Trial order causes systematic misestimation of economic choice traits in humans.

Nidhi V. Banavar[1] and Aaron M. Bornstein[1,2]

[1]Department of Cognitive Sciences, University of California, Irvine

[2]Center for the Neurobiology of Learning and Memory, University of California, Irvine

Corresponding Author

Nidhi V. Banavar. 3151 Social Sciences Plaza A, Irvine, CA, 92697.
nbanavar@uci.edu

Preprint Server

Classifications

Social Sciences: Psychological and Cognitive Sciences

Keywords

Abstract

In fields spanning policy, medicine, and finance, it is increasingly common to guide individual interventions on measures of a person's decision-making characteristics. As the practical application of these measures grows, so have efforts to improve their robustness to ostensibly irrelevant contextual factors such as time of day or ephemeral motivational state. Here, we examine whether such instruments exhibit a fundamental context-dependence: namely, the order in which decision problems are presented. In three datasets evaluating decision-making under different forms of uncertainty, we find systematic, meaningful effects of trial order, which in many cases qualitatively change the measurement's psychological interpretation (e.g. from risk-seeking to risk-averse). We further examine how trial properties modulate this phenomenon and provide an augmented modeling framework to reliably characterize and correct for these effects.

Significance Statement

Clinicians, policymakers, and practitioners are increasingly making important decisions using lab tasks that measure how an individual makes choices under uncertainty (e.g. risk aversion). These tasks are usually randomized, to avoid systematic confounds of presentation order. Here, we examine this assumption across three studies and constructs – risk attitudes, ambiguity attitudes, temporal discounting — and find that individuals remain sensitive to trial order, even in randomized experiments. Accounting for this influence leads to meaningful quantitative and qualitative changes in inferred choice traits. Our findings have important implications for the measurement of individual differences in choice traits. In particular, they suggest that choice measurements should account for the fact that humans may treat the world as inherently sequential.

2

1 Introduction

Increasingly, important decisions about an individual are being informed by an assessment of how that individual makes decisions under uncertainty. For example, asset managers measure their clients' risk tolerance when deciding how to allocate their investment portfolios [1], clinicians use the same property to evaluate their substance-using patients' likelihood to relapse [2], and aid workers use measures of intertemporal choice preferences to personalize vaccination incentives [3]. The success of these high-stakes interventions depends on identifying stable traits that will be robust across time and setting. For instance, an intervention with long-term consequences should ideally not depend on whether the measurement was taken on Monday morning versus Wednesday afternoon, or if a local sports team has just won a championship. Indeed, temporal and situational variability in these measures has been widely demonstrated [4]. Critically, this variability has in some cases been related to directly observable endogenous mechanisms [5], and to predict clinically-significant behavioral outcomes [6], supporting the construct validity of these measures and further emphasizing the need to distinguish relevant from irrelevant sources of variability.

One approach to isolating meaningful variability has been to incorporate insights from decision neuroscience by modeling the time it takes an individual to make a decision [7, 8]. This approach produces more robust estimates of individual decision characteristics because it reveals aspects of the decision-making process that are obscured when examining choice alone – e.g. long decision times may indicate near-indifference between two options. Indeed, scientists have demonstrated that sequential sampling models of response times (RT) not only make predictions that capture widely observed effects in RTs during economic choice [7], but that these models can also be used to derive descriptive theories of choice such as Random Utility [9]. Further, researchers have shown that models that jointly model choice and RT can better describe data across multiple experiments [10] and that subject RTs from one experiment can predict behavior on another [8].

Most decisions are made under at least some degree of uncertainty and under constraints such as limited time or partial information. Choice behavior under uncertainty has therefore been formalized across multiple disciplines with behavioral economics establishing some of the most widely used theoretical and empirical frameworks. Key behavioral economic and psychological parameters of interest include *risk tolerance* (α) which captures how people behave when they have complete information about the underlying probability distribution(s) of reward, *ambiguity tolerance* (β) which captures how people behave when they have partial or no information about probabilities, and *discount factor* (k) which captures how people trade off time and reward. Critically, the number inferred when estimating these parameters is tightly linked to a psychological interpretation – especially in the case of risk and ambiguity.

When making decisions under uncertainty and with limited time, humans and animals must balance efficiency with completeness. One mechanism through which humans may do so is through the efficient coding of valuation information [11]. The efficient coding hypothesis, originally formalized in perceptual neuroscience, says that resource-constrained organisms leverage environmental structure to maximize information and minimize redundancy [12]. A corollary of this hypothesis is that stimulus information is encoded relatively (i.e. what has changed now compared to the previous moment?). This suggests that individuals should be sensitive to trial order over the course of an experiment: the decision they were presented with (and the choice they made on the previous trial) could

impact how individuals perceive the current choice problem and the subsequent choice they make. Such behavior would be incongruent with standard behavioral economic tasks which deploy a randomized structure paired with explicit instructions to participants that they should "treat each trial independently and as if it were the only one that counts." These sorts of task-incongruent temporal dependencies have been widely established in the visual perception literature where such behavior is termed serial dependence [13]. These researchers specify that such a dependency is separate from priming and hysteresis, likely not driven by higher-order processes [13], and is adaptive [14].
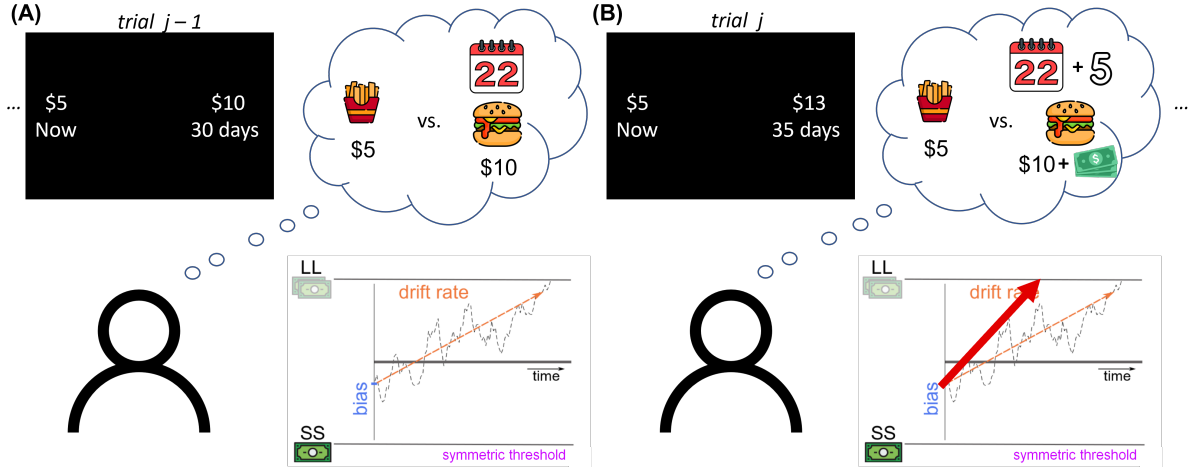

Figure 1: **Example of how trial order may influence decision making (choice and RT modeling parameters) in ITC**. The subject participates in the experiment on November 22nd. (A) On trial $j - 1$, they have to choose between \$5 today and \$10 in 30 days. They simulate their future 30 days from now, December 22nd, and imagine what they could buy. This is a difficult choice for the subject and deliberation is captured here by a slow rate of evidence accumulation. The subject ultimately decides to take the \$10 and wait. (B) On the next trial j, the choice is between \$5 today and \$13 in 35 days. Standard models assume that the previous trial should have no influence on the current problem. The spillover hypothesis predicts that a subject could perceive and evaluate the current decision relative to the preceding choice, effectively reusing the outcome of their previous simulation instead of starting afresh: here demonstrated by a steeper drift rate and then a repeated choice.

