qs and Ks) stands for 10 points. For the Aces, they can stand for either 1 point or 11 points, chosen in order to maximize the value of the hand. The value of a hand in descending order is as follows: an Ace and a single card with 10 points (also known as blackjack (abbrev. BJ)), 21 points (but not BJ), 20 points, 19 points, 18 points, 17 points, 4-16 points, more than 21 points (known as busted hand) for dealer, busted hand for player.

In the beginning of a round, the player is given two cards and the dealer is given one card faced up and one card faced down (the player cannot read the card faced down). The player plays before the dealer, except when the faced-up card of the dealer is an Ace, in which case we will cover later. He can choose to take an extra card or not to. If he chooses to take an extra card, a card will be dealt to him, and the same option will be presented to him again until his hand become busted. If he chooses not to, the turn will be passed to the dealer. The dealer will keep taking extra cards until his hand has more than or equal to 17 points. After the dealer finished playing, the round ends, and whoever has a hand with a higher value wins. When the player wins, he receives twice the amount of his original bet, thus
giving him a net win of the size of his bet. An exception is when he wins with a hand of blackjack, in which case he will in addition get an extra amount of 0.5 times of his bet. When the player loses, he receives nothing, thus giving him a net loss of the size of his bet. In the event where the hand of the player and the dealer has equal value, normally the player will get back his bet and thus winning or losing nothing. However, to simplify our analysis, in section 4.1.1 and 4.1.2.2, we assume that the player will instead throw a coin to decide if he wins or loses, giving him 50% chance of winning and 50% chance of losing an amount equal to the size of his bet.

If the face-up card of the dealer is an Ace, the player will be given an option to place a side bet of the size half his original bet, known as ‘the insurance’, to bet on whether the hand of the dealer is Blackjack. Before the player starts playing, the dealer will peek at the face-down card. If the dealer has blackjack, the player will get a net win of the amount 2 times of his side bet, which is equivalent to the size of his original bet. At this point, the original bet can also be resolved. If the dealer does not have blackjack, the player will lose his side bets and play resumes in order to resolve the original bet. The placement of the insurance side bet does not affect how the original bet is resolved.

Analysis for section 4.11 (16 vs 10 situation)

Here we are showing that $\delta_{\text{stand}} \geq \delta_{\text{hit}}$. First, we note that the size of $x$ has no role in affecting the rank of $\Delta_{\text{bet}}$ and $\Delta_{\not\text{bet}}$ since we can make a transformation to remove $x$ during the comparison between these two quantities. To simplify our analysis, we set $x = 1$ such that $\delta_{\text{stand}}$ and $\delta_{\text{hit}}$ is given by:

$$
\Delta_{\text{stand}} = p_0 f(1 - p_0) - (1 - p_0) k f(p_0)
$$

$$
\Delta_{\text{hit}} = - \left( 1 - \frac{p_0}{p_2} \right) k f(p_0) + \frac{p_0}{p_2} [f(p_2 - p_0) + p_2 f(1 - p_2) - (1 - p_2) k f(p_2)]
$$

(A3a)

(A3b)

$$
D = \Delta_{\text{stand}} - \Delta_{\text{hit}} = p_0 \left( 1 - \frac{1}{p_2} \right) k f(p_2) - f(p_0) + p_0 (f(1 - p_0) - f(1 - p_2)) - \frac{p_0}{p_2} f(p_2 - p_0)
$$

(A4)

Since $f$ is convex, $f(p_2 - p_0) \leq f(p_2) - f(p_0)$, eq. (A4) can then be rewritten as

$$
D \geq \frac{p_0}{p_2} (k - 1) (f(p_2) - f(p_0)) + p_0 [(f(1 - p_0) + k f(p_0)) - (f(1 - p_2) + k f(p_2))]
$$

(A5)

For $k = 1$, the 1st term in eq. (A5) vanishes. For the 2nd term, define $F(z) = f(1 - z) + f(z)$, $z \in [0,1]$.

Consider the conditions $0 \leq p \leq q \leq \frac{1}{2}$, using the properties of a convex function again, we have

$$
f(1 - p) - f(1 - q) \geq (q - p) f'(1 - q)
$$

$$
f(q) - f(p) \leq (q - p) f'(q) \leq (q - p) f'(1 - q) \leq f(1 - p) - f(1 - q)
$$

(A6)
Eq. (A6) means \( F(q) > F(p) \) for \( 0 \leq p \leq q \leq \frac{1}{2} \). For the domain \( \frac{1}{2} < z \leq 1 \), we can make use of the fact that by definition, \( F(z) = F(1 - z) \).

