Sequential Effects in Non-Sequential Tasks

Nidhi V. Banavar (nbanavar@uci.edu)
Michael D. Lee (mdlee@uci.edu)
Aaron M. Bornstein (aaron.bornstein@uci.edu)
Department of Cognitive Sciences, University of California, Irvine
Irvine, CA 92617 USA

Abstract

In behavioral economic experiments with randomized or unstructured choice sets, trial-level sequential dependencies at the level of choice behavior or reaction time are assumed to be present only in motor or perceptual operations, but not in the cognitive valuation processes themselves. Thus, these are not explicitly accounted for. We present a flexible Bayesian hierarchical model that allows us to test for the presence or absence of linear sequential effects on cognitive, perceptual, and motor parameters of interest and subsequent choice. We apply this model to two data sets: one intertemporal choice and one risky decision making. We demonstrate sequential effects on risk tolerance inference and on the deliberative evaluation of discounted value, with many individual differences. Our results suggest that data collected in sequence cannot be treated as if it were collected independently.

Keywords: sequential decision making; intertemporal choice; risky decision making; hierarchical Bayesian modeling

Introduction

Behaviors like *temporal discounting*, how people discount value over time, and *risk tolerance*, how individuals trade off known uncertainty, are highly context dependent. While they are precisely defined within behavioral economics, or other niche fields, extant literature is rife with evidence suggesting that human behavior does not necessarily correspond to these delineations. Psychologists have demonstrated how inferences of these parameters are sensitive to many other factors including development, arousal, and cognitive capacity (Lempert & Phelps, 2015; Frey et al, 2017). Thus the question of whether these parameters can even be treated as a single (multidimensional) latent variable is a question of active philosophical and empirical research.

Importantly, researchers have also demonstrated that these inferred parameters are sensitive to the method with which they are elicited (Lempert & Phelps, 2015; Frey et al, 2017; Pedroni et al, 2017). In particular, much volatility has been observed both within and across experiments (Frey et al, 2017).

While no measure is "pure," we must examine whether their measurement may be influenced by aspects of task structure that are unrelated, in principle, to the construct under examination. In particular, we focus on the fact that most experiments regarding intertemporal choice (ITC, infer discount factor) and risky decision making (Risk, infer risk tolerance) involve an individual making a sequence of choices, usually in one sitting. Commonly in ITC and Risk tasks, there is no ostensible structure and individuals are explicitly instructed to treat each decision independently and as if it were the only one that counts. Thus typical methods involve treating the data as if it were independently acquired and not actually a sequence of choices.

On the other hand, empirical data and analyses from the working memory and psychophysics literature have for decades demonstrated the effect of *serial dependence*: when stimulus and choice information from previous trials influence current choice behavior and generate systematic patterns in reaction time in the absence of explicit structure in the environment or stimulus sequence (e.g., Lockhead & King, 1983; Bertelson 1961). Further, theories of intertemporal choice that involve prospection—simulating the future (Peters & Buchel, 2010; Gabaix & Laibson 2017)—imply that already computed future values could be cached and re-used, especially if an individual has to make similar choices in sequence (Dasgupta et al, 2018). Studies have shown that episodic cues within an experiment can also influence risky decisions, suggesting a similar reliance on cognitive processes involved in simulation (Ludvig, Madan & Spetch, 2015), which could also lead to re-use.

In this paper, we develop a hierarchical Bayesian model that allows us to test for trial-level sequential influences of stimulus properties. We then apply this model to test for short-term (one-trial-back) influences of cognitive and motor perseveration in both choice behavior and response times.

Methods

Data

Inter-Temporal Choice (ITC) We model $n = 482$ adult subjects (in-person data collection, from Hunter et al, 2018) who made a sequence of 102 binary decisions between same-day monetary reward (SS: smaller sooner, range: \$1–\$85) and a larger reward in the future (LL: larger later, \$10–\$95). Delay (also indicated as T for time) between the SS and LL options ranged between 4 and 180 days. Stimuli were displayed numerically. SS and LL choices were counterbalanced to occur equally often on the Left or Right side of the computer screen. We model both choice behavior and reaction time for this data set. For brevity, however, we present results only for reaction time.

