Proceedings of the 30th Annual Conference of the Cognitive Science Society, Washington, DC, 2008

Revealing Individual Differences in the Iowa Gambling Task

Lee I. Newman (Leenewm@Umich.Edu)
Thad A. Polk (Tpolk@Umich.Edu)
Stephanie D. Preston (Prestos@Umich.Edu)
Department of Psychology, University of Michigan, 530 Church Street
Ann Arbor, MI 48109 USA

Abstract

The Iowa Gambling Task (IGT) is a well–studied experimental paradigm known to simulate both intact and impaired real-world decision making in choice tasks that involve uncertain payoffs. Prior work has used computational reinforcement learning models to successfully reproduce a range of task phenomena. In this prior work a set of models were fit to individual decision making data, the best-fit models were selected based on group-averaged metrics, and then theoretical conclusions were drawn based on group-averaged parameters. In the present work we investigate the performance of this class of reinforcement learning models in fitting individual data. This class of learning models has provided a useful starting point for characterizing decision making performance in the aggregate. However, we demonstrate that no one best-fit model aptly captures individual differences and our results caution against using aggregate parameters from best fit models to characterize decision making across populations as has been done in prior work.

Keywords: Reinforcement learning; decision making; Iowa Gambling Task; computational modelling.

Introduction

The Iowa Gambling Task (IGT) is a widely studied decision making task originally designed to simulate the real-world task of deciding among competing options when the payoffs associated with the options are uncertain and must be learned through experience.

Evidence for the ecological validity of the task derives from a large number of studies demonstrating that patients with known impairments in real-world decision making also exhibit impaired performance on the IGT (see Dunn, Dalgleish et al., 2006 for a review). One important population of patients are those with ventromedial prefrontal brain damage.

Interestingly, ventromedial prefrontal cortex is among the brain areas found to be involved in reinforcement learning (RL) in both animals and humans (Daw, 2007; Daw, O'Doherty et al., 2006; Fiorillo, Tobler et al., 2003; O'Doherty, Critchley et al., 2003; Rolls, 1996; Schultz, 2006; Schultz, Dayan et al., 1997; Schultz & Dickinson, 2000). This commonality in the brain areas involved in impaired IGT performance and in reinforcement learning has motivated the application of computational reinforcement learning models to the study of IGT performance. The approach taken in this body of work has been to (i) fit a set of RL models to individual decision data from the IGT, (ii) select a best-fitting model based on a group-averaged fit criterion, and (iii) characterize decision making behavior for a given population, or across populations, using group-averaged model parameters.

This approach has been useful in providing an explicit computational framework for studying IGT performance, and as a means for better understanding the nature of impaired decision making for a wide range of clinical populations. However, several concerns motivate a closer investigation of the performance of these models in capturing decision making at the level of the individual.

One concern is that while the models reported in the literature have been fit to individual data, the model selections and theoretical conclusions were based on group-averaged fits with large reported standard deviations across subjects. This suggests the possibility that models which best-fit a set of subjects may not accurately characterize decision making behavior for a large subset of subjects.

A second concern is that the IGT is a complex task as evidenced by the robust finding that a large number of normal subjects (typically 20-30%) routinely fail to perform the task successfully (for example. Bechara, Damasio et al., 1997). While this finding might be explained by extra-task factors (e.g. motivation, attention) it might instead result from individual differences. The presence of individual differences is further supported by recent empirical evidence demonstrating that subjects make use of explicit decision strategies and that these strategies vary across individuals (Maia & McClelland, 2004).

Taken together, these concerns motivate the present work in which we (i) use a base RL model to replicate the type of group-averaged results reported in the literature and (ii) use a diverse set of model variants to reveal individual differences in decision making and assess the ability of this class of RL model to capture such differences.