Taken together, it becomes plausible that one source of variability in inferring key behavioral economic parameters may come from trial order. Further, observed choice, RTs and sequential sampling parameters may also be meaningfully impacted by trial order: while trial-trial variability is accounted for in such models, rarely are influences of the recent past explicitly modeled (e.g. Figure 1). Such sequential dependencies are typically and unsurprisingly observed in "sequential" tasks like multi-armed bandits where learning rates over the course of an experiment are modeled (e.g. [15]). We emphasize that the phenomenon of interest here is a higher-order serial dependence or informational spillover, something that is usually not explicitly modeled in these higher-order tasks regardless of whether they are "sequential" or not. In this paper, we develop and deploy a framework that allow us to examine temporal dependencies as a function of decision problem properties, such as reward value, probability, and delay.

In this paper, we demonstrate that trial order, in particular relative differences in successive trial properties, affects decision-making under uncertainty. We show, across three experiments (total $n = 636$), that sequential effects modulate behavioral economic parameters estimated both by jointly modeling response time (RT) and choice behavior, and choice behavior alone (as is standard practice). Experiments 1 and 2 re-analyze previously collected data (as in [16], but with different model specifications). We further

designed and collected data for Experiment 3 to build on these analyses so that we could more closely examine the specific factors that give rise to sequential effects and how they can contribute to changes in parameter estimates. We demonstrate both high degrees of individual differences and high degrees of systematicity in these effects (i.e. consistency of the sign of sequential effects parameters in the task). Critically, in the third experiment, we show that accounting for such sequential effects not only results in different numerical parameter estimates, but also different psychological interpretations. These differences in interpretation go in both directions (i.e. some previously classified "ambiguity averse" subjects are reclassified as "ambiguity seeking" and vice versa). Taken together, these results suggest that widely-used measures of decision-making traits are mischaracterizing a sizable fraction of individuals, and that experiment randomization alone does not correct for this bias. We conclude with a discussion of how this systematic effect should be accounted for before such measures are used to guide meaningful interventions into individual lives.

2 Methods

2.1 Experiments and Data

None of the experiments presented involved the use of feedback where choice outcomes were realized over the course of the experiment (i.e. after each trial). In Experiments 2 (ITC) and 3 (AMB), subjects received feedback confirming only their choice selection (i.e. left or right option selected).

Experiment 1: Risky Choice (RISK). We model $n = 56$ subjects who made 80 choices between two lotteries in the gain and loss domain conditions (rewards range: Gain: $\$1 - \$100$, Loss: $-\$99 - \$0$, probability range: $1\% - 99\%$, Figure 2A). Gain and loss gambles were not intermixed (i.e. participants made their choices in two conditions: 40 in the gain domain, condition 1, and then 40 choices in loss). All data were collected previously on Amazon Mechanical Turk (for details, see [17]). The experiment was fully randomized with no experimentally-designed trial-level dependencies. Stimuli were displayed numerically (rewards and probabilities) and graphically (probabilities were also presented as pie charts). Subjects had an unlimited amount of time to make their choices, which they indicated via mouse click. In this experiment, subjects were instructed to maximize rewards in the gain domain, and minimize losses in the loss domain.

Experiment 2: Intertemporal Choice (ITC). We model $n = 482$ adult subjects who made a sequence of 102 binary choices between a same-day monetary reward (SS: "smaller sooner", range: $\$1 - \$85$) and a larger delayed reward (LL: "larger later", range: $\$10 - \$95$; delay range: $4 - 180$ days, Figure 2B). All data were collected previously in person (for details, see [18]). The experiment was fully randomized with no experimentally-designed trial-level dependencies. All stimuli were displayed numerically and counterbalanced so that the SS and LL options occurred equally often on either side of the computer screen. Subjects had 6s after stimulus onset to make a choice. Each choice was followed by 0.5s of feedback confirming the option selected and then a variable inter-trial-interval (ITI) between 3-5s. In this experiment, subjects were instructed to act in accordance with their genuine preference between the two choice options (i.e. there was no "correct"' answer unlike with Experiment 1 (RISK).) The task was also incentive compatible: a single trial was selected at random, realized, and paid out at the end of the experiment as a bonus.

Experiment 3: Risky and Ambiguous Choice (AMB). We model $n = 98$ adult subjects who made a sequence of 196 binary choices between a certain reward (range: $3 − $9.5) and a lottery (range: $5 − $24), in 4 blocks (Figure 2C). The amount a subject could win by choosing the lottery was almost always larger than the certain reward, except during 16 catch trials (4 per block). All data were collected on Amazon Mechanical Turk via psiturk [19]. Lotteries were either *risky* (1/7 of trials) where the full probability distribution was presented graphically (win probabilities: 25%, 50%, 75%) or *ambiguous* (6/7) where partial information was presented (ambiguity levels: 15%, 40%, 60%, 85%). While we maintained a general "non-sequential" structure, we ensured that 50% of successive trials increased in ambiguity, and 50% decreased. A risk trial followed by an ambiguous trial would be considered as an increase in ambiguity, as risky trials are unambiguous with respect to the probability of reward. Likewise, an ambiguous trial followed by a risk trial would be considered a decrease in ambiguity. Further, we controlled for median risk/ambiguity levels, lottery reward, and fixed reward across blocks. As with ITC, the stimulus options occurred equally often on either side of the computer screen. Subjects had up to 3s after stimulus onset to make a choice. Each choice was followed by 0.5s of feedback confirming the option selected and then a variable inter-trial-interval (ITI) between 0.5-2s. The task was also incentive compatible: a single trial was selected at random, realized, and paid out at the end of the experiment as a bonus. Participants on Amazon MTurk were bonused 10% of their winnings to be consistent with pay rates on the platform.
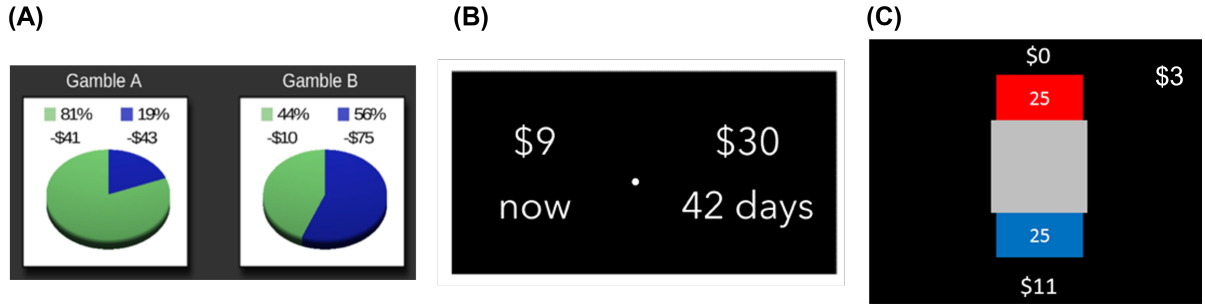

(A)   **(B)**   **(C)**

Figure 2: **Example trials for all three tasks**. (A) E1: Example RISK trial in the Loss domain where a subject has an unlimited amount of time from stimulus onset to choose between either the gamble on the Left or the Right. The Expected Value maximizing (and correct) choice is the gamble on the Left (figure from [17]). (B) E2: Example ITC trial where the subject has up to 6s to chose between $9 today or $30 in 42 days (figure from [18]). (C) E3: Example AMB ambiguous trial where the subject has up to 3 seconds to make a choice between a certain reward of $3 and a chance to win $11 by playing the lottery. The Expected Value maximizing (but not necessarily "correct") choice is the certain reward.