Risk We model $n = 56$ adult subjects (MTurk, from Guan et al, 2020) who made a sequence of 40 binary choices between gambles in the gain and loss domain separately, for a total of 80 trials. Each gamble was associated with two rewards and two probabilities summing to 1 (rewards range: Gain: $\$1 - \$100$, Loss: -$99 – $0, probability range: $1\% - 99\%$). Stimuli were displayed as pie charts with labels indicating

reward amount and probability. We model and present only choice behavior for this data set, as RT was unavailable.

Choice sets were randomized for both experiments, i.e. there were no explicit trial- or task-level sequential dependencies. No outcomes were realized during the tasks (no feedback). In the following sub-section, we develop a model that tests for linear sequential effects of stimulus properties and previous choices on current choice and reaction time.

Cognitive Models: Choice Behavior in Risky Decision Making

For all models, we implement hierarchical Bayesian models in JAGS (Plummer, 2003). Unless otherwise stated, all parameters are hierarchical Normals defined with hyperpriors: $\mu \sim Normal(0,1)$ and $\sigma \sim Normal(0,1)_+$. Thus a hierarchical parameter X is distributed: $X \sim Normal(\mu_X, \sigma_X^2)$. We use hierarchical specifications to better capture individual differences (Lee, 2018).

Subjective Value We model the Subjective Value (SV) of a choice in accordance with Subjective Expected Utility Theory. For individuals $i = 1,\ldots,n$ on trials $j = 1,\ldots,J$ in conditions $c = 1,2$:

$$SV_{(i,j,c)} = \begin{cases} \sum_{m=1}^{2} p_{m(i,j)} \cdot v_{m(i,j)}^{\alpha} & c = 1 \text{ (Gain)} \\ \sum_{m=1}^{2} p_{m(i,j)} \cdot -v_{m(i,j)}^{\alpha} & c = 2 \text{ (Loss).} \end{cases}$$

On a given trial, v is the dollar reward offered for each gamble and p is the probability of reward. As each gamble is associated with two separate rewards, to compute the SV, we multiply each exponentiated reward (v^{α}) and probability, and then sum them. We do this separately for the left and right gamble. The exponent α is interpreted as an individual's risk tolerance (the curvature of the utility function) and is inferred at the individual, not trial, level. Note that v always refers to the objective dollar reward and v^{α} always refers to a subjective dollar reward. We use hyperprior $\mu_{\alpha} \sim Gamma(2,1)$ for risk tolerance, with mode $= 1$ (risk neutrality). We further do not assume the curvature of the utility function is the same in both domains (i.e. infer $\alpha_{(i,c)}$).

Baseline. We implement a logistic choice rule to relate objective trial properties (e.g. dollar reward), subjective trial properties (e.g. SV) and choice behavior. Our baseline model includes no sequential effects. Specifically, the probability of choosing choice A vs choice B, $\theta_{A,B}$, is:

$$\theta_{A,B(i,j,c)} = \frac{1}{1 + \exp\left(\gamma_{(i,c)} + \beta_{(i,c)} \cdot SVD_{(i,j,c)} + \varepsilon_{(i,j,c)}\right)}$$

Here, $SVD_{(i,j,c)}$ represents the difference in $SV_{(i,j,c)}$ between the two options presented on any given trial. Then, $\gamma_{(i,c)}$ represents the shift, or bias, in a decision (towards Left or Right gamble). $\beta_{(i,c)}$ represents response variability, and we use hyperprior $\mu_{\beta} \sim Gamma(2,1)$, where the mode corresponds to probability matching. Finally, $\varepsilon_{(i,j,c)}$ represents effects of simple perseveration (repeat Left or Right choice). All parameters allow for variability at the individual and domain (gain or loss) level. We pair these prior specifications with a *Bernoulli* likelihood, as no two stimuli are presented together more than once.

