Increased Model Complexity Exhibits Suboptimal Compression of Task Features in Sequential Decision-Making

BBB Honors Thesis

Jonathan A. Levine, '19

Abstract

When faced with complicated decisions, humans use models of varying complexity to integrate information to guide behavior. We describe a general method for using the tools of information theory to quantify model complexity and optimality without needing knowledge of the underlying decision algorithm. This metric provides an interesting perspective on model-based/model-free learning, showing that sometimes more model-free agents may actually be more complex when analyzing the informational content of their models. In several simulated tasks, complex mental models provided more accuracy, but performed less optimal compressions of task information than simpler models, highlighting a tradeoff between cognitive effort and accuracy. This trend was corroborated using human data, where measured model complexity was significantly correlated with decreasing optimality in the Information Bottleneck (IB) space. Finally, we explore the potential for applying these methods to analyze population differences in task performance between Major Depressive Disorder (MDD) patients and healthy controls. The IB method, which highlights the multiplicity of optimal models for a given task, may be a useful tool in studying variability in decision making models across many clinical paradigms.

Introduction

The study of decision-making is inexorably cross-disciplinary – it spans neuroscience, economics, statistics, political science, and computer science. When facing everyday decisions, people use a slew of different strategies that attempt to combine sensory information from the present with their prior life experience in order to choose an appropriate action. Even under similar task conditions, individuals differ dramatically in the types and complexities of the strategies they employ.

In volatile or changing environments, decision-making requires non-trivial combinations of new sensory evidence together with potentially conflicting information from memory or bias. Mathematically, a decision module is performing a probabilistic inference to learn the temporal structure of the evidence generating process (Bishop, 2006), a computation that can be implemented by ensembles of spiking neurons (Legenstien and Maass, 2014). However, as Glaze and coauthors explore in their most recent paper (Glaze et al., 2018), a major problem for this framework is that the complexity of the mental model that might be used to estimate the latent task structure is not well-defined. They show that models of varying complexity in a two-alternative forced choice task (Fig. 1, 2A) exhibit a bias-variance tradeoff, where complexity was defined as the size of the prior hypothesis space over the latent rate of change of the source of evidence. While this bias-variance analysis was customized to the specific parameters of their task paradigm, the notion of using a data-driven technique to measure model complexity can be applied more generally. Here, we use similar tools of information theory introduced by Glaze et al. to explore this model-independent metric for complexity, both in their triangles task (Fig. 1) as well as in the Daw two-step decision task (Fig. 2).

Additionally, we expand the analysis to explore model-independent measures of decision-making optimality. While assessing the optimality of a given decision-making model, many cognitive scientists and psychophysicists have used the fundamentals of Bayesian inference to describe the statistically optimal solution to a problem (Chater and Oaksford, 2008). Using a Bayesian framework, "choice optimality" can be defined with statistical upper bounds on inference problems given a set of parameters, and "choice suboptimality" as departures from the statistically optimal behavior (Wyart and Koechlin, 2016). Yet, this framework places sole focus on reward maximization as the goal of cognitive systems and ignores the complexity of the underlying behavioral models. In reality, decision-making modules might have more diverse goals that include minimizing computation or cognitive effort (Ozcimder et al., 2017; Shenhav et al., 2017; Apps et al., 2015). Accordingly, the previous framework posits that there exists *an* optimal model for a given task, while in reality there might exist *many* optimal models, depending on the subject's desired model complexity. In this paper we argue for a more flexible definition of optimality, which incorporates balancing the minimization of cognitive effort with the maximization of reward. Our definition aims to consider the subject's performance as a function of their level of engagement in the task, and grade their optimality in terms of an Information Bottleneck between cognitive cost and predictive ability.

Methods

Triangles Task

Figure 1 shows the experimental procedure for the triangles task. For this inference task, subjects were required to indicate on each trial which of two sources, corresponding to two triangles laterally flanking the center of the screen, generated a randomly positioned visual star. On each trial, one triangle was chosen as the true source of the generated star, and that source's associated two-dimensional distribution was sampled to determine the position of the star. Each block of trials used a hazard rate (H) that governed the rate of switching between the two sources (triangles) and was chosen randomly from a set of three possible values (0.01, 0.3, 0.99). Hazard rates were chosen without replacement within sessions to ensure a change across blocks.

