The Exemplar Confusion Model: An Account of Biased Probability Estimates in Decisions from Description

Deborah Lin (djlin@student.unimelb.edu.au)
Christopher Donkin (christopher.donkin@gmail.com)
Ben R. Newell (ben.newell@unsw.edu.au)
School of Psychology, UNSW
Australia

Abstract

At the core of every decision-making task are two simple features; outcome values and probabilities. Over the past few decades, many models have developed from von Neumann' and Morgenstern's (1945) Expected Utility Theory to provide a thorough account of people's subjective value and probability weighting functions. In particular, one such model that has been largely successful in both Psychology and Economics is Cumulative Prospect Theory (CPT; Tversky & Kahneman, 1992). While these models do fit people's choice behavior well, few models have attempted to provide a psychological account for subjective value, probability weighting, and resulting choice behavior. In this paper, we focus on a memory confusion process as described in Hawkins et al.'s (2014) exemplar-based model for decisions from experience, the Exemplar Confusion (ExCon) model, and adapt it to account for biased probability estimates in decisions from description. Using Bayesian model selection techniques, we demonstrate that it is able to account for real choice data from a Rieskamp (2008) study using gains, losses, and mixed description-based gambles, and performs at least as well as CPT.

Keywords: Decisions from Description; Exemplar Model; Probability Estimation; Cumulative Prospect Theory; Bayesian Model Selection

Introduction

In a standard 'Decision from description' task, participants are asked to choose between two gambles, A and B. For example, gamble A may involve winning \$4 with probability of 0.80, or receiving nothing otherwise. Gamble B may involve winning \$3 with certainty. Expected utility theory (von Neumann & Morgenstern, 1947) suggests that a decision-maker combines probabilities of outcomes with subjective values to formulate a decision over which gamble they prefer, according to a set of axioms. In this scenario, an expected utility maximizer with a linear utility function would prefer gamble A, since

$$U(W) = cW$$
$$U_A = 0.8\times c\times 4 + 0.2\times c\times 0 = 3.2c \text{ and } U_B = 3c$$
$$U_A > U_B$$

where W is the value of the outcome and c is a constant.

However, violations of the axioms underlying this process have been observed, and thus the notion of a subjective non-linear weighting of probabilities was introduced by Savage (1954) and eventually incorporated into Kahneman and Tversky's (1979) well-known Prospect Theory.

One caveat of these mathematical, expected-utility models is that they are merely descriptive formulations of decision-making, and largely neglect the underlying cognitive processes and reasons for the phenomena (e.g. why people exhibit diminishing sensitivity to increases in value). While it may be argued that people do deliberately participate in complex mathematical processing when making a decision, several studies in which information search is tracked have shown that this is not the likely case even for people of higher cognitive ability (Payne & Braunstein, 1978; Cokely & Kelley, 2009; Glöckner & Herbold, 2011).

Apart from these more descriptive and mathematical models, there are some models that attempt to provide an explanatory, psychological account of decision-making. For example, the Priority Heuristic (Brandstätter, Gigerenzer, & Hertwig, 2006) provides a search-and-stop account of the way in which people evaluate aspects of a gamble, and the Primed Sampler Model (Erev, Glozman, & Hertwig, 2008) proposes that the mere presentation of outcomes and context impacts mental representations of gamble outcomes and probabilities. However, these psychological accounts remain few in number as the traditional focus of decision-making studies has been to develop models which fit more and more data, resulting in models such as Cumulative Prospect Theory (CPT; Tversky & Kahneman, 1992) and, on an extreme end, the Ensemble model (Erev et al., 2010) that tend to have many free parameters and potentially over-fit the data.

A secondary limitation of these mathematical theories lies in the fact that they are unable to account for decisions from experience, where participants sequentially sample outcomes from an unknown distribution to form an estimate of the outcome probabilities in a gamble. This is due to two main reasons. Firstly, the mathematical models developed in the description paradigm tend to utilize the explicit knowledge of the probabilities and outcomes in the gamble, which are not explicitly provided in the experience paradigm. Secondly, the rigidity of the formulae prohibit them from accounting for the inverse probability weighting function observed in decisions from experience (i.e., where people underweight instead of overweight small probabilities), without adding additional parameters and running the risk of over-fitting data.