We model both choice behavior and response times (RT) in the ITC and AMB tasks. We only model choice behavior in the RISK task, as RT data was not available. For ITC and AMB, we excluded any responses that were made in less than 300ms. We also excluded any missed trials and the trial immediately after them from the following analyses, in addition to the first trial in each block. This is because our primary analysis focuses on one-trial-back effects. Finally, we excluded any subjects that missed more than 25% of trials.

2.2 Models: Choice Behavior

2.2.1 Baseline ("Non-Sequential") Models

For tasks involving immediate uncertainty (RISK and AMB), we pair a logistic choice rule with models consistent with the Subjective Expected Utility Maximization Framework. We implement these models in a hierarchical Bayesian fashion using JAGS [20] to better capture individual differences [21]. Unless otherwise stated, all parameters are hierarchical Normals, with hyperprior specifications of mean $\mu \sim Normal(0,1)$ and standard deviation $\sigma \sim Uniform(0.01, 4)$. A parameter X is thus hierarchically distributed: $X \sim Normal(\mu_X, \sigma_X^2)$.

To infer risk and ambiguity values for the RISK and AMB tasks, we model Subjective Value as follows:

$$Subjective\ Value\ Lottery_{Risk} = p \cdot v^\alpha \qquad (1)$$

$$Subjective\ Value\ Lottery_{Ambig} = [p - \beta(A/2)] \cdot v^\alpha \qquad Subjective\ Value\ Fixed = v^\alpha$$

Here, p is the objective probability of a reward (risk level, or $p = 0.5$ on ambiguous trials per [22]). A represents the degree of ambiguity on the trial (ambiguity level, $A = 0$ on risk trials and for all of the RISK task). Finally, v represents the monetary reward associated with that lottery. The key behavioral economic parameters of interest, then, are α which is risk tolerance, and β which is ambiguity tolerance. In this formulation, both parameters are subject-specific. We use hyperprior $\mu_\alpha \sim Gamma(2,1)$ for risk tolerance, with mode $= 1$ (risk neutrality) and $\mu_\beta \sim Normal(0,1)$ for ambiguity tolerance, with mode $= 0$ (ambiguity neutral). For the RISK task, we further allow for the curvature of the utility function to differ as a function of domain c (i.e. infer $\alpha_{(i,c)}$), where $c = 1$ for gain domain and $c = 2$ for loss.

To infer temporal discounting for the intertemporal choice task, we model Subjective Value using a hyperbolic discounting function:

$$Subjective\ Value\ Future = \frac{v}{1 + k \cdot D} \qquad (2)$$

Here, v again represents the monetary reward associated with that lottery and D represents the delay with the future (LL) reward is offered. The key behavioral economic parameter of interest is k which is the individual's discount factor. We use hyperprior $\mu_k \sim Beta(1,1)$ to be "uninformative."

Our logistic choice rule, which has the same basic parametrization across all three tasks, for subject i on trial j is as follows, where $\theta_{A,B(ij)}$ is the probability of choosing Option A:

$$\theta_{A,B(ij)} = \frac{1}{1 + \exp\left(\gamma_i + \phi_i \cdot SVD_{ij} + \epsilon_{ij}\right)} \qquad y_{A,B(ij)} \sim Bernoulli(\theta_{A,B(ij)}) \qquad (3)$$

Here, SVD_{ij} represents the difference in the Subjective Value between the two options presented on any given trial. γ_i represents the shift, or bias, in a decision. ϕ_i represents response variability, and we use hyperprior $\mu_\phi \sim Gamma(2,1)$, where the mode corresponds to probability matching. Finally, ϵ_{ij} represents effects of simple choice – or motor – perseveration (repeat previous choice made). All parameters allow for variability at the individual and, in the case of the RISK task, domain (gain or loss) level. We pair these prior specifications with a $Bernoulli(\theta_{A,B(ij)})$ likelihood, as no two stimuli are presented together more than once.

2.2.2 Sequential Effects

Intuitively, we might imagine that there would be more (less) of an effect on a given parameter on sequential trials that present the subject with similar (different) values for the decision problem regardless of task structure or goals: e.g., if on RISK trial j-1, a subject decides between a 81% chance of winning $41 or a 55% chance of winning $39, and the next trial j asks the subject to choose between a 85% chance of winning $45 or a 55% chance of winning $37, there might be little need to re-deliberate, which could thus yield an effect on either choice or response time (refer to Figure 1 for example in ITC). This may be the case even if subjects are told to treat decisions independently, as they typically are in these "non-sequential" behavioral economic tasks, and trials themselves are randomized.

This task-incongruent transient reliance on recent history (henceforth also referred to as perseveration) can manifest in one of many ways, ranging from "lower-order" (e.g. motor) to "higher-order" (e.g. carrying over frontal cortex dependent computations) [23, 22]. We explicitly test for three qualitatively different types of perseverations (Table 1 *Lower*): motor or choice (i.e. repeat the same choice previously made), perceptual (i.e. on logistic bias or variability), and computational (i.e. influencing people's risk/ambiguity/impulsivity preferences). We define these perseverations as driven explicitly by cross-trial *differences* in the trial properties listed in Table 1 *Upper*, as these are the normative properties involved in choice computations. As we further expect these influences to be transient due to the complex and temporally constrained (ITC, AMB) choice environment, we restrict our analyses to one-trial-back (i.e. differences in current and immediately preceding trial properties).

Property Type	RISK	ITC	AMB
Uncertainty	Gamble Entropy (H)	Delay (D)	Risk/Amb levels (RA)
Reward		value	
Normative	Expected Value (EV)		EV
Interaction	EV x H	value x D	EV x RA

Parameter	
Logistic Bias	ϕ
Logistic Slope	γ
Choice Perseveration	ϵ
Risk Tolerance	α
Ambiguity Tolerance	β
Discount Factor	k

Table 1: **Trial properties considered for sequential effects and the choice parameters on which we test for these effects.** *Upper* Differences in trial properties considered as potential drivers of sequential effects. $H = -\Sigma p \log(p)$ is the Shannon Entropy of a gamble. *Lower* Parameters we simultaneously test for sequential effects by allowing them to vary trial-trial as a function of relative differences in properties as described in Table 1 a. We test the first three parameters in all three tasks. The final three are task-dependent (RISK: α; ITC: k; AMB: α, β).