As described by Glaze and coauthors, subjects differed in their use of the information contained in the hazard rate, showing variability in model complexity akin to cognitive effort. Our data includes a subset of the subjects data reported in Glaze et. al, as well as two simulated subjects using the sampling-based learning model described there.

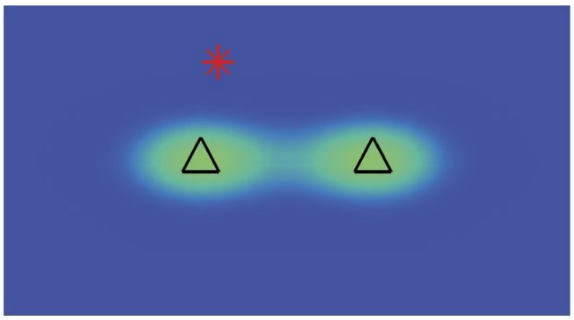


Fig. 1 | The Triangles Task (reprinted from Glaze et al., 2018)
Two Gaussian sources, indicated by green clouds, were centered on the two triangles. For each trial, a random sample from one of the two sources was shown (red star), and then the subject indicated the inferred source of that sample. The state switched at a constant probability on each trial, referred to as the hazard rate, H. The value of H was held constant for blocks of trials but then switched without warning or signal to a new value. For our analysis, there were three hazard rates considered, 0.01, 0.3, and 0.99.

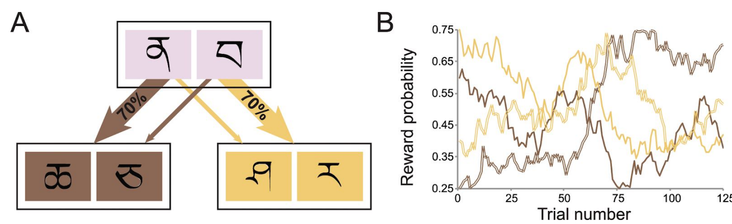


Fig. 2 | The Daw Two-Step Task (reprinted from Kool et. al., 2016)
(A) Participants make a series of choices between two visual stimuli in the first state, each of which leads probabilistically (shown by the arrows) to two possible second stages that offer different probabilities of obtaining reward.
(B) The reward probabilities of the second stages change slowly and independently throughout the task, following Gaussian random walks with reflecting bounds.

Daw Two-Step Task

Figure 2 shows the experimental procedure for the Daw two-step task. As shown by the state-diagram (Fig. 2A), participants make a series of choices between two visual stimuli in the first state, each of which leads probabilistically to one of two possible second stages, each offering different probabilities of obtaining reward. The reward probabilities of the second stages change slowly and independently using a zero-mean Gaussian walk. Our data was obtained using the MATLAB simulations for the *model-based_model-free_stochastic_2_choices* written by

Kool et al., published for public use on Github, found at https://github.com/wkool/tradeoffs. The simulation was run with reflecting reward bounds at 0.25 and 0.75 and a drift rate of 0.025. Each simulated subject's actions were based on a mental model that varied with their model-based/model-free weighting parameter, *w*. A *w* value of 1 is a completely model-based agent, while a value of 0 is entire model-free. In our analysis, we tested the hypothesis that model complexity and model optimality, as defined by the mutual information and Information Bottleneck, are related to this weighting parameter, *w*.

Analysis

All analysis was done in Python, using the Jupyter iPython notebook environment from Anaconda-Navigator. Our code can be found at https://github.com/alsfilip/Empirical_Information_Bottleneck

Predictive information

Predictive information is a data-driven approach in quantifying how much stimulus information a subject was encoding from past task features. This information provides a principled measure of the person's mental model complexity, a good proxy for the cognitive effort a person is putting into the task.

This encoding was done by measuring the mutual information between past stimulus features F_{past} and participant responses R_{now}. For the triangles task, following Glaze et al., we defined task features as combinations of the hazard rate on each trial H_t and stars X_t for any trial t.