In a player’s 16 vs dealer’s 10 situation, \( p_0 = 0.23, p_1 = \frac{5}{13}, p_2 = \frac{p_0}{p_1} = 0.6 \) (Shackleford, 2019b). By above, \( f(1 - p_0) + f(p_0) = F(0.23) > F(0.4) = F(0.6) = f(1 - p_2) + f(p_2) \), and thus \( D \geq 0 \).

For \( k > 1 \), we note that \( \Delta \delta \) is an increasing function with \( k \) since
\[
\frac{d \delta}{dk} = p_0 \left( \frac{1}{p_2} - 1 \right) \left( f(p_2) - f(p_0) \right) > 0
\]
(A7)

Therefore \( D \geq 0 \) also holds for \( k > 1 \).

Analysis for section 4.1.2.2 (taking side bets when the player has a non-blackjack good hand)

Here we are showing that \( \Delta_{\text{bet}} \geq \Delta_{\text{no bet}} \). We are assuming that both the options of taking and not taking side bets has the same expected value, this amounts to \( p_1 = \frac{1}{3} \) (in reality \( p_1 \) is slightly different at \( p_1 = \frac{4}{13} \)). Denoting the initial expected value of the hand, i.e before the dealer peeks for BJ by \( E_0 \). \( E_0 \) can be expressed as
\[
E_0 = (1 - p_1) \left[ \frac{x}{2} p_2 - \frac{3x}{2} (1 - p_2) \right] = \frac{2x}{3} \left( 2p_2 - \frac{3}{2} \right) = x \left( \frac{4}{3} p_2 - 1 \right)
\]
(A8a)

Like in the previous section, the size of \( x \) has no role in affecting the rank of \( \Delta_{\text{bet}} \) and \( \Delta_{\text{not bet}} \) since we can make a transformation to remove \( x \) during the comparison between these two quantities. To simplify our analysis, we set \( x = 1 \) such that
\[
E_0 = \frac{4}{3} p_2 - 1
\]
(A8b)

Let \( E_1 \) be the expected value after the dealer peeks for BJ if the player takes the side bet.
\[
E_1 = -\frac{E_0}{(1-p_1)} = \frac{3}{2} E_0
\]
(A9)

Let \( E_2 \) be the expected value after the dealer peeks for BJ if the player does not take the side bet.
\[
E_2 = E_1 + \frac{1}{2} = \frac{3}{2} E_0 + \frac{1}{2}
\]
(A10)

It is trivial that \( E_2 \geq E_0 \). However, whether it is \( E_1 \geq E_0 \) or \( E_1 \leq E_0 \) depends on the sign of \( E_0 \). Therefore, we divide the problem in the separate cases: \( E_0 < 0 \) and \( E_0 \geq 0 \).

Case 1: \( E_0 < 0 \)

When \( E_0 < 0 \), \( E_1 < E_0 \). \( \Delta_{\text{bet}} \) and \( \Delta_{\text{no bet}} \) can be expressed as
\( \Delta_{bet} = p_1gf(E_0) + (1-p_1) \left[ -kf \left( \frac{E_0}{2} \right) + p_2f \left( \frac{1}{2} + \frac{3E_0}{2} \right) - (1-p_2)kf \left( \frac{3}{2} + \frac{3E_0}{2} \right) \right] \) \hspace{1cm} (A11a)

\( \Delta_{no\ bet} = -p_1kf(1+E_0) + (1-p_1) \left[ f \left( \frac{1}{2} + \frac{E_0}{2} \right) + p_2f \left( \frac{1}{2} - \frac{3E_0}{2} \right) - (1-p_2)kf \left( \frac{3}{2} + \frac{3E_0}{2} \right) \right] \) \hspace{1cm} (A11b)

Substituting in \( p_1 = \frac{1}{3} \), we have

\[ D = \Delta_{bet} - \Delta_{no\ bet} = \frac{1}{3} \left[ \left( f(-E_0) + kf(1+E_0) \right) - 2 \left( kf \left( \frac{E_0}{2} \right) + f \left( \frac{1}{2}(1+E_0) \right) \right) \right] \] \hspace{1cm} (A12)

For \( k = 1 \), using the properties of a convex function, we have

\[ f(-E_0) \geq 2f \left( \frac{E_0}{2} \right) \] and \( f(1+E_0) \geq 2f \left( \frac{1}{2}(1+E_0) \right) \).

It follows that \( D \geq 0 \).