Methods

Overview of the IGT

In a widely used version of the IGT (Bechara, Damasio et al., 1994), subjects choose cards from four decks (A, B, C, and D). Every card delivers a positive dollar-denominated gain, but some cards include a monetary loss. Subjects are told in advance that the cards will involve gains, and sometimes losses, and that their goal is to select freely among the decks to maximize their profit over the course of the task which typically involves 100 trials. See (Bechara, Tranel et al., 2000) for a detailed description of the task instructions.

Critically, the payoff schedule (Table 1) is designed to impose a tradeoff in decision making: to perform well subjects must learn that the decks which offer larger immediate gains (A and B) also have larger negative future consequences that make the expected value (EV) of these decks less profitable overall than the decks which offer smaller immediate gains (C and D).

Table 1: Deck Payoff Schedule.

Deck	Deck Type	Fixed Gains	Variable Losses	Loss Freq.	EV per card
A	Bad	+ $100	- $150 to -$350	50%	-$25
B	Bad	+ $100	- $1250	10%	-$25
C	Good	+ $50	- $25 to - $75	50%	+$25
D	Good	+ $50	- $250	10%	+$25

The primary dependent variable upon which IGT performance is assessed is the total percentage of cards selected from the good decks. Healthy subjects who perform advantageously select more than half the cards from the good decks (C, D) over 100 trials.

Data Collection

The behavioral data used in the present work were previously reported by Preston, Buchanan, Stansfield, and Bechara (2007). Forty one university students with no known brain damage or decision making impairments performed the IGT as in the paradigm described above and as reported in (Bechara et al., 1994).

The Base Model

Formally, the IGT is an instrumental conditioning task that can be modeled as a reinforcement learning problem in which subjects must learn, via experienced rewards, the relative values of the four decks in order to bias their selections towards the decks which deliver larger expected values. Action-value models (also known as delta-rule or single-state Q-learning models) are a standard approach to modeling reinforcement learning problems of the form of the IGT and it is this class of model that has been the focus of the research reported in the literature (for example: Busemeyer & Stout, 2002; Yechiam, Busemeyer et al., 2005). Consistent with this prior work, the base model (Model 1) we investigated has three components: (i) a *learning component* responsible for storing and updating internal representations of the value of each deck based upon experienced payoffs, (ii) a *selection component* that chooses among the decks based on their stored value representations, and (iii) a *reward function* that translates experienced monetary payoffs into an internal representation of reward.

Learning Component In learning to choose cards advantageously, the model assumed that subjects internally process an estimated value (V_A, V_B, V_C, V_D) for each deck (A, B, C, D) based on the payoffs experienced from each deck. These values were initialized to zero and updated on each trial according to the following learning equation:

$$V_i(t+1) = (1 - \alpha)\, V_i(t) + \alpha\, r(t) \;:\; i \in \{A, B, C, D\} \quad (1a)$$

In this equation, r(t) is the internal reward experienced by the subject after receiving the monetary payoff from the deck selected on trial t. The learning rate α governs the relative influence of the current value estimate $V_i(t)$ and the current reward r(t) on the updated value $V_i(t+1)$.

Selection Component The base model assumed that on each trial, subjects select a deck probabilistically based on the current learned values associated with each deck. The assumption of probabilistic choice in instrumental conditioning tasks was originally proposed by Herrnstein (1961) and has since being codified mathematically in the well-known *matching rule.* The applicability of the matching rule to human behavior has been empirically demonstrated (O'Doherty et al., 2003). Additional work has clarified the functional form of matching-based selection, and found that a softmax selection rule best fits behavioral choice data (Daw et al., 2006; Lau & Glimcher, 2005). In this work, the probability $P_d(t+1)$ that a subject selects deck *d* on trial *t+1* was computed according to the following softmax rule (Sutton & Barto, 1998):

$$P_d(t+1) = e^{\theta \cdot V_d(t)} \Big/ \sum_{j=1}^{D} e^{\theta \cdot V_j(t)} : \textit{where j indexes the decks} \quad (2)$$

In this action selection equation, the sensitivity parameter θ determined the way in which differences in the deck values V_i are transformed into differences in deck selection probabilities. When θ is zero, selection probabilities are uniform regardless of the value estimates and thus selection is random. Large values of θ lead to large differences in selection probabilities when value estimates differ.