To capture the amount of information participants were encoding about the hazard rate, we computed the mutual information (in bits) between some past stimulus feature, and a window of subject responses R_{0-n}, where 0 indicates the current trial, and n indicates some window into the future. n was chosen to be 1 for the triangles task.

I_{past} is calculated as:

$$I(F_{-1}; R_{0-n}) = \sum_f \sum_r p(f,r) \log_2 \left[\frac{p(f,r)}{p(f)p(r)} \right] \tag{1}$$

The extent to which subject responses are predicting the future stimuli (i.e., I_{future}) is measured similarly as

$$I(R_{0-n}; F_{n+1}) = \sum_r \sum_f p(r,f) \log_2 \left[\frac{p(r,f)}{p(r)p(f)} \right] \tag{2}$$

I_{past} is a data-driven quantification of the model complexity, describing how much of the stimulus space is being encoded by the subject. I_{future} then quantifies their use of this information to predict the future.

<u>Information Bottleneck</u>

We represent a mental model M as a compression of task features $F = \{f_1, f_2, ..., f_n\}$. The "optimal" compression of F in this scheme is one which provides the most predictive information about future observations Y. The predictive information contained within a sequence can be measured by the mutual information between past and future observations, or $I_{pred} \equiv I(F; Y)$. Therefore, the optimal compression M is one which minimizes the information retained from the past $I_{past} \equiv I(F; M)$, while maximizing the predictive information about the future $I_{future} \equiv I(M;Y)$. The Information Bottleneck (Tishby, Pereira, Bialek, 2000) provides a method to obtain this optimal compression. It can be obtained by minimizing the following equation:

$$L[p(M|F)] = I(F;M) - \beta I(M;Y) \tag{3}$$

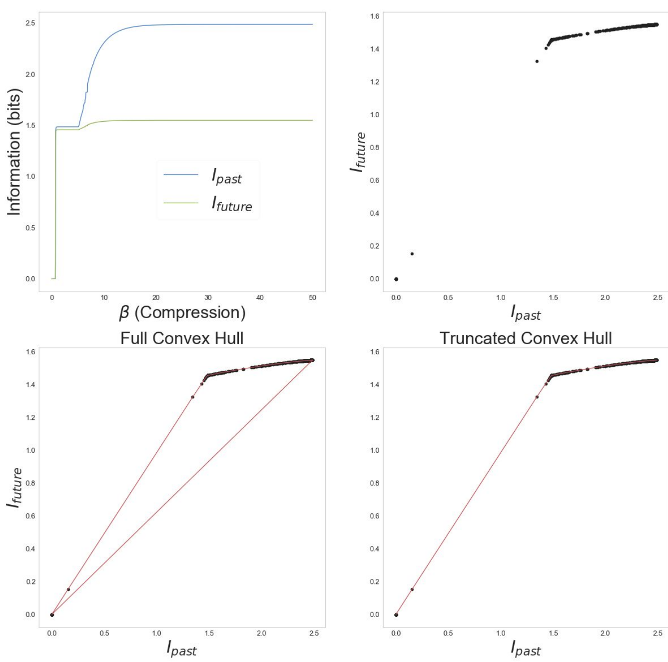


Fig. 3 | The Information Bottleneck

The IB iterates through different values of compression, β, while trying to minimize I_{past} and maximize I_{future} (Top Left). For each point of beta, a point is generated (Top Right) showing the optimal information given some compression of the past. The values generated by the algorithm are then fit to a polygon using a Convex Hull (Bottom Left), which is truncated to create the optimal bound used in our analysis (Bottom Right).

---

where $L[p(M|X)]$ indicates the quantity to be minimized, and β is a Lagrange multiplier. If we solve this equation for a range of values of β, we obtain an Information Bottleneck curve, which indicates the best compression possible for different sizes of I_{past} (Fig. 3). When $\beta \to 0$, none of the past is retained and therefore the compression has no predictive power. This is equivalent to the subject expending zero cognitive effort. Conversely, $\beta \to \infty$ implies that all of the past is being retained (i.e., no compression), and the curve saturates at $I(F;Y)$.