Sequential Effects: Properties 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: 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 80% 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. We consider the influence of previous (one-trial-back) and current stimulus properties and choices on representation and subsequent decision on the current trial. We compare stimulus properties by taking the absolute difference between given properties on trial j and $j-1$. In particular, we consider the cross-trial differences in the following properties:

Property	ITC	Risk
Value (v)	$v_{LL} - v_{SS}$	$v_R - v_L$
Delay	T	
Entropy (H)		$(H_R + H_L)/2$
Composite		$EV_R - EV_L$
Heuristic (1)		max / min(v or H)
Heuristic (2)	v and T	v and H

Table 1: Stimulus properties considered as indicator variables for the presence of sequential effects.

$H = -\Sigma p \log(p)$ is the Shannon Entropy of a gamble, and $EV = \sum_l p_l \cdot v_l$ is the Expected Value (assuming risk neutrality) of a gamble.

Specifically, we define all these properties as indicator variables (π), using a median split to determine whether the properties being considered in a given model (x) are large or small in difference ($x' = x_j - x_{j-1}$). Then, for every individual i on trial j and condition c:

$$\pi_{High(i,j,c)} = \begin{cases} 1, & |x_{(i,j,c)} - x_{(i,j-1,c)}| > \text{median(all } x') \\ 0, & \text{otherwise,} \end{cases}$$

and vice versa for $\pi_{Low(i,j,c)}$. For example, suppose we were interested in ITC trials with large delay differences (DD). Then, if $T_{(i,j)} = 100$ and $T_{(i,j-1)} = 6$, $\pi_{HighDD(i,j)} = |100 - 6| > 86.5 = 1$.

Sequential Effects: Model We augment our baseline model by allowing the above-mentioned properties to exert linear influences on parameters previously only inferred at the individual level. Each model considers one trial property from Table 1 (under the column Risk) at a time, but tests simultaneously for its influence on the following parameters:

Parameter	
Logistic Bias	$\gamma_{(i,c)}$
Logistic Slope	$\beta_{(i,c)}$
Risk Tolerance	$\alpha_{(i,c)}$
Perseveration	$\varepsilon_{(i,j,c)}$

Table 2: Parameters simultaneously tested for sequential effects in the Risk task.

For example, we use:

$$\alpha'_{(i,j,c)} = \alpha_{(i,c)} + \delta_{(i,c)} \cdot \pi_{(i,j,c)}$$

instead of $\alpha_{(i,c)}$ in our $SV_{(i,j,c)}$ computation, where $\delta_{(i,c)}$ is a continuous variable representing the weight of the sequential effect. By this formulation, δ is actually a $4 \times n$ weighting matrix. Thus, the new $\alpha_{(i,c)}$ is the sequential effect adjusted risk tolerance for individual *i*. Our primary question of interest, then, centers around the posterior values the respective δ parameters take (in particular, zero vs non-zero).

Latent Mixture Finally, we use a latent-mixture model to allow for contaminant behavior. We assume that, for each trial, every individual belongs to one of two groups, or mixtures: task compliant or non-compliant. Specifically, if for any given trial the model infers that $\theta = 0.5$ is more likely (i.e., the subject is guessing) then that trial is considered to be non-compliant and is not included in the regular analysis. We use a $Uniform(0,1)$ prior for the base-rate of each group, paired with a *Bernoulli* likelihood.

Cognitive Models: Reaction Time in Intertemporal Choice

Researchers have used response times (RT) to improve the modeling of discount factors (Peters & D'Esposito, 2020). Previous work has also related components of the Drift Diffusion Model (DDM): both drift rate and bias to discount factor (Hunter et al., 2018). We therefore might expect that sequential effects which do not present themselves in choice outcomes might still be observable in response times.

Thus we implement a modified hierarchical Bayesian approximation of the DDM as presented in Bogacz et al. (2006). The approximation uses a shifted and scaled logistic function (tanh), and we allow for trial level variability in both the bias and drift rate terms. As with choice behavior, unless otherwise specified, all parameters can be assumed to be hierarchical and Normally distributed with independent priors. We use a *Lognormal* likelihood to fit RT.

Baseline. We present three different versions of the Bogacz approximation before considering sequential effects. First, we augment the formulation specified in the original paper with an explicit bias term (1). We assume symmetric thresholds $z'_{(i)}$ and use prior $bias \sim Normal(0,1)T(-z'_{(i)}, z'_{(i)})$, where the bias is restricted to values that fall between $(-z_{(i,j)}, z_{(i,j)})$. A positive bias indicates a preference for LL, while a negative bias for SS. We set $\mu_A \sim Unif(-0.9, 0.9)$ for the drift rate hyperprior.