Reward Function The base model dissociates the monetary payoffs (stimuli) obtained during the task from the internal representation of reward *r(t)* experienced by the subject. In the base model, trial rewards were computed according to the equation:

$$r(t) = G \cdot gain(t) + L \cdot loss(t) \quad (3)$$

In the base model, the rewards represent the net monetary payoff obtained on a given trial, with the parameters G and L allowing for both relative weighting of the contribution of gains and losses to the reward as well as absolute weighting of the levels of reward experienced by a subject.

Model Variants

For the purposes of investigating how well this class of model is able to fit individual subject data, we explored ten variants of the base model. The theoretical and empirical motivations and algorithmic instantiations of these models are detailed in Sutton and Barto (1998). These ten models varied along five dimensions: (i) the functional form of the learning rate (exponential vs. simple averaging), (ii) the nature of action selection (softmax vs. pursuit), (iii) the nature of prediction error (delta-rule versus reinforcement comparison), (iv) the presence of decay, and (v) the functional form of the reward function.

Averaging Method in Learning. In the base model the learning rate α was constant, which exponentially weighted past rewards, with more recent rewards having exponentially greater influence on learned values than more remote rewards. In a variant model (Model 2), we allowed α to vary inversely with the number of times a deck has been selected. This resulted in learned values for each deck that were simple arithmetic averages of all rewards experienced.

Pursuit Action Selection In the base model, action values are stored and updated on every trial. Action selection probabilities were computed on-line using these stored values. There is evidence that action selection processes and value learning and storage processes are supported by different neural substrates (O'Doherty et al., 2003). This possibility was captured by a variant model (Model 4) in which an action selection probability P_j was stored for each deck, and these probabilities were updated according to the equations:

$$P_{j^*}(t+1) = P_{j^*}(t) + \beta\,[1 - P_{j^*}(t)\,] \;:\; j^* \text{ is deck with highest value } Vj \qquad (4a)$$

$$P_j(t+1) = P_j(t) + \beta\,[0 - P_j(t)\,] \quad \forall j \neq j^* \qquad (4b)$$

In the Pursuit model, action values were computed and stored as in the base model, but on every trial the probability of the action with the highest value was moved towards 1 (thus the name “pursuit”), and the probabilities all other actions were moved towards 0. The parameter β determined the rate at which stored probabilities were updated.

Reinforcement Comparison In the base model, the learning was based value on the prediction error computed as the difference between the current reward r(t) and the current value estimate $V_i(t)$ (see Equation 1b). Framing effects, the dependence of choice on a context-specific reference point, are well documented in the psychological literature on decision making (for example, Kahneman & Tversky, 1984; Tversky & Kahneman, 1981). We investigated the possibility that deck selections in the IGT might be affected by framing effects. This possibility was instantiated computationally using a Reinforcement Comparison model. In this model (Model 5), we modified Equation (1b) so that prediction error was based on the difference between the current reward *r(t)* and a reference reward $\overline{r}(t)$:

$$V_i(t+1) = V_i(t) + \alpha\,[\,r(t) - \overline{r}(t)\,] \;:\; i \in \{A, B, C, D\} \qquad (5a)$$

$$\overline{r}(t+1) = \overline{r}(t) + \beta\,[\,r(t) - \overline{r}(t)\,] \qquad (5b)$$

The reference reward was updated on every trial, based on the rewards experienced from all decks. The parameter β governed the rate at which reference rewards were updated.