Tishby and colleagues prove that the solution can be obtained using the Blahut-Arimoto algorithm. For a source X, this algorithm converges on an optimal compression $\hat{X}$ that minimizes the source's distortion $d(X,\hat{X})$ using the joint distribution $p(X,\hat{X})$. It is implemented by iterating between these two equations until convergence:

$$p_{t+1}(\hat{X}) = \sum_{x \in X} p_t(x) p_t(\hat{x}|x) \quad (4)$$

$$p_{t+1}(\hat{X}|X) = \frac{p_t(\hat{x}) \exp(-\beta d_t(x,\hat{X}))}{\sum_{\hat{x} \in \hat{X}} p_t(\hat{x}) \exp(-\beta d_t(x,\hat{x}))} \quad (5)$$

where β indicates the targeted slope in the rate distortion curve. Tishby and colleagues demonstrate that the distortion $d(x,\hat{x})$ is equivalent to $D_{KL}[p(Y|X) \| p(Y|\hat{X})]$, where D_{KL} is the Kullback–Leibler divergence. Therefore, in order to obtain the minimum $p(M|F)$ that maximizes $I(M;Y)$ for a given value β, we can iterate through the equations 6-7:

$$p(M|F) = \frac{p(M)}{Z(F,\beta)} \exp\left[-\beta \sum_{y \in Y} p(y|F) \log \frac{p(y|F)}{p(y|F)}\right] \quad (6)$$

(where $Z(X,\beta)$ represents the normalizing factor:

$$Z(X,\beta) = \sum_{m \in M} p(m) \exp\left[\sum_{y \in Y} p(y|f) \log \frac{p(y|f)}{p(y|m)}\right] \;)$$

$$p(Y|M) = \frac{1}{p(m)} \sum_{f \in F} p(y|f) p(m|f) p(f) \quad (7)$$

Iterating through these three equations provides a distribution $p(M|X)$ for a given value of β that minimizes I_{past}, while maximizing I_{future}. The curve is then created using a convex hull fitting algorithm (Fig. 3 Bottom).

Results

The Triangles Task

Figure 4 shows data from two different model simulations (Fig. 4A) doing the triangles task:

1) **Wide Prior:** A more complex model with a wide prior over hazard rate. This model tries to infer the hazard rate in the task using a wide prior over the rate space. This makes it very adaptable to changes in hazard rate, but more variable in its decisions.

2) **Narrow Prior:** A less complex model with narrow prior over the hazard rate space. This model explores a much more restricted range of hazard rates, leading to behavior that essentially amounts to following the visual evidence.