The other two models decompose the now deterministic drift rate to incorporate stimulus properties explicitly into the model. First, we implement a simple linear regression style on decomposition modeling Subjective Value Difference (2). Here, however, Subjective Value is defined using a non-linear hyperbolic discount function, and $k_{(i)}$ is the discount factor. We also fit a model that does not include an integrated value-delay signal and instead trades off value difference and delay separately as in Hunter et al. (2018) (3). All "regression" weights have $Normal(0,1)$ prior distributions. Then, for threshold $z_{(i,j)}$, drift rate $A_{(i,j)}$ and $c^2 = 1$:

$$\text{LL} : z_{(i,j)} = z'_{(i)} - bias_{(i,j)}, \quad \text{SS} : z_{(i,j)} = z'_{(i)} + bias_{(i,j)}$$

$$DT_{(i,j)} = \frac{z_{(i,j)}}{A_{(i,j)}} \tanh\left(\frac{A_{(i,j)} z_{(i,j)}}{c^2}\right) \quad (1)$$

$$A_{(i,j)} = \beta_{0(i)} + \beta_{1(i)}\left(\frac{v^{\alpha}_{LL(i,j)}}{1 + k_{(i)} T_{(i,j)}} - v^{\alpha}_{SS(i,j)}\right) \quad (2)$$

$$A_{(i,j)} = \beta_{0(i)} + \beta_{1(i)}(v_{LL(i,j)} - v_{SS(i,j)}) + \beta_{2(i)} \log^{-1}(T_{(i,j)}) \quad (3)$$

$$RT \sim \log\text{Normal}(\log(DT_{(i,j)}), \sigma^2_{RT(i)})$$

Sequential Effects As in *Choice Behavior*, we augment our baseline models by allowing the properties listed in Table 1 (under the column ITC) to exert linear influences on the parameters of interest: bias and drift. For example, we use:

$$\beta'_{0(i,j)} = \beta_{0(i)} + \delta_{(i)} \cdot \pi_{(i,j)}$$

for the intercept term in the drift rate decomposition. Here, δ becomes a 3 or $4 \times n$ matrix depending on which model was fit. The sequential effect adjusted term is the newly inferred $\beta_{0(i)}$. Again, our analysis centers around the posterior estimates of δ.

Statistical Analysis

We quantify evidence in favor of either hypothesis by using the Savage-Dickey ratio to approximate the Bayes factor as we test the two hypotheses: H_0: no sequential effect and the alternative H_a: non-zero sequential effects. The Bayes Factor (BF) quantifies the relative strength of evidence in the data: where $BF > 3$ indicates moderate or greater evidence in favor of the hypothesis being considered (Lee & Wagenmakers, 2013). Values lower than 3 indicate that there is not enough evidence in the data to make strong statements in favor of either the null or the alternative. The Savage-Dickey ratio, then, allows us to test nested models at a particular point in the parameter space: namely 0, where there is no sequential effect. In this paper, any "evidence in favor of" a particular hypothesis reported means that the estimated Bayes Factor is greater than 3. In our analyses of hierarchical parameters, we also consider the "representative subject", which is inferred behavior for an individual that contains all the variability of

previous experiment participants. This is distinct from the group mean, and can be thought of as answering the question "what might the next person who walks in to do the experiment look like?"

Results

Choice Behavior in Risky Decision Making

For this task, we consider sequential effects on all parameters listed in Table 2. 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 Table 1: Composite) was followed by a trial with a *low* difference in EV — "easy" then "difficult" in sequence — 4 individuals showed moderate to strong evidence of a negative sequential effect on risk tolerance, but only in the loss domain (see Figure 1). A negative sequential effect implies that the parameter, when inferred without sequential effects, has been underestimated. The true value, then, is greater: for example, for a specific subject, $\alpha = 1.043$ updated to $\alpha = 1.205$ when adjusted for this sensitivity. Importantly, the magnitude of α 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.