For \( k > 1 \), we again study the derivative of \( \Delta\delta \) with respective to \( k \)

\[ \frac{dD}{dk} = \frac{1}{3} \left[ f(1+E_0) - 2f \left( \frac{E_0}{2} \right) \right] \] \hspace{1cm} (A13)

If \( E_0 \geq \frac{-1}{2} \), we have \( f(1+E_0) \geq 2f \left( \frac{1}{2}(1+E_0) \right) \geq 2f \left( \frac{E_0}{2} \right) \) such that \( D \) increases with \( k \) and remains positive at \( k > 1 \).

From eq. (A8b), \( E_0 \geq \frac{-1}{2} \) when \( p_2 > \frac{3}{8} \). Since a large \( p_2 \) means that there is a large probability of winning with the hand, which, in order words, means the hand is a good hand. The results suggest that players prefer to take side bets when they have good hands.

**Case 2: \( E_0 \geq 0 \)

Since in the previous section, we established that \( E_0 \geq \frac{-1}{2} \) is always considered good hands, we have to show that when \( E_0 > 0 \), \( \Delta\delta \geq 0 \) unconditionally. Noting that when \( E_0 \geq 0 \), \( E_1 \geq E_0 \), and that from eq. (A8b) \( E_0 \) has an upper bound of \( \frac{1}{3} \), \( \Delta_{bet} \) and \( \Delta_{no\ bet} \) is given by

\( \Delta_{bet} = -p_1kf(E_0) + (1-p_1) \left[ f \left( \frac{E_0}{2} \right) + p_2f \left( \frac{1}{2} + \frac{3E_0}{2} \right) - (1-p_2)kf \left( \frac{3}{2} + \frac{3E_0}{2} \right) \right] \) \hspace{1cm} (A14a)

\( \Delta_{no\ bet} = -p_1kf(1+E_0) + (1-p_1) \left[ f \left( \frac{1}{2} + \frac{E_0}{2} \right) + p_2f \left( \frac{1}{2} - \frac{3E_0}{2} \right) - (1-p_2)kf \left( \frac{3}{2} + \frac{3E_0}{2} \right) \right] \) \hspace{1cm} (A14b)

Substituting in \( p_1 = \frac{1}{3} \), we have

\[ D = \Delta_{bet} - \Delta_{no\ bet} = \frac{1}{3} \left[ k(f(1+E_0) - f(E_0)) - 2 \left( f \left( \frac{1}{2} + \frac{E_0}{2} \right) - f \left( \frac{E_0}{2} \right) \right) \right] \] \hspace{1cm} (A15)

For \( k = 1 \), again using the properties of a convex function, we have

\[ f(1+E_0) - f \left( \frac{1}{2} + \frac{E_0}{2} \right) > \left( \frac{1}{2} + \frac{E_0}{2} \right)f' \left( \frac{1}{2} + \frac{E_0}{2} \right) \]

\[ f \left( \frac{1}{2} + \frac{E_0}{2} \right) - f \left( \frac{E_0}{2} \right) < \frac{1}{2}f' \left( \frac{1}{2} + \frac{E_0}{2} \right) \]

\[ f(E_0) - f \left( \frac{E_0}{2} \right) < \frac{E_0}{2}f'(E_0) < \frac{E_0}{2}f' \left( 1 + \frac{E_0}{2} \right) \]
This gives
\[ f(1 + E_0) - f\left(\frac{1}{2} + \frac{E_0}{2}\right) > f\left(\frac{1}{2} + \frac{E_0}{2}\right) - f\left(\frac{E_0}{2}\right) + f(E_0) - f\left(\frac{E_0}{2}\right) \]
\[ f(1 + E_0) - f(E_0) > 2 \left( f\left(\frac{1}{2} + \frac{E_0}{2}\right) - 2f\left(\frac{E_0}{2}\right) \right) \]

It follows that \( D \geq 0 \).

From (A15), it is obvious \( D \) increases with \( k \) such that \( D \geq 0 \) still holds when \( k > 1 \).