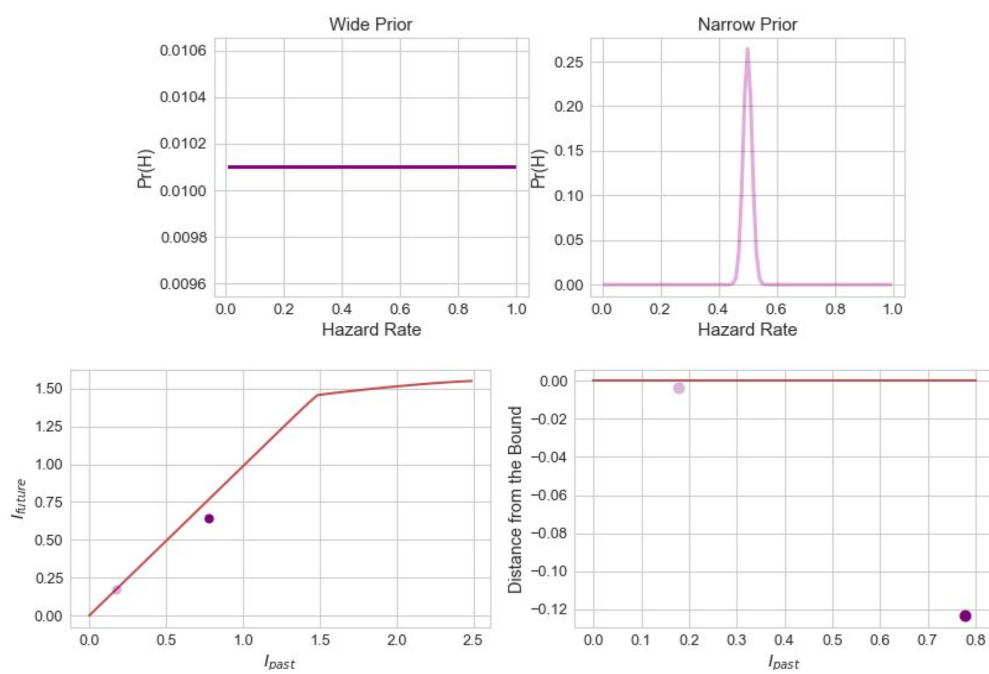


Fig. 4 | The Simulated Triangles Task in the Information Bottleneck
Top: We simulated two different subjects on the triangles task, varying the size of the range of considered hazard rates. Each model tries to infer the hazard rate using the prior probability distributions as shown. (Wide = Dark Purple; Narrow = Light Purple). **Bottom Left:** The predictive information is shown for both subjects, plotted in the IB space. The red curve was generated using a convex hull to enclose the points generated by the Information Bottleneck algorithm written by Tishby et al., as described above. The complexity of the model can be quantified using I_{past}. As expected, the subject with a wider prior over the hazard rates has a higher value for I_{past} and thus uses a more complex model. **Bottom Right:** Optimality as measured by distance from ideal bound. The more complex model lies further from the bound, and is thus a less optimal model.

The model with a narrower prior over hazard rate had a lower value for I_{past} (Fig. 4B), indicating a less complex model. This is consistent with Glaze et al., 2018, where subjects encoding less of the hazard rate information were best fit by a model with a narrow prior. This subject also had a low value for I_{future}, showing that a compression of a limited amount of the feature space can only predict a small amount of the future. When plotting the simulated subjects in the IB space, we found that the more complex model was also less optimal (Fig. 4 C), as it fell further from the ideal bound. This indicates that a more complex model, although more accurate under certain task conditions, is less optimal in its compression of the past into useful information.

To test this relationship between complexity and optimality on the human subjects tested by Glaze et al, we plotted each subject's I_{past} and I_{future} in the IB space, and calculated the distance between that point and the empirical bound for their respective trial sequences. We found that the complexity of mental models for subjects in the triangles task was negatively correlated with optimality (Fig. 5; spearman r = -0.74, p < 0.00001). Thus more complex models, while sometimes more accurate, are less optimal in their ability to efficiently encode information from the past.

<u>The Daw Two-Step Task</u>

For the two-step task, we defined the feature vector, F_{past}, as unique combinations of the following values:

$R1_t$ = subject's first choice (0 or 1) on trial *t*
$R2_t$ = subject's second choice (0 or 1) on trial *t*
Rw_t = reward received (0 or 1) on trial *t*
$S2_t$ = second state left or right (0 or 1) on trial *t*
$R1^*_t$ = first choice with highest probability of receiving a reward on trial *t*

and participant responses, R_{now}, as their first choice (*R1)*. We did not include the second response in the response vector, as the difference between model-based and model-free strategies only carries consequences for the first stage, since the second stage values are identical for both strategies.