Similarly, subjects demonstrated a sensitivity to sequences that were low in entropy in both domains, where the sequentially adjusted logistic slope was higher: reduced response variability than originally inferred.

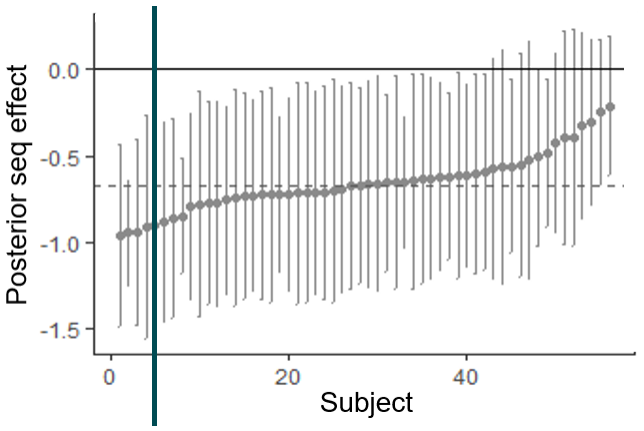

Figure 1: Sequential effects of high EV difference trials followed by low EV difference trials in the Loss domain. The dashed horizontal line is the posterior mean for the representative subject. The bold vertical line separates individuals with $BF > 3$ (left of the line) in favor of sequential effects from those with $1 < BF < 3$ (right of the line). The corresponding plot in the Gain domain not shown as there were no sequential effects in this block of the task.

We also find evidence in favor of the null for an overall effect of motor perseveration: individuals did not systematically "stay" or "switch" their choices in both the gain and loss domain. Interestingly, this is the only result in this data set that holds at the group level and the representative subject level. All other tests for the presence of sequential effects on any parameter involved substantial individual differences: Bayes factors for representative subjects were consistently between 1 and 2, which is interpreted as anecdotal evidence in favor of the hypothesis being tested.

Finally, the model finds little evidence for guessing behavior in the data: a total 12 trials were identified as contaminant by the model (4 gain, 8 loss). This tells us that subjects were largely compliant in the task.

Given the length of the task and the size of this data set, we interpret the few subjects that do show sequential effects for specific trials as a demonstration of the model's capability in identifying individual differences, rather than making more general claims.

Reaction Time in Intertemporal Choice

Out of $n = 482$, subjects, $n = 185$ individuals missed between $1-5$ trials. These missed trials, and the completed trial that immediately followed a missed trial were excluded from the analysis (approximately 1% of total trials).

Baseline. We find that models that fit the deterministic drift rate decomposition and incorporate trial properties, (2) and (3), perform much better ($DIC_1 = 3967717; DIC_2 = 466889.3; DIC_3 = 365194.4$) than the model with purely stochastic drift rate (1). In particular, we find that the aggregate posterior estimates for individuals in Models 2 and 3 are similar for the primary parameters of interest: threshold, bias, and drift rate (see Table 3). Individuals in this data set, on average, appear to have a slight bias towards the Smaller Sooner option (see Table 3).

We further see that all β weights are close to zero, but with considerable individual differences. These low parameter values, however, are to be expected given how small the average inferred drift rate is.

Model 2 tests the hypothesis that as individuals accumulate information, they are considering a unified signal of value and delay which, in this case, is the difference in Subjective Value between the two options presented on the screen. Model 3, on the other hand, tests the hypothesis that individuals separately consider these properties. Our posterior estimates suggest that trading off value and delay independently may be what the subjects are doing, as Model 3 infers a very low number for the value difference parameter. On average, then, individuals are faster to make up their minds the larger the delay between the current and future options. All else held constant, this translates to lower reaction times. As such, and given superior performance in model comparison, we tested for sequential effects using the Model 3 parameterization of drift rate.