Appendix 3: Ellsberg paradox

Here we are exploring the condition in which the unambiguous urn would be preferred over the ambiguous urn. We assumed that a symmetric prior probability that satisfies \( p(m) = p(1 - m) \) for \( m \in \left\{ 0, \frac{1}{2n}, \frac{2}{2n}, \ldots, 1 \right\} \). The surprise value for picking the unambiguous urn (\( \Delta_1 \)) and the ambiguous urn (\( \Delta_2 \)) is given by eqs. (6a) and (6b) in the main text. For convenience, we are repeating them here:

\[ \Delta_1 = \frac{1}{2} \delta \left( \frac{1}{2} \right) + \frac{1}{2} \delta \left( -\frac{1}{2} \right) = \frac{1}{2} f\left(\frac{1}{2}\right) - \frac{k}{2} f\left(\frac{1}{2}\right) \]
\[ = \sum_m p(m) \left[ \frac{1-k}{2} f\left(\frac{1}{2}\right) \right] \]
\[ = (1 - k) \sum_{m<1/2} p(m) f\left(\frac{1}{2}\right) + p\left(\frac{1}{2}\right) \frac{1-k}{2} f\left(\frac{1}{2}\right) \]  
(A16a)

\[ \Delta_2 = \sum_m p(m) \left[ \delta \left( m - \frac{1}{2} \right) + m \delta (1 - m) + (1 - m) \delta (-m) \right] \]
\[ = \sum_{m<1/2} p(m) \left[ -k f\left(\frac{1}{2} - m\right) + mf(1 - m) - k(1 - m)f(m) \right] + p\left(\frac{1}{2}\right) \frac{1-k}{2} f\left(\frac{1}{2}\right) \]
\[ + \sum_{m>1/2} p(m) \left[ f\left(\frac{1}{2} - m\right) + mf(1 - m) - k(1 - m)f(m) \right] \]
\[ = \sum_{m<1/2} p(m) \left[ -k f\left(\frac{1}{2} - m\right) + mf(1 - m) - k(1 - m)f(m) \right] + p\left(\frac{1}{2}\right) \frac{1-k}{2} f\left(\frac{1}{2}\right) \]
\[ + \sum_{m<1/2} p(1 - m) \left[ f\left(\frac{1}{2} - m\right) + (1 - m)f(m) - kmf(1 - m) \right] \]
\[ = (1 - k) \sum_{m<1/2} p(m) \left[ f\left(\frac{1}{2} - m\right) + mf(1 - m) - (1 - m)f(m) \right] + p\left(\frac{1}{2}\right) \frac{1-k}{2} f\left(\frac{1}{2}\right) \]  
(A16b)

Note that we used \( \sum_{m>1/2} F(1 - m) = \sum_{m<1/2} F(m) \) for general function \( F \).

Hence,
\[ D = \Delta_2 - \Delta_1 = (1 - k) \sum_{m<1/2} p(m) \left[ f\left(\frac{1}{2} - m\right) + mf(1 - m) - (1 - m)f(m) - f\left(\frac{1}{2}\right) \right] \]  
(A17)
Condition 1: $f$ is not strongly convex

Since $k \geq 1$, a sufficient condition for $D > 0$ is

$$f\left(\frac{1}{2} - m\right) + mf(1-m) + (1-m)f(m) > f\left(\frac{1}{2}\right) \\forall 0 \leq m \leq \frac{1}{2} \tag{A18}$$

Define $F(m) = f\left(\frac{1}{2} - m\right) + mf(1-m) + (1-m)f(m)$. Note that

$$F(0) = F\left(\frac{1}{2}\right) = f\left(\frac{1}{2}\right)$$

$$F'(m) = -f'\left(\frac{1}{2} - m\right) - mf'(1-m) + f(1-m) + (1-m)f'(m) - f(m)$$

$$F'\left(\frac{1}{2}\right) = -f'(0) < 0$$

Based on these results, since $f$ (and hence $F$) is a continuous function, (A18) would be true if $F''(m) < 0 \\forall 0 \leq m \leq \frac{1}{2}$

$$F''(m) = (1-m)f''(1-m) + f''\left(\frac{1}{2} - m\right) - 2f'(1-m) - 2f'(m) \tag{A19}$$

It is not true that $F''(m) < 0$ at $0 \leq m \leq \frac{1}{2}$ for any convex function $f$. Here we would explore possible conditions where $F''(x) < 0$ is true. We note that

$$F''(m) = (1-m)f''(1-m) + f''\left(\frac{1}{2} - m\right) - 2f'(1-m) - 2f'(m)$$

$$= (1-m)[f''(m) - 2f'(m)] + m[f''(1-m) - 2f'(1-m)] - [2mf'(m) + (2-2m)f'(1-m) - f''\left(\frac{1}{2} - m\right)] \tag{A20}$$