Decay In the base model, the only value updated on each trial was the value of the selected deck. Inherent in this model was the assumption that the previously learned values for all other decks are maintained. The task demands of the IGT are substantial enough that 20-30% of healthy subjects typically fail to perform advantageously. It is therefore plausible that not all subjects were able to maintain the value estimates for decks which have gone unselected for multiple trials. We instantiated this possibility in a variant model (Model 3) in which for every trial that a deck goes unselected, the value estimate for the deck decayed.

$$V_{i^*}(t+1) = V_{i^*}(t) + \alpha\,[1 - V_{i^*}(t)\,] \;:\; i^* \text{ is deck selected on trial } t \qquad (6a)$$

$$V_i(t+1) = \beta\, V_i(t) \quad \forall i \neq i^* \qquad (6b)$$

In this model, value updates for the selected deck (Equation 6a) were identical to the base model (Equation 1a). The value of unselected decks were decayed by the fraction β which we set at 0.8 (i.e., constant decay of 20%).

Form of Reward Function In the base model, the reward function transformed monetary payoffs into rewards assuming that it is the net payoff that is internally rewarding to subjects. Net payoff, however, is not the only possible basis of reward. Electrophysiological animal studies provide evidence that payoffs can be evaluated along multiple dimensions, including frequency and temporal delay in addition to magnitude (Shizgal, 1997). We modeled three alternate reward metrics in which rewards are based on: (i) the variance in net payoffs, (ii) the variance in losses, and (iii) the frequency of losses. In each of these variants (Models 6, 7 and 8), we assumed that lower variance or frequency is more rewarding than higher variance or frequency. We instantiated these models by computing, on every trial, a 4-trial mean variance or frequency.

Hybrid Models We also tested two hybrid models (Models 9 and 10) in which the rewards were a weighted linear combination of net payoffs and either loss variance or loss frequency. We parameterized the contribution of the alternate reward metric thus allowing the contribution of variance or frequency to be fit to individual subjects.

Model Fitting and Comparison

To investigate the ability of this class of RL models to account for decision making of individual subjects, we (i) fit the models to individual decision data using maximum likelihood methods, and (ii) compared the model fits using the Bayesian Information Criteria (BIC). This approach was chosen for consistency with other work reported in the literature (for example: Busemeyer et al., 2002).

Maximum Likelihood Fitting The parameters for each of the models were fit to the deck selection histories independently for each of the 41 subjects using maximum likelihood methods. The likelihood function is given below:

$$ln(L_m) = \sum_{t=1}^{100} ln(P(d = d^* | t, \vec{H})) \quad (7)$$

The likelihood function captured for each of the 100 trials, the estimate by model *m*, of the probability $P(d=d^*|t, \vec{H})$ of choosing the deck d=d* actually selected by the subject on trial t, given the entire selection history $\vec{H}$ up to trial t.

The parameters of each model were numerically fit to the subject data using the Nelder-Mead Simplex numerical optimization algorithm available in the Mathematica programming language (Wolfram Research, 1998-2005).

Model Comparison We used the Bayesian Information Criterion (BIC) which strongly penalizes free parameters relative to other criteria. After performing model fits, we computed a *relative* BIC score for each model *m* compared to a nominal model $m_{nominal}$. The nominal model assumed fixed selection probabilities for each deck, with the probabilities computed based on the proportion of selections from each of the decks. The nominal model thus had three free parameters (the fourth probability can be imputed based on the other three). This model replicates marginal selection probabilities, but does not account for any temporal patterns in the history of selections. We computed a relative BIC score for each model *m* as compared to the nominal model according to Equation (8), in which L_m is the likelihood computed according to Equation (7) and k_m is the number of free parameters in model *m*.

$$\delta BIC_{(m)} = 2\ ln(L_m - L_{nominal}) - (k_m - 3)\ ln(100) \quad (8)$$

The $\delta BIC_{(m)}$ score therefore provides an indication of the goodness of fit for a given model. A positive $\delta BIC_{(m)}$ indicates that a model provides a better fit to the data than the nominal model, a condition which is possible only if the model captures some of the temporal aspects of subject's selection history not captured by the nominal model.

In addition to investigating the goodness of fit for each model relative to the nominal model, we tested for significant differences between δBIC scores of each variant and the base model using the Wilcoxon Signed Rank test from which we generated standard two-sided p-values.