Parameter	M1 Mean (95)	M2 Mean (95)	M3 Mean (95)
Threshold	1.98 (1.38,2.63)	1.59 (1.2,2.07)	1.56 (1.17,2.09)
Bias - S	-0.03 (-0.24,0.2)	-0.027 (-0.1,0.06)	-0.01 (-0.09,0.06)
Bias - T	-0.03 (-0.3,0.27)	-0.025 (-0.2,0.14)	-0.01 (-0.18,0.17)
Drift Rate	0.15 (-0.9,1.18)	-0.015 (-0.82,0.86)	-0.009 (-0.89,0.79)
β_0		-0.028 (-0.97,0.90)	-0.015 (-0.95,0.84)
β_1		-0.00245 (-0.45,0.4)	-1.40e-05 (-0.01,0.01)
β_2			0.008 (-0.5,0.56)
Drift Rate	0.015	-0.015	-0.009 (-0.5,0.56)
σ_{RT}	0.14 (0.04,0.3)	0.27 (0.15,0.41)	0.28 (0.14,0.4)

Table 3: Aggregate posterior estimates for DDM parameters. Bias - S is inferred bias at the subject-level, while Bias - T is the subject- and mean trial-level bias.

Sequential Effects. For the four parameters tested for sequential effects (drift rate β weights and trial-level bias), we find that 134 subjects show evidence for *non-zero* sequential effects on at *at least one parameter*. In particular, we present inferences about sequential effects driven by value, delay, or value and delay (see Table 1). This carves the stimulus space into 8 "regions" ($\pi_{(i,j)}$) of sequential effects (See Table 4).

	High Value	High Delay	Low Value	Low Delay
Low Delay	X	-	X	X
Low Value	-	X	X	
High Delay	X	X		
High Value	X			

Table 4: Specific stimulus properties that elicited sequential effects in subjects. An 'X' indicates a trial property or combination we explicitly modeled, and '-' is undefined or a combination that has already been marked.

Of the 134, 69 subjects showed sequential effects on the bias term, 76 on the β_0 drift rate intercept term, 36 on the β_1 drift rate value term and 11 on the β_2 drift rate delay term. We note that 41 subjects have more than one non-zero sequential effect (30 subjects with 2, 10 subjects with 3 and 1 subject with all 4 sequential terms non-zero), and again that this is across all combinations of stimulus properties. Importantly, these sequential effects were distributed roughly evenly between "main" effects driven only by differences in value or delay ($n = 83$) and "interactions" ($n = 89$), with $n = 38$ showing sequential effects for both. That is, unlike the results from modeling choice behavior in the Risk task, DDM parameters seem more susceptible to a broad range of stimulus sequences.

We also found that all 482 subjects showed evidence for *no* sequential effects on *at least one parameter* for some $\pi_{(i,j)}$. This suggests, again, that there are extensive individual differences in both the presence or absence of sequential effects, and in how and when they manifest.

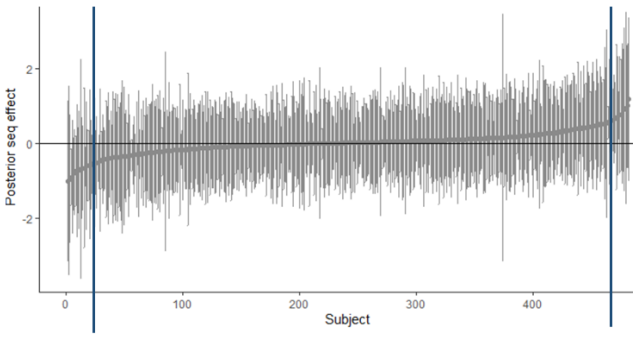

Figure 2: Sequential effects of high delay and low value difference trials on the drift rate intercept parameter. Subjects on the left and right side of the bold vertical have $BF > 3$ in favor of sequential effects. The remainder of parameters and sequential effects are not shown due to space considerations.

In sum, we find that 28% of subjects show evidence of sequential effects in DDM parameters as some function of stimulus properties value and delay. We note that the absence and presence of sequential effects are not the only conclusions we reach from the data: each subject, for some combination of stimulus properties, also had parameters where the strength of evidence was not strong enough to favor either hypothesis.

Discussion

We have introduced a flexible, generative framework to test for the presence of sequential effects on choice behavior and reaction time in explicitly non-sequential, or unstructured, environments. Our model assumes linear influences of current and previous (one-trial-back) stimulus properties on current representations, Drift Diffusion Model (DDM) parameters, and choice. Our results demonstrate evidence of stimulus-driven short-term sequential effects in both choice behavior and on reaction time related parameters in two different economic decision-making tasks. Importantly, these sequential effects were restricted to specific stimulus properties for choice behavior, but were much more widespread for parameters modeling reaction time.