Specifically, we use Indicator Variables to subset increases or decreases in specific sequences of trials, resulting in a 8-fold tiling of trial property space for all three experiments (e.g. for ITC: increase in delay, decrease in delay, increase in value, decrease in value, increase in value and delay etc.) We then augment our baseline models by allowing these

trial properties to exert linear additive influences on the parameters of interest. For example, if we consider ITC trials that increase in value from trial j–1 to trial j:

$$\gamma'_{0,ij} = \gamma_{0,i} + \eta_i \cdot 1([V_{a,j} - V_{b,j}] > [V_{a,j-1} - V_{b,j-1}]) \tag{4}$$

Thus, in Equation 4, $\gamma_{0,i}$ becomes the sequential-effect-adjusted logistic bias for individual i and the indicator variable is 1 if there is an increase in value difference from trial $j - 1$ to trial j. The posterior value of η_i tells us how much an individual is weighting relative changes in trial properties (i.e. increase or decrease). We define η_i hierarchically: $\eta_i \sim \text{Normal}(\mu_\eta, \sigma_\eta)$.

Our fundamental analysis is structured around a hypothesis test on η_i: $H_0 : \eta_i = 0$ vs. $H_a : \eta_i \neq 0$. We quantify the strength of evidence in favor of the null and alternative by using the Savage-Dickey ratio to estimate the Bayes Factor (BF). This ratio compares prior and posterior density at any point in parameter space (i.e. $\eta_i = 0$). As others have done, we interpret any values of $BF > 3$ as evidence in favor of our alternative hypothesis [24].

2.3 Models: Response Times

Recall that we only have response time data for ITC and AMB.

2.3.1 Baseline Models

We implement a hierarchical Bayesian drift diffusion model (DDM) to model response times using the Wiener module [25] in JAGS [20] for both the ITC and AMB tasks. That is, for subject i and trial j, we model observed response time as Wiener first passage time *(wfpt)* distributed:

$$RT_{ij} \sim wfpt(\alpha_i, \tau_i, \beta_i, \delta_{ij})$$

Here, α_i represents the subject-level threshold or boundary separation, τ_i is the subject-level non-decision time (processes ostensibly unrelated to the value-based decision process), β_i is the subject-level bias ($\beta_i < 0.5$ bias towards immediate option in ITC and towards the fixed option in AMB), and δ_{ij} is the subject-and-trial-level drift rate (the rate of evidence accumulation). We model all these parameters as hierarchical Normals in order to better capture individual differences [21]. For $\alpha_i, \tau_i,$ and β_i, we use the same prior and hyperprior specifications for both tasks, referencing [25] for mean hyperpriors and using 'noninformative' priors for the standard deviation:

$$\mu_\alpha \sim Uniform(0.001, 3) \quad \mu_\tau \sim Uniform(0, 0.6) \quad \mu_\beta \sim Uniform(0.01, 0.99)$$
$$\sigma_\alpha, \sigma_\tau, \sigma_\beta \sim Uniform(0.01, 4)$$

Similar to previous work [10], we take a cognitive psychometrics approach to modeling the drift rate. Critically, however, we allow the drift rate to be driven by *objective* trial properties (i.e. not Subjective Value or even differences in SV) and normative combinations of these trial properties (e.g. Expected Value). This is because explicitly relating untransformed trial properties to elements of the decision process is critical for our question of interest. Incorporating transformed trial properties like Subjective Value might perpetuate the very biases we seek to mitigate as they would be inferred without accounting for potential effects of trial order. We keep the broad functional relationship between

trial properties as dictated by behavioral economic models of choice behavior (e.g. allowing an inverse relationship between the drift rate and delay for ITC). We also normalize all stimulus properties such that they fall between 0 and 1. Then, for subject i and trial j:

$$\delta_{ITC,ij} = \beta_{0,i} + \beta_{1,i} \cdot (value_{LL,ij} - value_{SS,ij}) + \beta_{2,i} \cdot delay_{ij}^{-1} \tag{5}$$

$$\delta_{AMB,ij} = \beta_{0,i} + \beta_{1,i} \cdot (EV\ Diff) + \beta_{2,i} \cdot A_{ij} \tag{6}$$

In Equation 6, $EV\ Diff$ represents the difference in Expected Value ($EV_{Lottery} - EV_{fixed}$) between the two options presented. $EV_{Lottery} = p \cdot v$, where p is the objective probability of reward (i.e. risk level) and v retains its interpretation of objective reward. For an ambiguous trial, $p = 0.5$ as in [22]. The Expected Value of a certain reward is simply v. On ambiguous trials, A is the function of the lottery that is occluded by the grey bar as seen in Figure 2C. Recall that on risky trials, A=0. Finally, we allow all drift rate decomposition parameters βs to be hierarchical Standard Normals.

2.3.2 Sequential Effects

We allow all sequential effect parameters to be hierarchical Standard Normals. We simultaneously assess the influence of relative trial properties on all *drift rate decomposition* (i.e. βs) and *bias* parameters (same properties as in choice, Table 1A). For example:

$$\beta'_{0,ij} = \beta_{0,i} + \eta_i \cdot 1([V_{a,j} - V_{b,j}] > [V_{a,j-1} - V_{b,j-1}]) \tag{7}$$

Thus, in Equation 7, $\beta_{0,i}$ becomes the sequential-effect-adjusted drift rate "regression" intercept for individual i and the indicator variable is 1 if there is an increase in value difference from trial $j-1$ to trial j. Just as in Choice Behavior (section 2.2.2), we test whether the sequential effect parameters (i.e. η_i) is non-zero using the Savage-Dickey ratio to approximate the Bayes Factor (BF). We interpret any $BF > 3$ as evidence in favor of sequential effects.

3 Results

We analyze data from two previous experiments (Experiment 1 (RISK): Risky decision-making task in gain and loss domains n=56; Experiment 2 (ITC): Intertemporal choice task n=482) and one new experiment (Experiment 3 (AMB): Risky and ambiguous decision-making task in gain domain n=98) (Figure 2). We explicitly incorporate sequences of trial properties into standard choice and response time models using hierarchical Bayesian modeling, considering one "sequence type" at a time. Critically, this approach allowed us to estimate the reliability of sequential effects at an individual-subject level.

3.1 Sequential Effects in Choice Behavior

We first assessed whether choices exhibited trial property-dependent sequential effects at a broad, "model-free," level. Our primary assessment of model free signatures of sequential effects in the RISK and AMB tasks involved comparing whether the proportion of times a subject selected a choice differed across trial sequence types. The trial properties of interest include cross-trial differences in Expected Value and entropy (see table 1A for complete list). Specifically, we look at pairs of trials as differences of differences: for

10

example, an increase in Expected Value difference between successive trials means that the choice options on the current trial are more distinct than the choice options on the previous trial. This could correspond to a current trial being relatively easier than the previous. This interpretation is consistent with the experimental data on a whole.