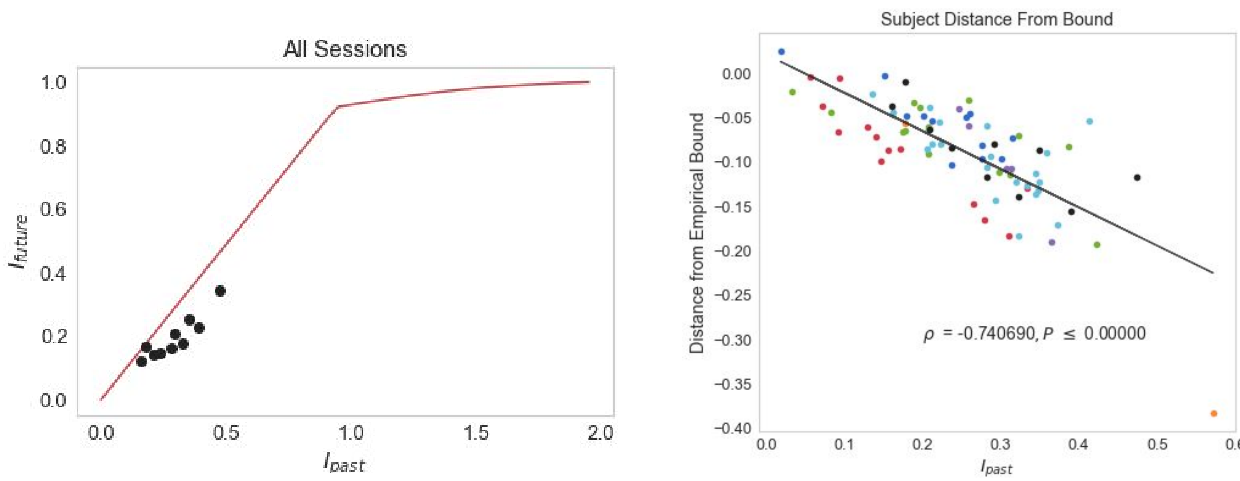
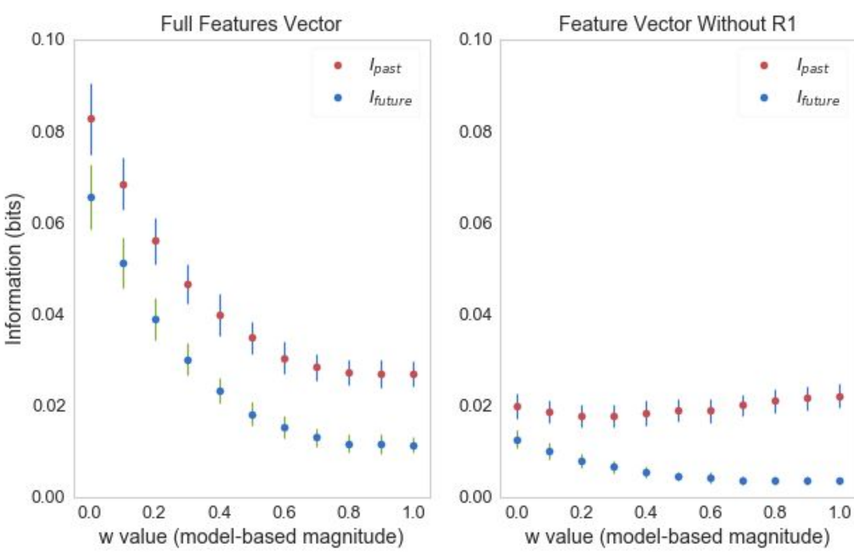


Fig. 5 | Subject Data in the Triangles Task
Increasing model complexity in the triangles task decreases model optimality
Subjects fell along the IB curve depending on the complexity of their strategies (Left; data shown for subset of subjects who completed all sessions). The complexity of the mental model of a given subject was negatively correlated with the optimality of that model (Right; data shown for all subjects in all sessions).

The I_{past} was defined as $I\ (F_{past};R_{now})$, and the I_{future} was defined as $I\ (R_{now};F_{future})$. In this scheme, the *R1** feature acts as the "true" value of the trial, corresponding to the source triangle in the triangles task. In this case, the ideal bound is calculated using $I\ (F_{past}\,;\ M)$ and $I\ (M;\ F_{past})$ for some compression, *M*. The empirical bound is calculated separately for each subject, since each subject's session of trials depend on their responses.