Thus, instead of extending a mathematical model of decisions from description to account for decisions from experience, we aim to do the opposite with a process model. In this paper, we apply the underlying mechanism of how people form probability estimates of outcomes in an exemplar-based process model for decisions from experience – the Exemplar Confusion Model (ExCon; Hawkins et al., 2014) – to decisions from description. Our choice of model was largely based on convenience, and we view this work as something of a proof-of-concept.

We show that the model can reproduce the complex pattern of over-estimation and under-estimation of probabilities in decisions from description and experience respectively. Finally, we use Bayesian model selection to show that ExCon performs well when compared to a widely accepted benchmark model in the description paradigm – CPT. We begin by defining the models more formally.

Cumulative Prospect Theory

In Cumulative Prospect Theory (Tversky & Kahneman, 1992), people prefer the gamble with the highest weighted utility, *U*, which is calculated using subjective values and probabilities.

$$U(\cdot) = \sum v(x)\pi(p)$$

The value function in CPT, *v(x)*, exhibits diminishing sensitivity to increases in absolute values in the gain and loss domain.

$$v(x) = \begin{cases} -\lambda(-x)^{\beta}, & x < 0 \\ x^{\alpha}, & x \geq 0 \end{cases}$$

The probability weighting function, *π(p)*, is such that small probabilities are over-weighted and large probabilities are under-weighted.

$$\pi(p) = \begin{cases} \dfrac{p^{\delta}}{(p^{\delta} - (1-p)^{\delta})^{\frac{1}{\delta}}}, & x < 0 \\ \dfrac{p^{\gamma}}{(p^{\gamma} - (1-p)^{\gamma})^{\frac{1}{\gamma}}}, & x \geq 0 \end{cases}$$

Finally, the probability that gamble *A* is chosen over gamble *B* is a *softmax* transformation of the utilities of the two gambles

$$P(A,B) = \frac{1}{1 + e^{\varphi[V(A) - V(B)]}}$$

While CPT has generally been hailed to be a successful, benchmark theory in both psychology and economics (Abdellaoui, 2000; Erev et al., 2010; Camerer, 1998), it is not without problems. Aside from the afore-mentioned issue of being a largely descriptive rather than an explanatory model, more in-depth analysis of CPT parameters reveal that its parameter space is not well-constrained. In a parameter recovery study, Nilsson, Rieskamp, and Wagenmakers (2011) found that the effect of loss aversion could be created even without the loss aversion parameter (λ) by allowing the value function parameters (α and β) to take on different values. Similarly, Scheibehenne & Pachur (2014) found that the choice sensitivity parameter (θ) appeared to tradeoff with the value function parameter (α). With such potential trade-offs between parameter values, it is evident that the CPT cannot reliably provide psychological insight into decision-making behavior.

Regardless, the flexibility of the parameter space does allow the CPT to fit data relatively well and it will serve the purpose of model comparison in this paper.

Exemplar Confusion Model

The Exemplar Confusion (ExCon) model developed by Hawkins et al. (2014) provides a process explanation for biased probability estimates in decisions from experience. In the 'Decisions from experience' paradigm, participants are not given a description of the outcomes and probabilities for gambles *A* and *B*, but rather must learn about them through experience. Participants are thus presented with outcomes from gambles, and must infer the probability of those outcomes. Finally, after a suitable number of samples from each gamble have been experienced, the participant chooses between gambles *A* and *B*.

In the ExCon model for experience, people store a memory trace for each outcome they encounter. However, like other models of risky-choice (Erev, Glozman, & Hertwig, 2008), participants are assumed to have an imperfect memory. In ExCon, there is an assumed limit on the accuracy with which memory traces are stored. With probability, p_I, the participant commits a *confusion error*, and fails to store the outcome that should be associated with the current sample, and instead stores something else.

Other models have included process-based mechanisms to produce sub-optimal performance. For example, Bhatia's (2014) added a distraction process to the evidence accumulation process of Decision Field Theory (DFT; Busemeyer & Townsend, 1993). Machiori, Guida and Erev's (2015) Noisy Retrieval Model (NRM) posits that biased probability estimates arise from both a reliance on small samples and confusion with previously encountered outcome distributions in the retrieval process. However, unlike Bhatia's (2014) extension of DFT and the NRM, we posit that the error occurs in memory storage and not in the attentional or retrieval processes, respectively.