In RISK, subjects chose the EV maximizing option significantly more often not just when a trial was relatively easy with respect to the rest of the experimental choice set, but also when there was a relative increase in EV from the previous trial to the next (EV increase 71%, vs. EV decrease 64%, Wilcoxon Rank Sum $V = 1276.5$, $p < 0.001$), but no significant difference when comparing proportions of EV maximizing choices when trials increase or decrease in gamble entropy (H increase 67%, H decrease 68%, $V = 705$, $p = 0.45$, Figure 3A). In AMB, we find a similar pattern: subjects chose the lottery option significantly more often when there was a relative increase in EV from the previous trial to the next vs. when there was a relative decrease (EV increase 50%, EV decrease 47%, Wilcoxon Rank Sum $V = 3693$, $p < 0.001$), but not when considering only increases or decreases in gamble ambiguity (Amb increase 48%, Amb decrease 50%, $V = 2441$, $p = 0.43$, Figure 3B). We find no differences in choice proportions as a function of trial properties (value, delay increase/decrease) in the ITC task. This simple analysis suggests that, in at least two of the three types of choices evaluated, people are not just sensitive to trial order but more specifically to the relative differences in normative calculations - the relative ease or difficulty of the current choice problem.
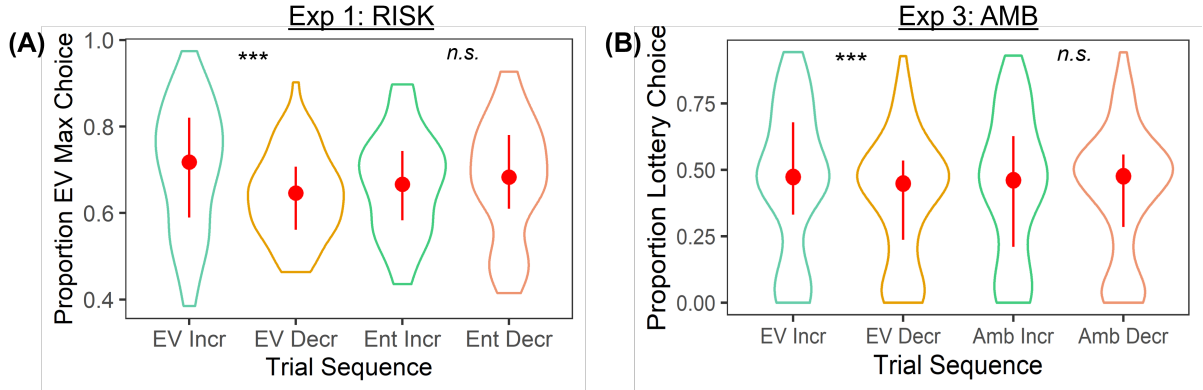
Figure 3: **Participants in RISK and AMB tasks make distinct choices depending on trial sequence.** (A) In the RISK task, they are more likely to make the EV maximizing choice between two gambles when the previous trial was more difficult (i.e. choices less distinct). (B) In AMB, they are more likely to pick the lottery option as opposed to the fixed when the previous choice was easier (i.e. less difference between options). Conversely, there was no significant difference between EV maximizing choices when considering relative increases or decreases in trial entropy (A) or trial ambiguity (B). Overlaying the violin plots are the median and IQR.

For our model-based analyses, we augment logistic choice rules to include such relative cross-trial differences on multiple parameters as described in Methods. That is, we allow for trial order to potentially manifest as a variety of qualitatively different types of "perseverations" – from motor, to perceptual, to cognitive. In RISK, we observed reliable sequential effects on logistic slope and risk tolerance for 7% of individuals. Critically, and consistent with these effects being cognitively specific, these individuals only had non-zero sequential effects for specific sequences of trials: when a trial with a high difference in Expected Value between the two options was followed by a trial with a low difference in EV – "easy" then "difficult" in sequence. For example, for a specific subject, an initial risk tolerance of 1.043 updated to 1.205 when adjusted for this sensitivity. Importantly,

the magnitude of risk tolerance was not the only changing factor: the interpretation of the individual's risk tolerance changed from risk neutral to risk averse in the loss domain. For the other two experiments, unlike our RISK analysis, our hypothesis test of the posterior values of the sequential effect terms did not result in any evidence in favor of trial-trial sensitivity. However, we note that we also did not find strong enough evidence in favor of the null hypothesis.

3.2 Choice and Response Time

3.2.1 Drift Rate Decompositions

ITC	Mean(CI)	AMB	Mean(CI)
β_0	$-0.66(-1.26, -0.06)$	β_0	$-0.90(-1.12, -0.68)$
β_1	$3.45(1.19, 5.70)$	β_1	$1.51(1.04, 1.94)$
β_2	$2.34(-2.15, 6.84)$	β_2	$0.03(0.001, 0.1)$
α	$2.47(1.70, 3.24)$	α	$1.84(0.96, 2.71)$
τ	$0.76(0.36, 1.16)$	τ	$0.50(0.03, 1.03)$
$bias$	$0.51(0.39, 0.63)$	$bias$	$0.50(0.4, 0.6)$

Table 2: **ITC + AMB: Drift rate decompositions capture meaningful variance in both ITC and AMB.** Each cell shows aggregate posterior means (95% Credible Intervals) for Drift Rate decompositions and other DDM parameters. Bolded parameters are ones we also test for sequential effects.

We find that for the ITC task, subjects tend to accumulate evidence more quickly when the value difference increases (β_1), all else held constant. Similarly, subjects tend to accumulate evidence more quickly when the delay decreases (β_2) – recall that we parameterize delay as $delay_{ij}^{-1}$ – all held constant. Both make sense intuitively, as larger value differences might push individuals towards selecting the LL option, and delayed rewards offered in the far future may not be worth the wait.

With the AMB task, the average subject's drift rate increases as the Expected Value difference between choice options increases (β_1), all else held constant. Subjects seem nominally sensitive to the degree of Ambiguity during evidence accumulation (β_2). We also highlight the credible interval ranges to confirm that there are considerable individual differences, as we might expect. We note that we considered many different combinations of trial properties as drivers of drift rate and present these model results (Equation 5, 6) as they best fit the data.

We interpret the threshold and non-decision time parameters being greater for ITC to be a reflection of differences in task structure. Recall that subjects have up to 6s to respond in the ITC task, whereas they only have up to 3s in the AMB task. Finally, aggregate posterior means suggests that subjects are generally unbiased in both tasks, which is in contrast with their choice behavior: in ITC subjects chose the smaller sooner reward more frequently and in AMB chose the certain reward more frequently.

3.2.2 Sequential Effects

As described above, response time models provide more reliable estimates of individual-subject decision processes. We therefore analyzed trial-type-specific effects using an augmented drift-diffusion model in place of the logistic choice rule. We performed these analyses for ITC and AMB as RT data was unavailable for RISK. This approach allowed

us to identify sequential effects at the level of both bias (pre-choice inclination) and deliberation (evaluative processing of choice properties).

We observed reliable trial-property-driven sequential effects on the DDM decomposition (Equations 5, 6) and bias terms for both ITC and AMB tasks. For almost every possible combination of stimulus sequences, 100% of subjects showed evidence of sensitivity to previous stimuli in the ITC task (Table 3, Table S1). Effects in the AMB task are more specific to the combination of trial properties considered – in particular, individuals seem to be particularly sensitive to relative cross-trial differences in Expected Value (i.e. the choice becomes easier or harder) (Table 3, Figure S1, Table S2). Interestingly, while in the ITC data we find effects on both the drift rate and bias terms, in AMB we only find evidence of sequential effects on the drift rate terms which suggests differences in the cognitive process that may not only drive spillover but also relate to the choice problem at hand (evaluating future uncertainty vs. present).