We explored the relationship between the model-based/model-free parameter, *w*, defined in Kool et al., and predictive information. Although it has been reported that the difference between model-based and model-free agents in two-step tasks do not exhibit an accuracy-demand tradeoff (Kool et. al, 2019; Akam et. al, 2015), we test the hypothesis that a more data-driven metric for complexity might highlight differences in the two strategies. Since the model-free strategy is insensitive to the structure of the task, it will simply increase the likelihood of performing an action if it previously led to reward, regardless whether this reward was obtained after a high probability or low probability transition. Conversely, the model-based strategy reflects an interaction between the transition type and reward on the previous trial, and

---



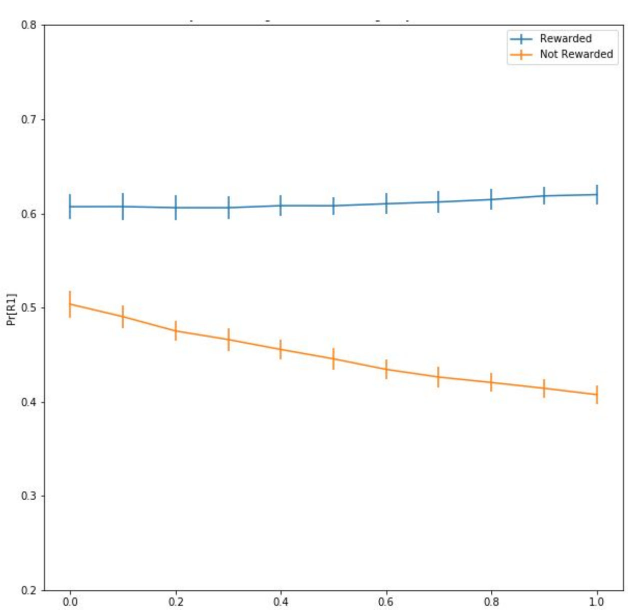


Fig. 6 | Predictive Information in the Daw Two-Step Task

More Model-Free agents (low *w* value) had less complex models due to their differential use of R1 to encode information

Top: Complexity and information encoding decreased as subjects became increasingly model-based (Left). This trend disappeared when R1 was removed from the feature space (Right). The data shown are (mean +/- STD) from 100 simulations of 10,000 trials each, for 11 subjects with varying *w* values.

Bottom: Model-based and model-free agents differed in their "stay" or "switch" probability following a trial in which they were not rewarded. Model-based agents' decisions were based on common/rare transition structure, lending themselves to lower probability/lower variance choices. Model-free agents did not distinguish between common and rare transitions and thus have more complexity due to this variance.

will change its tendency of repeating first-stage actions based on whether the transition was rare or common (Kool et al., 2016).

While the prevailing hypothesis indicates that the model-based strategy is more computationally expensive than the model-free strategy, we found the inverse relationship when defining complexity using the mutual information. In our scheme, the model complexity as measured by I_{past} significantly decreased with increasing w value, indicating that the model-free agents were actually employing a more complex strategy (Fig. 6). We hypothesize that this trend stems from the increased dependency on task features of more model-free agents, since they contain no prior information about the task structure. Because of this difference in a priori knowledge, they need more complex temporal models to integrate noisy information from the reward structure. Thus although the model-based agent might be seen as more complex in terms of its pre-task planning, it actually proves to be less complex in terms of encoding information from the task stimuli.

In order to analyze the optimality of model-based/model-free scheme, we applied the Information Bottleneck method to the two-step task. The complexity and the optimality of the subjects did not show any clear relationship (Fig. 7). While the most complex subject ($w=0.0$) did fall farthest from the bound, there was no systematic optimality curve as a function of complexity. Similarly, the model-free/model-based magnitude parameter w did not show a clear relationship to optimality. Further analysis is required to distinguish between complex models that exhibit the falling-off-the-bound trend and similarly complex models that do not.