Results

Base Model – Aggregate Results

Before analyzing how well the models fit individual decision data, we first confirmed that the results of fitting our base model to the data were consistent with results reported in the literature for similar subject populations.

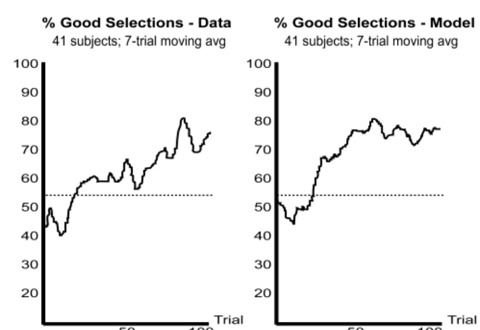
The aggregate results of fitting the base model to each of the 41 subjects are summarized in Table 2. Consistent with simulation results in the literature, and evidenced by the positive mean and median δBIC scores, the model captured temporal aspects of subject selections not captured by the nominal model. The model generated positive δBIC scores for 61% of the 41 subjects. Also consistent with other studies, the standard deviation of the fit across subjects was large (31.6) suggesting that the base model did not provide a good fit for many of the subjects.

Table 2: Base Model Fitting Results

Criteria	Value
δBIC Mean	+13.6
δBIC Median	+3.5
δBIC SD	31.6
% δBIC>0	61%

Figure 1 shows observed and simulated selection histories from the advantageous decks (C and D), pooled over all 41 subjects, and smoothed using a 7 trial window. Consistent with other simulations reported in the literature, the model closely reproduced *aggregate* selection behavior, with early disadvantageous selections shifting to advantageous selections after 20-30 trials, and continuing thereafter and peaking at approximately 75%.

Figure 1: Selection Histories – Model vs. Data.



Model Variants – Aggregate Results

The aggregate results of fitting each of the models pooled across the 41 subjects are summarized in Table 3. Overall, the Hybrid Loss Frequency model, the Base Model, and the Decay model were the best fitting models. These models generated large positive δBIC scores, had positive δBIC scores for a large percentage of the subjects and were the first or second best fits for a large percentage of subjects relative to other models. It is evident from the aggregate results shown in Table 3 that many of the models with inferior performance in aggregate fit were the first- or second- best fitting models for some subjects. For example, the Pursuit and Loss Frequency models (Models 4 and 8) generated low or negative mean δBIC scores, yet together were the best models for a substantial subset (49%) of the subjects.

Table 3: Aggregate Fits for Model Variants

Models	δBIC Mean (p-value [a])	% Subj. δBIC >0 [b]	% Subj. Best Fit [c]
1. Base	+13.6	61%	34%
2. Simple Avg.	+5.1 (<0.003)	51%	24%
3. Decay	+12.8 (<0.38)	56%	22%
4. Pursuit	+4.4 (<0.01)	39%	29%
5. Rein. Comparison	+8.8 (<0.006)	51%	7%
6. Net Payoff Variance	-16.6 (0.000)	17%	7%
7. Loss Variance	-20.3 (<0.000)	20%	2%
8. Loss Frequency	-8.0 (<0.001)	22%	20%
9. Hybrid Loss Variance	+10.4 (<0.004)	49%	10%
10. Hybrid Loss Frequency	+17.9 (<0.22)	63%	44%

[a] p-value for the test of model δBIC versus base model δBIC.
[b] Percent of subjects for which a model had positive δBIC scores.
[c] Percent of subjects for which model was the 1st or 2nd best fit.

Model Variants – Individual Subject Results

To consider more closely how the set of models fit individual subjects, we undertook a subject level analysis the results of which are summarized in Table 4. It is clear from these results that no single model was able to capture the decision making patterns of individual subjects, despite having free parameters that were fit to individual subject data. Across the 41 subjects, nine of the ten models were able to best fit at least one subject. Furthermore, for some subjects (e.g., subjects 2, 5, 6, 8, 13, 21, 23, and 36) no model provided a good fit as evidenced by the negative δBIC score for the models that best-fit these subjects.