ITC	Proportion	AMB	Proportion
$value \uparrow$	1	$value \uparrow$	0.80
$value \downarrow$	0.998	$value \downarrow$	0.93
$delay \uparrow$	1	$amb \uparrow$	0.05
$delay \downarrow$	1	$amb \downarrow$	0.01
$v. \uparrow d. \uparrow$	1	$v. \uparrow a. \uparrow$	0
$v. \uparrow d. \downarrow$	1	$v \uparrow a. \downarrow$	0.18
$v. \downarrow d. \uparrow$	1	$v. \downarrow a. \uparrow$	0
$v. \downarrow d. \downarrow$	0.84	$v. \downarrow a. \downarrow$	0.57

Table 3: **ITC + AMB: Most participants are sensitive to trial sequences.** The majority of subjects show sensitivity to trial sequences in both ITC and AMB tasks. Each cell shows the proportion of subjects that demonstrated sequential effects ($BF > 3$) on any one of the drift rate decomposition parameters or bias. Each row represents specific successive trial properties (e.g. $value \uparrow$ subsets successive trials that increased in value difference as noted in Equation 4). The top four rows can be thought of as "main effects" of specific trial properties and the bottom four "interactions."

3.2.3 Changes in parameter estimates.

The reader, especially if they are interested in applications of behavioral economics, may wonder why accounting for sequential effects should matter beyond finding varied individual differences. First, magnitude information is valuable especially for researchers interested in temporal fluctuations in ambiguity/risk attitudes. We further argue that adjusting for these effects can meaningfully change inferences on the parameters we care about. We thus examined how accounting for sequential effects can reveal qualitative changes in the characterization of individuals. We fit effect terms capturing spillover across all "main" (Table 3) trial sequences at the same time (as opposed to one-at-a-time), generate 1000 datasets from these model fits, and redeploy a logistic choice rule to infer risk and ambiguity tolerance as in the standard approach. That is, for the AMB task, Equation 7 is instead written as follows (see Table 1a for properties for all trials):

$$\beta'_{0,ij} = \beta_{0,i} + \eta_{i,01} \cdot \text{EV Increase} + \eta_{i,02} \cdot \text{EV Decrease} + \eta_{i,03} \cdot \text{Amb Increase} + \eta_{i,04} \cdot \text{Amb Decrease} \tag{8}$$

We refer to these models as "stacked" as they include all potential "main effects" of trial properties. We report the models that contained only the main effects for two reasons: 1) these were the primary drivers of sequential effects in our above analysis for both tasks – especially AMB – and 2) we wanted to avoid "over-parametrization" by including the interaction terms. As this parametrization adds 12 more variables to the
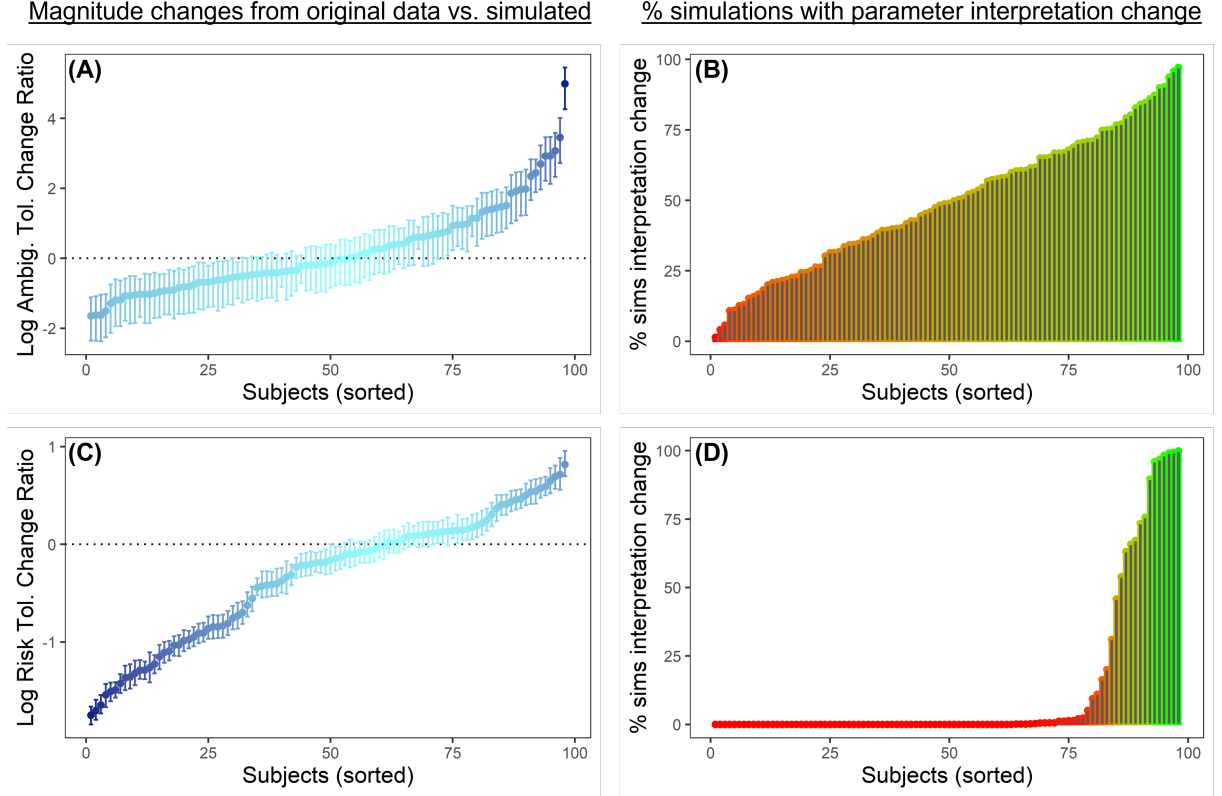
Figure 4: **AMB task: Magnitude and interpretation changes in (A, B) ambiguity and (C, D) risk tolerance.** *Left*: Ratio of (A) ambiguity tolerance and (C) risk tolerance estimates: log(simulated choice set/observed data). We plot median ratios and IQRs from 1000 simulated choice sets. *Right*: The percentage of simulation-fit parameters that change interpretations in (B) ambiguity and (D) risk attitudes when compared to parameter fits in the original data. Subjects re-sorted by effect size in each plot.

model, we define the sequential-effect terms (ηs) as non-hierarchical Standard Normals: $\eta_{i,xx} \sim \text{Normal}(0, 1)$ to aid model convergence. All models reported converged according to standard metrics.

Differences between the original parameter estimates and those from the simulated data would strongly suggest that cross-trial temporal dependencies are capturing something fundamental in human decision making under uncertainty. We report ratios between simulated data parameter estimates and the baseline model. If $0 < |\text{Ratio}| < 1$, then the sequential effect adjusted parameter estimates are numerically smaller than than than the initially inferred, and $|\text{Ratio}| > 1$ the converse. The median ratio change for risk tolerance across all subjects is 0.857 ($IQR = 0.716$), suggesting that models without sequential effects tend to overestimate an individual's risk tolerance. Likewise, the absolute value of median ratio change for ambiguity across all subjects is 0.574 ($IQR = 1.13$).

Critically, we find evidence not just of differences in sequential-effect-adjusted parameter estimates (Figure 4A,C) but also widespread qualitative changes in interpretation (Figure 4B,D). For ambiguity tolerance, this change moves in both directions (i.e. some subjects are newly classified as ambiguity averse or ambiguity tolerant), but for risk tolerance, all reclassified subjects move from risk seeking to risk averse.