ExCon Model for Experience

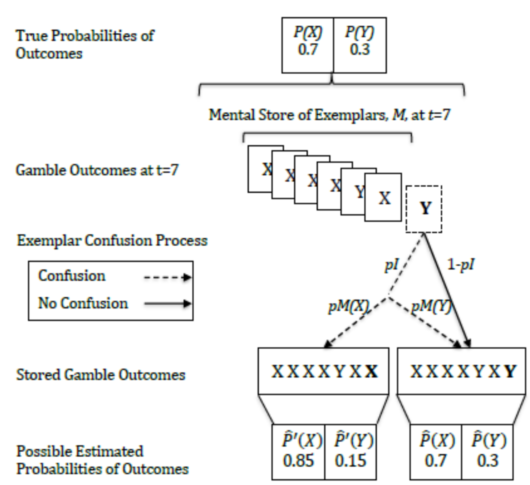

Figure 1. The ExCon Model Process for Decisions from Experience

In decisions from experience, when a confusion error occurs, the new exemplar (i.e. outcome) is confused with a previously stored exemplar. As illustrated by Figure 1, at the 10th draw of an outcome from a gamble, the new exemplar (in this case, the rare outcome 'Y') will be confused with probability, p_I. If there is no confusion, then the outcome will be accurately stored in memory. However, if there is confusion, then one of the previously encountered outcomes is stored instead. The probability that a previously encountered outcome is stored is equal for all previously encountered items in memory. So in the case of the example in Figure 1, the probability that the participant will encode the outcome 'Y' if a confusion occurs, p_M, is 0.5.

ExCon Model for Description

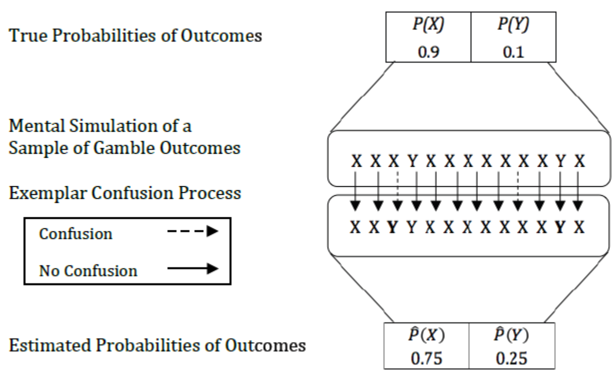

Figure 2. The ExCon model process for Decisions from Description

In this paper, we adapt the ExCon model for the description paradigm. The ExCon for description assumes that people mentally simulate a set of sample gamble outcomes. The number of mental simulations for each gamble, K, varies across individuals. The confusion process occurs with probability p_I. As shown in Figure 2, when a confusion error occurs, the outcome for a given mental simulation is a random choice of one of the possible outcomes.

Note that this is different than in the ExCon for Experience, where outcomes can only be confused with memory traces (i.e., outcomes already sampled). In ExCon for Description, because participants have been presented with all possible outcomes when the gambles are described, we assume that confusions can occur for all possible outcomes regardless of whether they have been previously mentally simulated.

Description-Experience Gap

The "Description-Experience (DE) Gap" is a phenomenon that has been widely observed across decision-making studies, whereby people choose as if they underweight small probabilities in decisions from experience, while choosing as if they overweight them in decisions from description (Hertwig et al., 2004; Hau et al., 2008; Camilleri & Newell, 2013).

The ExCon model for Description naturally predicts an overestimation of small probabilities. However, the ExCon for Experience, as described thus far, also predicts an overestimation of small probabilities (for all but the smallest of probabilities and number of samples). However, it is possible to adapt the ExCon model for Experience to produce an underestimation of small probabilities. The overestimation of small probabilities occurs because confusion errors lead to the sampled outcome being replaced by either outcome with equal probability, once both outcomes have been observed. That is, the confusion error leads to the storage of previously observed outcomes with a probability of 0.5 when there are 2 stored outcomes. As such, ExCon will overestimate the probability of an outcome whenever its true probability is smaller than 0.5.