Discussion

A growing body of computational work has used a class of RL models to successfully reproduce a range of phenomena observed in the aggregate behavior of subjects performing the IGT. Group-averaged parameters from a best fitting model have been used to characterize the decision making processes thought to underlie task performance for healthy subjects and this approach has been extended to characterizing the impairments observed in a wide range of clinical populations. Inherent in this approach are two important assumptions: (i) that individuals share a common decision making "architecture" captured by the functional form of the model, and (ii) that individual differences in task performance can be captured by differences in the parameter values fit to this common architecture. In the present work, we sought to investigate the soundness of these two assumptions by looking more closely at the performance of this class of RL model in fitting individual decision data.

Table 4: Best Fitting Models for Individual Subjects

Subject*	% Good Selections	Mean δBIC (All Models)	Best Model	δBIC Best Model
15	70	8.8	**1**	17.5
16	73	-0.2	**1**	22
17	56	40	**1**	69.1
20 (F)	46	2.1	**1**	14.6
29	76	19.2	**1**	30.3
41 (F)	27	24.9	**1**	44.1
2	52	-11.1	**2**	-4.7
13	56	-5.2	**2**	-1
18	61	-3.3	**2**	2.8
22	69	-1.7	**2**	5.6
33	67	1.5	**2**	12.6
37	69	1.7	**2**	11.1
8 (F)	23	-14.4	**3**	-5.5
26	68	-0.3	**3**	16.7
30	59	8.5	**3**	13.8
4	76	52.7	**4**	99.1
5 (F)	39	-15.2	**4**	-4.3
9 (F)	18	-9.5	**4**	-0.9
23 (F)	37	-11.8	**4**	-2.6
3	75	-0.5	**5**	36.7
14	83	22.8	**5**	45.6
35	76	20.1	**5**	55.4
7 (F)	32	-8.9	**6**	2
19	58	-2.7	**6**	2
1	55	-4.8	**8**	40.9
11 (F)	46	-4.2	**8**	11
36	52	-10.3	**8**	-2.5
12	65	-6.1	**9**	-2
6	55	-28.5	**10**	-12.1
10	61	-5.8	**10**	46.3
21	69	-26	**10**	-6.2
24	56	-0.6	**10**	7.9
25	84	2.1	**10**	32.8
27 (F)	39	1	**10**	17.5
28	51	4.6	**10**	33
31	78	-22.1	**10**	15.9
32	76	109	**10**	151.4
34	55	2.9	**10**	13.4
38	74	17	**10**	38.2
39	70	30.5	**10**	68
40	71	21.6	**10**	57.8

***** The (F) indicates a subject did not perform the task successfully.

We showed that a model consistent with other models reported in the literature reproduced the *aggregate* behavior of our subjects. However, we found that not one of ten variant models yielded a good fit across all subjects. The model that yielded the best group-averaged fit (Model 10) was the best-fitting model for less than half (44%) of the subjects. Furthermore, many of the models that yielded poor fits across subjects yielded the best fits for sizable subsets. These results were found for subjects who performed the task successfully as well as unsuccessfully and therefore cannot be attributed to the inability of the RL models to fit poorly performing subjects.

What model-based inferences can be drawn by considering the fits at the individual level (Table 4)? First, the fact that the Hybrid Loss Frequency model best fit the most subjects suggests that for these subjects, the avoidance of losses is an important component of reward in addition to the receipt of net monetary payoffs. Moreover, the fact that the two models that consider *only* loss avoidance (Models 7 and 8) provided good fits for 22% of the subjects further supports a role for loss avoidance in decision making. This is consistent with the idea of "loss aversion" that has been well-documented in the decision theory literature (Kahneman & Tversky, 1979). Second, favorable results for the Pursuit model suggest that for some subjects, decisions may be driven in part by positive reinforcement of previously chosen actions, somewhat independent of the trial-by-trial fluctuations in the learned values associated with each deck. Lastly, that many subjects (22%) were well fit by the Decay Model suggests that individual differences in working memory play a role in performance on the IGT. This result is consistent with behavioral (Pecchinenda, Dretsch et al., 2006) and biological data (Lee & Seo, 2007).