14

4   Discussion

Measures of individual decision-making traits are becoming increasingly used in applied settings. However, there are many sources of variability that can bias these measures. In this paper, we present a quantitative behavioral analysis of one of these potential sources of bias: trial-trial dependencies in experiments without feedback. We build off a neuroeconomics literature that demonstrates the importance of trial order in how neurons represent value signals [11], and visual perception studies [13] that establish serial dependence. By explicitly incorporating trial-order information into behavioral models we find: a) differential sensitivity across experiments that involve different types of stimulus-level uncertainty, b) greater prevalence across subjects when jointly modeling choice and response time, c) evidence accumulation is impacted by trial-order across tasks, and d) individual-level posterior inferences on key parameters of interest can be meaningfully changed by accounting for these effects.

Our analyses highlight that sequential effects in both choice behavior and reaction time are a function of individual differences, with non-trivial changes in parameter magnitude and interpretation. This work has deep theoretical and empirical implications. Firstly, our analyses suggest that these sequential effects are not noisy artefacts but are instead the consequence of a systematic influence of trial properties on components of the decision process. This suggests a potential need for the theoretical re-conceptualization of *experimentally-inferred* parameters as explicitly dynamic and sensitive to (highly) local contexts and not *solely* a static and psychologically interpretable end. Secondly, much work has shown that parameters inferred from experiments tend to correspond poorly with real-world behavior and with other tasks that purportedly measure the same construct (e.g. sequential risk tasks vs non-sequential) [26] and the approach demonstrated here may be one way to more closely reconcile these discrepancies. Thirdly, there is much valuable information that can be gained from the joint modeling of choice and response times, and scientists should aim to collect RTs whenever possible [7, 8, 27]. Finally, the widespread nature of our results could suggest that susceptibility to sequential effects in complex choice under uncertainty could be a by-product of the rational use of limited resources [28].

It is tempting to interpret these sequential effects as "corrupting" the choice process and parameter inference, despite possible rational justifications [28, 29]. While we show that our posterior parameter estimates do indeed differ when we account for relative differences in trial properties, we suggest three empirical alternatives to how we can move forward with this information. Firstly, as suggested in the previous paragraph, we do explicitly consider the degree of sequential dependency to be informative and report two parameters with every task: the parameter the experiment was built around (e.g. risk tolerance) *and* some parameter or summary that reflects sequential sensitivity. Secondly, experimenters could consider interleaving distractor tasks to clear the working memory buffer that may generate sequential effects. Thirdly, in a similar vein, experimenters may consider merging relatively unrelated experiments together so that each subsequent trial may come from any one of n tasks. While some may argue that this strategy is folly to the same idea behind how experiments are currently designed – that randomized structure does more than what we in this paper demonstrate it does – this is why we suggest unrelated but relevant (i.e. non-distractor) tasks. In this paper, we remain agnostic about the circumstances and mechanisms through which sequential dependencies are generated.

Results from the AMB study suggest that differences in reward magnitudes play an

important role in these effects: people seem to be evaluating current rewards relative to what they had immediately seen before. One possibility as to why this could be the case is if humans, as other work suggests, encode magnitudes noisily and logarithmically [30]. A logarithmic internal representation of rewards presented in the experiment would have smaller rewards more closely represented on the "number line" (allowing for greater discrimination) and larger rewards to be disproportionately further apart. A natural consequence of using such a number line would mean that relative differences between successive sets of choice options, especially for options on opposing ends of the line, would be perceived to be greater than their absolute magnitudes. This is akin to a transient [numerical] contrast effect. In fact, this line of work by Khaw and colleagues argues that hallmarks of human behavior under uncertainty – like risk aversion in the gain domain – can actually be better explained by these (Bayes optimal) noisy logarithmic coding models relative to standard power utility models [30].

This is not to suggest that relative differences in how uncertainty in these experiments is represented internally do not also contribute to the generation of sequential effects. In this paper, we have considered three qualitatively different types of uncertainty: immediate with complete probability information (risk), immediate with incomplete probability information (ambiguity), and temporal with 100% probability of reward (temporal discounting). The cognitive processes invoked in managing these types of uncertainty, especially in sequence, may vary as a function of task – in particular risk/ambiguity vs. temporal discounting. Indeed, popular theories of intertemporal choice involve simulating the future [23]. This is a potentially a resource-intensive process that conflates uncertainty and concreteness – simulations further out in the future are also less likely to be concrete [31]. This may lead to the reuse or indeed reification of a simulated future that is n or greater days from the present. While such a future may not directly lead to one-trial-back sequential effects as patently as logarithmic representations of numbers may, it may form some sort of reference point that can influence perceptions of relative increases or decreases in delay. It is also worth noting that in this ITC experiment, as is standard, delay is represented only in a numerical form (i.e. the delayed reward will be offered n days in the future, see Figure 2). This is in contrast to RISK and AMB experiments, where there are both numerical and graphical representations of uncertainty. Thus it may be reasonable to interpret the ubiquity of sequential effects in ITC relative to AMB as possibly also driven by the purely numerical presentation of uncertainty (suggesting that uncertainty could be coded on a similar noisy logarithmic scale) and/or the re-use of computationally expensive simulated futures. Conversely, we may interpret the relative lack of sequential effects driven by differences in ambiguity in the AMB task as partially a consequence of scale anchoring due to the graphical presentation of risk/ambiguity. More consequentially, the experiment design included 50% trial sequences that increased in ambiguity and 50% trial sequences that decreased in ambiguity. If sequential effects are a type of adaptation effect, they may be washed out by this experiment design (the only factor we explicitly manipulated to be 50-50). Finally, unlike with ITC, there are no clear theories as to what cognitive processes are invoked during the decision process, lending further credence to the potentially important role of graphical presentation – which itself can vary over experiments – of uncertainty.

We are further agnostic about the processes involved in propagating these sequential dependencies. Researchers studying serial dependence in visual perception argue that attention plays a critical role [13] in addition, possibly, to working memory [14]. The relationship between attention and value-based decision making has been well studied

16

[32], with some researchers arguing that attention plays a causal role in the formation of value. Further, researchers have also demonstrated evidence for "last fixation bias," where the first fixation on a given trial is shaped by what the individual was last looking at on the previous trial, in value-based decision making [27]. In related but distinct work, researchers have used eye tracking (the most commonly used proxy for measuring attention) to demonstrate individual differences in strategies for an ITC task [33]. In particular, Khaw and colleagues demonstrate not only that search strategy is a predictor of an individual's discount rates, but also that search strategy can be shaped by tweaks in experimental design [33]. Taken together, attentional processes may indeed play a critical role in generating, or propagating, these experiment-design-dependent but task-incongruent spillover effects. Less studied, though equally plausible, is the relationship between working memory and economic decision making under uncertainty. Researchers have suggested that persistent activity in cortex, a typical signature of working memory, also supports value-based decision making [34]. As attention and working memory putatively operate on different time scales, it is entirely possible that both processes are involved in the phenomenon studied in this paper. Further research, and perhaps multi-modal data is necessary to be able to tease apart differential contributions of either process.

Finally, we return to the potential clinical implications of this research. Scientists have established for decades that there are exist meaningful differences between "healthy" controls and clinical populations in working memory and attention (e.g. aging [35, 36] and schizophrenia [37, 38]). We note that the degradation of working memory in aging appears to be more established, and less selective, than that of attention and aging. This opens up an interesting line of research in both health and disease: how to study the presence of sequential effects, recoverable through the joint modeling of choice and RT, to make more nuanced subject level inferences about health status, especially in disorders that demonstrate degradations in the processes that putatively support sequential effects?