A simple change to the confusion error process in ExCon for Experience allows the model to underestimate small probabilities. We assume that whenever a confusion error occurs, the participant stores a random sample from the set of outcomes already in memory. So, instead of storing previously observed outcomes with equal probability (e.g., 0.5 when there are two outcomes), the previously observed outcomes are stored with probability equal to the rate at which the outcomes have been observed up until this point. Returning to Figure 1, the probability that the outcome 'Y' is stored is no longer 0.5, but is now 1/6, as the participant had stored 5 'X' outcomes and 1 'Y' outcome.

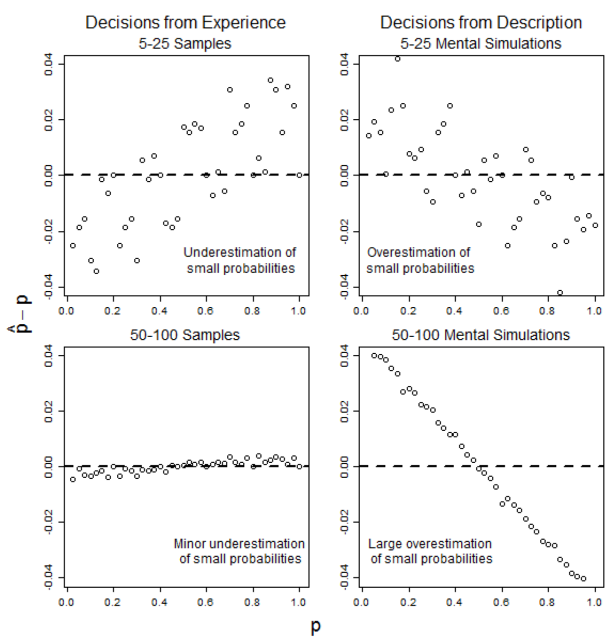

Figure 3. Simulation of the ExCon model produces biased probability estimates in decisions from experience and description. The confusion probability was 0.1 for these simulations.

The left panel of Figure 3 shows that this adapted version of the ExCon for Experience underestimates small probabilities, especially for small samples. In larger samples, however, the probability estimates become less biased, as the memory store of exemplars necessarily have a sample probability, p_M, which is closer to the true probability, on average (Figure 3). In this scenario, even if the exemplar is confused, the probability of a rare outcome being stored as itself is approximately the true probability of the rare outcome occurring, $p_M \approx p$.

On the other hand, the ExCon model for Description generates probability estimates that over-estimate small probabilities. Small probabilities are over-estimated because a confusion error leads to a rare events having an inflated probability of being stored. As a result, the probability estimates for both rare and likely outcomes tend towards 0.5, overestimating small probabilities and underestimating large probabilities (right column of Figure 3). With a larger number of mental simulations, this effect becomes more pronounced because there is a relatively larger proportion of exemplar confusions in the set.

Comparing ExCon and CPT

In addition to producing the characteristic overestimates of low probability events, the ExCon for Description is able to account for real data. To show that ExCon predicts participants' choices, we compare to a well-established benchmark model, the CPT. We use the data from Study 2 in Rieskamp (2008) to compare the ExCon and CPT. Nilsson et al. (2011) showed that a hierarchical extension of CPT was capable of fitting well the data from Rieskamp's (2008) study, thus ensuring that the CPT model provides a good yardstick for the ExCon model. Also, since the gambles in this data set were generated so as to span a wide range of outcomes and probabilities, they should provide a good basis for comparison of the two models. The data set consists of 30 participants, who each contribute 180 pairs of gambles; of which, 60 had only positive outcomes, 60 had only negative outcomes, and 60 had both positive and negative outcomes.

We use Bayes factors to compare the two models. The Bayes factor tells us how much more likely the observed data is under Model A than Model B. Formally, the Bayes factor is the ratio of the marginal likelihood of the observed data, D, for each model, M_i, such that $BF_{AB} = \frac{P(D|M_A)}{P(D|M_B)}$. The marginal likelihood for each model is given by $P(D|M) = \int P(D|\theta, M)P(\theta, M)d\theta$, where $P(D|\theta, M)$ is the likelihood of a set of parameters θ, and $P(\theta, M)$ is the prior probability of those parameters. The marginal likelihood can be interpreted as the likelihood of the model for all parameter values of the model, weighted by the prior probability of those parameter values.