Note that if $f''(m) \leq \frac{3}{2}f'(m)$, when $0 \leq m < \frac{1}{4}$

$$2mf'(m) + (2-2m)f'(1-m) > \frac{3}{2}f'(1-m) \geq \frac{3}{2}f'\left(\frac{1}{2} - m\right) \geq f''\left(\frac{1}{2} - m\right),$$

and when $\frac{1}{4} \leq m \leq \frac{1}{2}$,

$$2mf'(m) + (2-2m)f'(1-m) \geq 2f'(m) \geq 2f'\left(\frac{1}{2} - m\right) > f''\left(\frac{1}{2} - m\right).$$

Therefore, one can infer from eq. (A19) that

$$F''(m) < 0 \text{ if } f''(m) \leq \frac{3}{2}f'(m) \\forall 0 \leq m \leq \frac{1}{2}. \tag{A21}$$

Please note that eq. (A21) is a sufficient condition but not a necessary condition for $D > 0$. As we have shown in Figure 5, there are functions $f$ which do not fulfil eq. (A21) but still allows $D > 0$.

Eq. (A21) suggests that if is not very convex then, $D > 0$ is guaranteed. On the other hand, $D > 0$ can also be achieved if $f$ is strongly convex, as we will show below.
**Condition 2: f is strongly convex**

We investigate what happens if \( f \) is strongly convex. Starting from eq. (A17),

\[
D = (k - 1) \sum_{m < 1/2} p(m) \left[ f \left( \frac{1}{2} - m \right) + mf(1 - m) + (1 - m)f(m) - f \left( \frac{1}{2} \right) \right]
\]

\[
> (k - 1) \sum_{m < 1/2} p(m) \left[ mf(1 - m) - f \left( \frac{1}{2} \right) \right]
\]

(A22)

because \( mf(1 - m) \) would dominate the other neglected terms for strongly convex \( f \).

Next, we make use of the fact that \( f \) is convex. There are multiple ways of doing it, which will lead to different constraints on \( f \) for ambiguity aversion. Here, we show an example.

Applying the Jensen’s inequality, we have

\[
E^{(0)}[f(1 - m)] \geq f(1 - E^{(1)}[m]),
\]

where \( E^{(c)}[\cdot] \equiv \frac{\sum_{m < 1/2} m^c p(m)}{\sum_{m < 1/2} m^c p(m)} \). This leads to

\[
D > (k - 1) \left( \sum_{m < 1/2} p(m) \right) \left[ E^{(0)}[m]f(1 - E^{(1)}[m]) - f \left( \frac{1}{2} \right) \right].
\]

(A23)

Note that \( 0 < E^{(c)}[m] < \frac{1}{2} \) since \( 0 < m < 1/2 \) for both \( c = 0 \) and \( 1 \). Hence, \( D > 0 \) is guaranteed if \( f \) is sufficiently convex. More specifically,

\[
E^{(0)}[m]f(1 - E^{(1)}[m]) > f \left( \frac{1}{2} \right)
\]

(A24)

For example, for \( f(m) = m^r \) and \( p(m) = \frac{1}{2n+1} \) for all \( m \), one can easily show that \( D > 0 \) if

\[
r > \frac{\log \frac{4n}{3n+1}}{\log \frac{4n+1}{3n}}.
\]

(A25)

which implies a strongly convex function \( f \) with \( r > 4.82 \) in the limit of large \( n \). We would like to emphasize again that eqs. (A24) and (A25) are not necessary condition for \( D > 0 \). In fact, as we have shown in our numerical study (Figure 5), \( D > 0 \) can be achieved with a much less strongly convex function than eqs. (A24) and (A25) suggest.

**Reference**

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263–292. https://doi.org/10.2307/j.ctv1kr4n03.21

Shackleford (2019a). Blackjack. https://wizardofodds.com/games/blackjack/basics/#rules

Shackleford (2019b). Blackjack Hand Calculator. https://wizardofodds.com/games/blackjack/hand-calculator/
Supplementary figures:

Figure S1: The difference in surprise values between the branching scheme when event grouping is present and when it is absence for Option 2 in Allais paradox Problem 2 (left) and Option 2 in Problem 1 adopted from Birnbaum (2008) (right), which are discussed in the main text. The dark red line is the boundary where the options are equally preferred, i.e. \( \Delta_{\text{group}} = \Delta_{\text{no group}} \). Consistent with our analysis, event grouping reduces the appeal for Option 2 in Allais paradox for all \( k, r > 1 \). On the other hand, event grouping increases the appeal for Option 2 in the Birnbaum problem with a slightly more stringent condition that \( r \) is not too close to 1 and \( k \) is not too large. As with the rest of our numerical studies on these problems, we set \( f(x) = x^r \).