Our results suggest that there are significant individual differences in IGT performance and that these differences cannot yet be captured simply by differences in the parameter values of a single best model within the class of RL models that have been applied to the task. Our results also suggest that parametrically characterizing decision processes across subject populations may be premature until more robust models of IGT performance are identified.

References

Bechara, A, Damasio, AR, Damasio, H, & Anderson, SW. (1994). Insensitivity to Future Consequences Following Damage to Human Prefrontal Cortex. *Cognition, 50*(1-3), 7-15.

Bechara, A, Damasio, H, Tranel, D, & Damasio, AR. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293-1295.

Bechara, A, Tranel, D, & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain, 123*, 2189-2202.

Busemeyer, JR, & Stout, JC. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the bechara gambling task. *Psychological assessment, 14*(3), 253-262.

Daw, ND. (2007). Dopamine: at the intersection of reward and action. *Nat Neurosci, 10*(12), 1505-1507.

Daw, ND, O'Doherty, JP, Dayan, P, Seymour, B, et al. (2006). Cortical substrates for exploratory decisions in humans. *Nature, 441*(7095), 876-879.

Dunn, BD, Dalgleish, T, & Lawrence, AD. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience and biobehavioral reviews, 30*(2), 239-271.

Fiorillo, CD, Tobler, PN, & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science, 299*(5614), 1898-1902.

Herrnstein, RJ. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *J. Experimental Analysis Behavior, 4*, 267-272.

Kahneman, D, & Tversky, A. (1979). Prospect theory - analysis of decision under risk. *Econometrica, 47*(2), 263-291.

Kahneman, D, & Tversky, A. (1984). Choices, Values, and Frames. *American Psychologist, 39*(4), 341-350.

Lau, B, & Glimcher, PW. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Experimental Analysis Behavior, 84*(3), 555-579.

Lee, D, & Seo, H. (2007). Mechanisms of Reinforcement Learning and Decision Making in the Primate Dorsolateral Prefrontal Cortex. *Annals of the New York Academy of Sciences, 1104*(1), 108-122.

Maia, TV, & McClelland, JL. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences of the United States of America, 101*(45), 16075-16080.

O'Doherty, J, Critchley, H, Deichmann, R, & Dolan, RJ. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *Journal of neuroscience, 23*(21), 7931-7939.

Pecchinenda, A, Dretsch, M, & Chapman, P. (2006). Working memory involvement in emotion-based processes underlying choosing advantageously. *Experimental Psychology, 53*(3), 191-197.

Preston, SD, Buchanan, TW, Stansfield, RB, & Bechara, A. (2007). Effects of anticipatory stress on decision making in a gambling task. *Behav Neurosci, 121*(2), 257-263.

Rolls, ET. (1996). The orbitofrontal cortex. *Philosophical transactions of the Royal Society of London. Series B: Biological sciences, 351*(1346), 1433-1443; discussion 1443-1434.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Ann Rev Psychol, 57*, 87-115.

Schultz, W, Dayan, P, & Montague, PR. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599.

Schultz, W, & Dickinson, A. (2000). Neuronal coding of prediction errors. *Ann Rev Neurosci, 23*, 473-500.

Shizgal, P. (1997). Neural basis of utility estimation. *Current opinion in neurobiology, 7*(2), 198-208.

Sutton, RS, & Barto, AG. (1998). *Reinforcement learning : an introduction*. Cambridge, Mass. :: MIT Press.

Tversky, A, & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science; Science, 211*(4481), 453-458.

Yechiam, E, Busemeyer, JR, Stout, JC, & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Science, 16*(12), 973-978.

Proceedings of the 30th Annual Conference of the Cognitive Science Society, Washington, DC, 2008