The Bayes factor requires that we specify the prior probability of the parameters of each model. We now define the prior distributions we placed on the parameters of each model. We chose to use moderately informative priors, based on values that are commonly observed in the literature (e.g., Nilsson et al., 2008). Figure 4 contains a plot of the prior distributions we used for each of the model parameters.

The CPT model we fit had 6 parameters. The top row of Figure 4 contains a plot of the prior distributions for the parameters of the CPT model. For the value function parameters, α and β, we used $N(0.5,0.15)$ distributions[1]. For the loss aversion parameter, λ, we used an $F(5,20)$ distribution that was shifted to begin at 0.35 and scaled by 1/2. For the probability weighting function parameters, γ and δ, we used $N(0.55,0.05)$. For the choice sensitivity parameter, φ, we used an $F(3,5)$ distribution shifted to start at 0.05 and scaled by 1/3.

We used an ExCon model with 4 parameters. The model as described thus far has just two parameters – the probability of a confusion, p_I, and the number of mental simulations that the participant undertakes before making their decision, K. For the p_I parameter we used a Beta distribution with shape parameters of 3 and 20. For the K parameter we used an $F(5,20)$ distribution whose output was multiplied by 30. We also found it necessary to include a value function in ExCon, and so assumed that

$$v(x) = \begin{cases} -(-x)^{\alpha}, & x < 0 \\ x^{\alpha}, & x \geq 0 \end{cases}$$

Finally, we assumed that the output of ExCon, the average utility of each gamble based on the mental simulations, was transformed into a choice probability via the same *softmax* function as used in CPT. The prior distributions for the α and ψ parameters are the same in ExCon as they are in CPT.

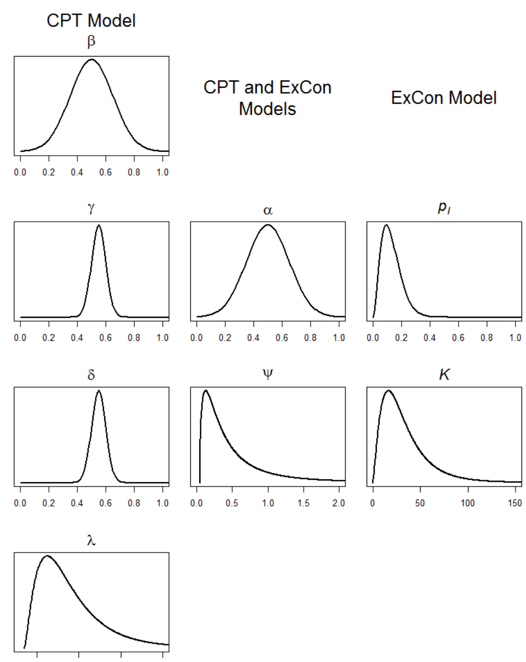

Figure 4. Prior distributions for the CPT and ExCon models. The left column contains the parameters that are only in the CPT model. The center column contains the parameters found in both models. The right column contains the parameters found in only the ExCon model.

[1] Note that parameters with bounds were truncated at those bounds. For example, the α and β parameters were truncated to be between 0 and 1.

We estimate the Bayes factors for each model by evaluating the likelihood of the model across a wide range of parameter values. In particular, for each parameter in each model, we took the 0.05, 0.15,…,0.95 quantiles of the relative prior distribution. We then evaluated the likelihood of each of the 30 participants' data for each model at all combinations of these parameter values. Each of the resultant likelihood values was then multiplied by the prior probability of the parameter values. Finally, the average across all of these weighted likelihoods was taken to give the marginal likelihood for each model. The ratio of these marginal likelihoods gives us the Bayes factor for each of the 30 individuals.

The likelihood of a set of parameters in the CPT model can be calculated analytically. However, in the ExCon model, the probability that a gamble is chosen depends on the particular sequence generated via mental simulation, with stochasticity in both the simulated gambles and the confusion process. We know of no analytical solution for such a doubly-stochastic process. Therefore, we use simulation to generate likelihoods from the model, using 5000 simulations per parameter set.