The presence of such effects reinforces the sequential processing nature of the brain and adds to decades long research showing that even if stimuli in an experiment are decorrelated, they are implicitly related by time (Kiyonaga et al., 2017). This corresponds to our results, that parametric inferences even in higher order cognition can be influenced by the linear passage of time and tells us that trials completed in sequence should *not* be treated independently.

Finally, sequential effects in both choice behavior and reaction time showed overwhelming individual differences, with non-trivial changes in parameter magnitude and interpretation. For example, the interpretation of all subjects in the Risk task that presented non-zero sequential effects when "easy" trials preceded "hard" ones changed from risk neutral to risk averse in the loss domain. For the DDM parameters in intertemporal choice, we found similar changes on adjusted bias and drift rate parameters. For example, a subject whose bias term changed from positive to negative was initially interpreted as generally preferring the delayed option (and thus perhaps more patient), when, in actuality, that apparent patience was an artefact of the structure of the choice set. This is particularly important because both the magnitudes of these parameters and their resulting interpretations can be used to explain and predict real world behavior in health and clinical populations (Konova et al, 2020).

Our future directions include expanding the coding of stimulus properties to a continuous kernel: moving beyond indicator variables to continuous parameters and allowing for n-trial-back analyses. We also plan to apply this framework to larger data sets in order to establish the presence or absence of consistent stimulus driven sequential influences across individuals in economic decision making.

Acknowledgements

The authors are grateful to Catherine A. Hartley for providing the intertemporal choice data.

This work used the infrastructure for high-performance and high-throughput computing, research data storage and analysis, and scientific software tool integration built, operated, and updated by the Research Cyberinfrastructure Center (RCIC) at the University of California, Irvine (UCI). The RCIC provides cluster-based systems, application software, and scalable storage to directly support the UCI research community. https://rcic.uci.edu

References

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology*, *13*, 90-102.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*, 700-–765.

Dasgupta, I., Schulz, E., Goodman, N. D., & Gershman, S. J. (2018). Remembrance of inferences past: Amortization in human hypothesis generation. *Cognition*, *178*, 67-81.

Frey, R., Mata, R., Rieskamp, J., & Hertwig, R. a. (2017). Trisk preference shares the psychometric structure of major psychological traits. *Science Advances*, *10*.

Gabaix, X., & Laibson, D. (2017). Myopia and discounting. *National Bureau of Economic Research*.

Guan M., V. J., Stokes R., & Lee, M. (2020). A cognitive modeling analysis of risk in sequential choice tasks. *Judgement and Decision Making*, *15*, 823–850.

Hunter, L. E., Bornstein, A. M., & Hartley, C. A. (2018). A common deliberative process underlies model-based planning and patient intertemporal choice.

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

Konova, A. B., Lopez-Guzman, S., Urmanche, A., Ross, S., Louie, K., Rotrosen, J., & Glimcher, P. W. (2020). 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.

Lee, M. D. (2018). Bayesian methods in cognitive modeling. In J. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience.* John Wiley & Sons.

Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Lempert, K. M., & Phelps, E. A. (2016). The malleability of intertemporal choice. *Trends in Cognitive Sciences*, *20*, 64–74.

Lockhead, G. R., & King, M. C. (1983). A memory model of sequential effects in scaling tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 461-473.

Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2015). Priming memories of past wins induces risk seeking. *Journal of Experimental Psychology: General*, *144*, 24-29.

Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. *Nature Human Behaviour*, *11*, 803-809.

Peters, J., & Büchel, C. (2011). The neural mechanisms of inter-temporal decision-making: Understanding variability. *Trends in Cognitive Sciences*, *15*, 227–239.

Peters, J., & D'Esposito, M. (2020). The drift diffusion model as the choice rule in inter-temporal and risky choice: A case study in medial orbitofrontal cortex lesion patients and controls. *PLOS Computational Biology*, *16*.

Plummer, M. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing.* Vienna, Austria.