Overall, in this paper we have demonstrated that the near-ubiquitous assumption in modeling economic choice, that choices made in sequence can be treated as independently made due to experiment randomization, is false. We show that this is the case through the joint modeling of choice and RT and show that our model puts forth meaningfully different parameter estimates than standard choice models. This work takes an important initial step in quantifying the effect of a heretofore underexplored source of variability in the inference of risk tolerance, ambiguity tolerance, and discount factor.

Data Sharing Statement

All data and analysis code will be deposited in a public github repository.

Competing Interest

The authors declare no competing interest.

Funding

References

[1] M.S. Kumar and A. Persaud. Pure contagion and investors' shifting risk appetite: analytical issues and empirical evidence. *International Finance*, 5(3):401–436, 2002.

[2] L.E. Kwako, R. Momenan, R.Z. Litten, G.F. Koob, and D. Goldman. Addictions neuroclinical assessment: a neuroscience-based framework for addictive disorders. *Biological psychiatry*, 80(3):179–189, 2016.

[3] J. Andreoni, M. Callen, K. Hussain, M.Y. Khan, and C. Sprenger. Using preference estimates to customize incentives: an application to polio vaccination drives in pakistan. *Journal of the European Economic Association*, 2016.

[4] K.M. Lempert and E.A. Phelps. The malleability of intertemporal choice. *Trends in cognitive sciences*, 20(1):64–74, 2016.

[5] S.C. Lazzaro, R.B. Rutledge, D.R. Burghart, and P.W. Glimcher. The impact of menstrual cycle phase on economic choice and rationality.

[6] A.B. Konova, S. Lopez-Guzman, A. Urmanche, S. Ross, K. Louie, J. Rotrosen, and P.W. Glimcher. Computational markers of risky decision-making for identification of temporal windows of vulnerability to opioid use in a real-world clinical setting. *JAMA psychiatry*, 77(4):368–377, 2020.

[7] J.A. Clithero. Response times in economics: Looking through the lens of sequential sampling models. *Journal of Economic Psychology*, 69:61–86, 2018.

[8] A. Konovalov and I. Krajbich. Revealed strength of preference: Inference from response times. *Judgment and Decision making*, 14(4), 2019.

[9] R. Webb. The (neural) dynamics of stochastic choice. *Management Science*, 65(1): 230–255, 2019.

[10] J. Peters and M. D'Esposito. The drift diffusion model as the choice rule in intertemporal and risky choice: a case study in medial orbitofrontal cortex lesion patients and controls. *PLoS computational biology*, 16(4):e1007615, 2020.

[11] J. Zimmermann, P.W. Glimcher, and K. Louie. Multiple timescales of normalized value coding underlie adaptive choice behavior. *Nature communications*, 9(1):1–11, 2018.

[12] H.B. Barlow et al. Possible principles underlying the transformation of sensory messages. *Sensory communication*, 1(01), 1961.

[13] J. Fischer and D. Whitney. Serial dependence in visual perception. *Nature neuroscience*, 17(5):738–743, 2014.

[14] A. Kiyonaga, J.M. Scimeca, D.P. Bliss, and D. Whitney. Serial dependence across perception, attention, and memory. *Trends in Cognitive Sciences*, 21(7):493–497, 2017.

[15] A.M. Bornstein, M.W. Khaw, D. Shohamy, and N.D. Daw. Reminders of past choices bias decisions for reward in humans. *Nature Communications*, 8(1):1–9, 2017.

[16] N.V. Banavar, M.D. Lee, and A.M. Bornstein. Sequential effects in non-sequential tasks. In *Proceedings of the 19th International Conference on Cognitive Modeling*.

[17] M. Guan. *A cognitive modeling analysis of risk in sequential choice tasks*. University of California, Irvine, 2019.

[18] L.E. Hunter, A.M. Bornstein, and C.A. Hartley. A common deliberative process underlies model-based planning and patient intertemporal choice. *bioRxiv*, page 499707, 2018.

[19] T.M. Gureckis, J. Martin, J. McDonnell, A.S. Rich, D. Markant, A. Coenen, D. Halpern, J.B. Hamrick, and P. Chan. psiturk: An open-source framework for conducting replicable behavioral experiments online. *Behavior research methods*, 48 (3):829–842, 2016.

[20] M. Plummer et al. Jags: A program for analysis of bayesian graphical models using gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing*, volume 124, pages 1–10. Vienna, Austria., 2003.

[21] M.D. Lee. Bayesian methods in cognitive modeling. *The Stevens' handbook of experimental psychology and cognitive neuroscience*, 5:37–84, 2018.

[22] I. Levy, J. Snell, A.J. Nelson, A. Rustichini, and P.W. Glimcher. Neural representation of subjective value under risk and ambiguity. *Journal of neurophysiology*, 103 (2):1036–1047, 2010.

[23] J. Peters and C. Büchel. The neural mechanisms of inter-temporal decision-making: understanding variability. *Trends in cognitive sciences*, 15(5):227–239, 2011.

[24] M.D. Lee and E-J. Wagenmakers. *Bayesian cognitive modeling: A practical course*. Cambridge university press, 2014.

[25] D. Wabersich and J. Vandekerckhove. Extending jags: A tutorial on adding custom distributions to jags (with a diffusion model example). *Behavior research methods*, 46(1):15–28, 2014.

[26] R. Frey, A. Pedroni, R. Mata, J. Rieskamp, and R. Hertwig. Risk preference shares the psychometric structure of major psychological traits. *Science advances*, 3(10): e1701381, 2017.

[27] I. Krajbich, D. Lu, C. Camerer, and A. Rangel. The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in psychology*, 3:193, 2012.

[28] I. Dasgupta, E. Schulz, N.D. Goodman, and S.J. Gershman. Remembrance of inferences past: Amortization in human hypothesis generation. *Cognition*, 178:67–81, 2018.

[29] N.D. Daw, A.C. Courville, and P. Dayan. Semi-rational models of conditioning: The case of trial order. *The probabilistic mind*, pages 431–452, 2008.

[30] M.W. Khaw, Z. Li, and M. Woodford. Cognitive imprecision and small-stakes risk aversion. *The review of economic studies*, 88(4):1979–2013, 2021.

[31] D. Leiser, O.H. Azar, and L. Hadar. Psychological construal of economic behavior. *Journal of Economic Psychology*, 29(5):762–776, 2008.

[32] I. Krajbich. Accounting for attention in sequential sampling models of decision making. *Current opinion in psychology*, 29:6–11, 2019.

[33] M.W. Khaw, Z. Li, and M. Woodford. Temporal discounting and search habits: evidence for a task-dependent relationship. *Frontiers in Psychology*, 9:2102, 2018.

[34] C.E. Curtis and D. Lee. Beyond working memory: the role of persistent activity in decision making. *Trends in cognitive sciences*, 14(5):216–222, 2010.

[35] T.A. Salthouse. The aging of working memory. *Neuropsychology*, 8(4):535, 1994.

[36] E. Commodari and M. Guarnera. Attention and aging. *Aging clinical and experimental research*, 20(6):578–584, 2008.

[37] N.F. Forbes, L.A. Carrick, A.M. McIntosh, and S.M. Lawrie. Working memory in schizophrenia: a meta-analysis. *Psychological medicine*, 39(6):889–905, 2009.

[38] S.J. Luck and J.M. Gold. The construct of attention in schizophrenia. *Biological psychiatry*, 64(1):34–39, 2008.