Figure 5 plots the log Bayes factors for each of the 30 participants from Rieskamp (2008). For 19 of the 30 participants, the ExCon model provides an unequivocally better account of the data than the CPT model. Two participants had Bayes factors for whom the evidence for both models were 'equivalent' (i.e., between 1/3 and 3). The remaining 9 participants were clearly better fit by the CPT model.

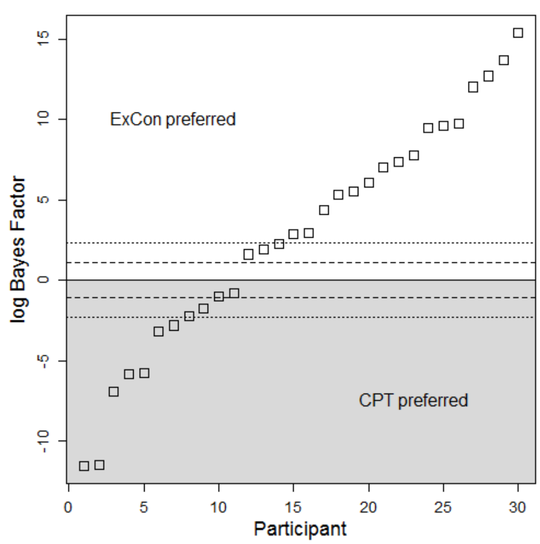

Figure 5. Log Bayes factors for each of the 30 participants from Rieskamp (2008). Values greater than 0 correspond to Bayes factors suggesting that the ExCon model was more likely to have generated the observed data than the CPT model. The dotted lines indicate Bayes factors of 3 and 10 (and $1/3$ and $1/10$).

Discussion

We have shown that a relatively simple model of decision-making, originally developed to account for decisions from experience, is also able to account for data in decisions from description. The model naturally produces an overestimation of small probabilities, and is shown to predict empirical data at least as well as the CPT model – a common benchmark model of decision-making.

Bayesian model selection is dependent on priors. Most concerning is our prior on the number of mental simulations. More work must be done to investigate the nature of the mental simulations, since the number of simulations will depend critically on the mental simulation process.

We must also work to provide further justification for why the confusion processes operate differently under decisions from description and experience. In the description condition, when memory confusion occurs, outcomes are replaced with equal probability. One potential explanation is that observers are influenced by the outcomes that remain onscreen, placing equal attention to both outcomes. In decisions from experience, participants must rely on their memories of sampled outcomes, and so outcomes are replaced based on the contents of memory.

When fitting the ExCon to the decisions from description data, we found it necessary to transform raw outcome values into utilities, and to include a softmax decision rule. Without these extra assumptions, the ExCon model was far too deterministic in its predictions, when compared with the more 'random' behavior of participants.

By including the decision rule and the utility function into ExCon, the model becomes very similar to CPT. The distinguishing feature of ExCon is that it places a process-based account of the under- and over-estimation of probabilities. By constraining the probability-weighting function to follow a particular process, the ExCon model is also constrained in the range of predictions that it makes. Since the observed data are relatively consistent with those predictions, the Bayes factor prefers the ExCon model over the CPT model.

A promising avenue for future work would be to attempt to develop process-based accounts for the value function and decision rule components of CPT (and ExCon). The decision-rule, for example, could be replaced with process models for decisions. Though less explored than the decision rule, it seems possible that the value function may come about through sequential effects. For example, the *assimilation* effect says that stimuli and responses on the current trial are more like those from the previous trial. As such, the discount in utility of rare, large outcomes may come about because their utility is assimilated towards the more common, smaller outcomes.

In decisions from experience, another potential avenue of study would be to develop a psychological account for exploration-exploitation strategies. In a repeated choice paradigm where each choice is consequential, the switch from exploration to exploitation strategies is gradual, while exploitation only occurs with the single choice made only at the end in a sampling paradigm. As each choice is consequential, participants have to decide when to stop exploring the gambles' outcome distributions and start exploiting the preferred gamble for the best total reward. One model that attempts to capture this switch is the Instance-Based Learning model (IBL; Gonzalez & Dutt, 2011), which tracks the rate of alternation between choices instead of the rate of choosing a choice. The IBL also utilizes an inertia parameter at the start of each trial and suggests that the probability of exhibiting inertia in choice, *pInertia*, is selected from a uniform distribution between 0 and 1 each time. However, in reality, it is likely that the inertia increases over time as people become more confident in their preference, thus resulting in a gradual transition from exploration to exploitation in the repeated choice paradigm. Finally, the performance of the ExCon with more specific choice paradoxes and potential manipulation of the confusion process (for example, changing the presentation of outcomes to make confusion less likely) could be explored. If support for methods that lessened memory confusion in decision-making were found it would lend weight to the processing assumptions made in the ExCon framework.

Acknowledgments

BRN was supported by an Australian Research Council Future Fellowship (FT110100151). CD was supported by an Australian Research Council DECRA Fellowship (DE130100129). DL was supported by a UNSW Summer Vacation Research Scholarship.

References

Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. *Management Science, 46(11),* 1497-1512.

Bhatia, S. (2014). Sequential Sampling and Paradoxes of Risky Choice. *Psychonomic Bulletin and Review, 21(5),* 1095-1111.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The Priority Heuristic: Making Choices without Trade-offs. *Psychological Review, 113(2),* 409-432.

Camerer, C. (1998). Bounded Rationality in Individual Decision making. *Experimental Economics, 1*, 163-183.

Camilleri, A.R. & Newell, B.R. (2013).The long and short of it: Closing the description-experience "gap" by taking the long run view. *Cognition, 126,* 54-71.

Cokely, E. T. & Kelley, C. M. (2009). Cognitive Abilities and Superior Decision Making Under Risk: A Protocol Analysis and Process Model Evaluation. *Judgment and Decision making, 4,* 20-33.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., Hertwig, R., Stewart, T., West, R., & Lebiere, C. (2010). A Choice Prediction Competition: Choices from Experience and from Description. *Journal of Behavioral Decision Making, 23*, 15-47.

Erev, I., Glozman, I., & Hertwig, R. (2008). Context, Mere Presentation, and the Impact of Rare events. *Journal of Risk and Uncertainty, 36,* 153-157.

Glöckner, A., & Herbold, A. K. (2011). An Eye-Tracking Study on Information processing in Risky Decisions: Evidence for Compensatory Strategies based on Automatic Processes. *Journal of Behavioral Decision making, 24*, 71-98.

Gonzalez, C., & Dutt, V. (2011). Instance-Based Learning: Integrating Sampling and Repeated Decisions from Experience. *Psychological Review*, *118(4)*, 523-551.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science, 15(8),* 534-539.

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The Description-Experience Gap in Risky Choice: The Role of Sample Size and Experienced Probabilities. *Journal of Behavioral Decision Making, 21*, 493-518.

Hawkins, G. E., Camilleri, A. R., Heathcote, A., Newell, B. R., & Brown, S. D. (2014). Modeling Probability Knowledge and Choice in Decisions from Experience. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 595-600). Austin, TX: Cognitive Science Society.

Kahneman., D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica, 47,* 263-291.

Machiori, D., Guida, S. D., & Erev, I. (in press). Noisy Retrieval Models of Over- and Undersensitivity to Rare Events. *Decision.*.

Nilsson, H., Rieskamp, J., & Wagenmakers, E. (2011). Hierarchical Bayesian Parameter Estimation for Cumulative Prospect Theory. *Journal of Mathematical Psychology, 55*, 84-93.

Payne, J. W., & Braunstein, M. L. (1978). Risky Choice: An Examination of information Acquisition behavior. *Memory and Cognition, 6,* 554-561.

Rieskamp, J. (2008). The Probabilistic Nature of Preferential Choice. *Journal of Experimental Psychology, 34(6)*, 1446-1465.

Savage, L. J. (1954). *The Foundations of Statistics*. New York: Wiley

Scheibehenne, B., & Pachur, T. (2014). Using Bayesian Hierarchical Parameter estimation to Assess the Generalizability of Cognitive Models of Choice. *Psychonomic Bulletin & Review*.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty, 5,* 295-323.

von Neumann, J., & Morgernstern, O. (1947). *Theory of Games and Economic Behavior* (2nd Edition). Princeton: Princeton University Press.