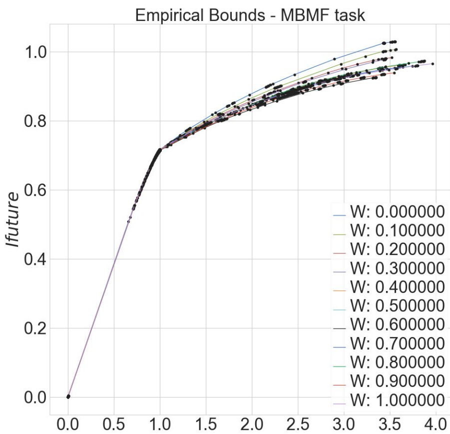
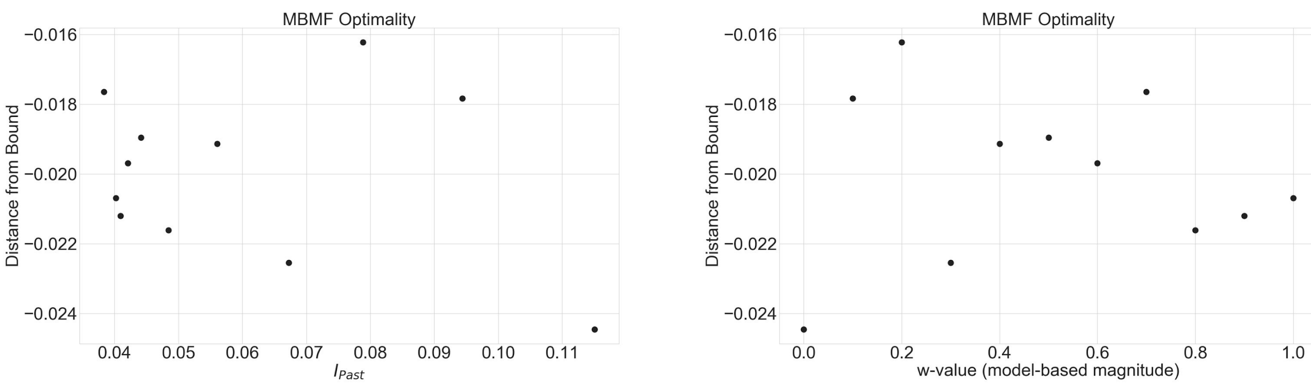


Fig. 7 | The Daw Two-Step Task in the Information Bottleneck

We calculated empirical bounds using the IB method for each simulated subject. The optimal curves changes slightly based on the specific task sequence seen by each simulated subject (Top). We tested the relationship between complexity and optimality: Left: Information-based measures of complexity did not show a clear relationship with optimality. This data set did not exhibit the same trend as the two other data sets in this paper. Right: The model-based magnitude metric, *w*, also did not show any clear, continuous relationship with optimality. The most model-free agent, shown above to be the most complex, did fall farthest from the bound, yet the trend did not extend to mixed strategies.



<u>MDD Reinforcement Learning Task</u>

As reported by Mukherjee et. al., 64 clinically depressed and 64 healthy controls participated in a reward and punishment-based RL probabilistic task. Depressed individuals made significantly poorer choices on both the reward (d = .37, p=.03) and the punishment (d = .58, p = .003) reversal learning task. Depressed participants made significantly less win-stay choices, i.e., making the same choice, given the choice led to a reward or no punishment in the previous trial. Choice behavior in depressed individuals was best fit by model free RL while model based learning best described choice behavior of healthy controls. We attempted to explore these results further using the methods described above. We found that both depressed patients and healthy controls exhibited the complexity-optimality tradeoff in the IB space (Fig. 8 Top). Furthermore, depressed patients had significantly less complex models than their healthy counterparts, and were therefore significantly more optimal as well (Fig. 8 Bottom). While more analysis will be required to understand the clinical relevance of these complexity-optimality population differences, the initial results suggest that the poorer performance by depressed patients on sequential learning tasks may be in part due to their simpler, yet not necessarily less optimal compression of the features in the environment.

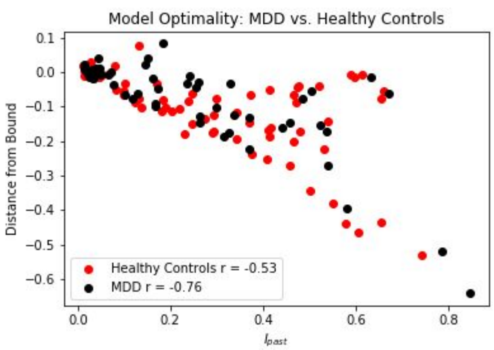
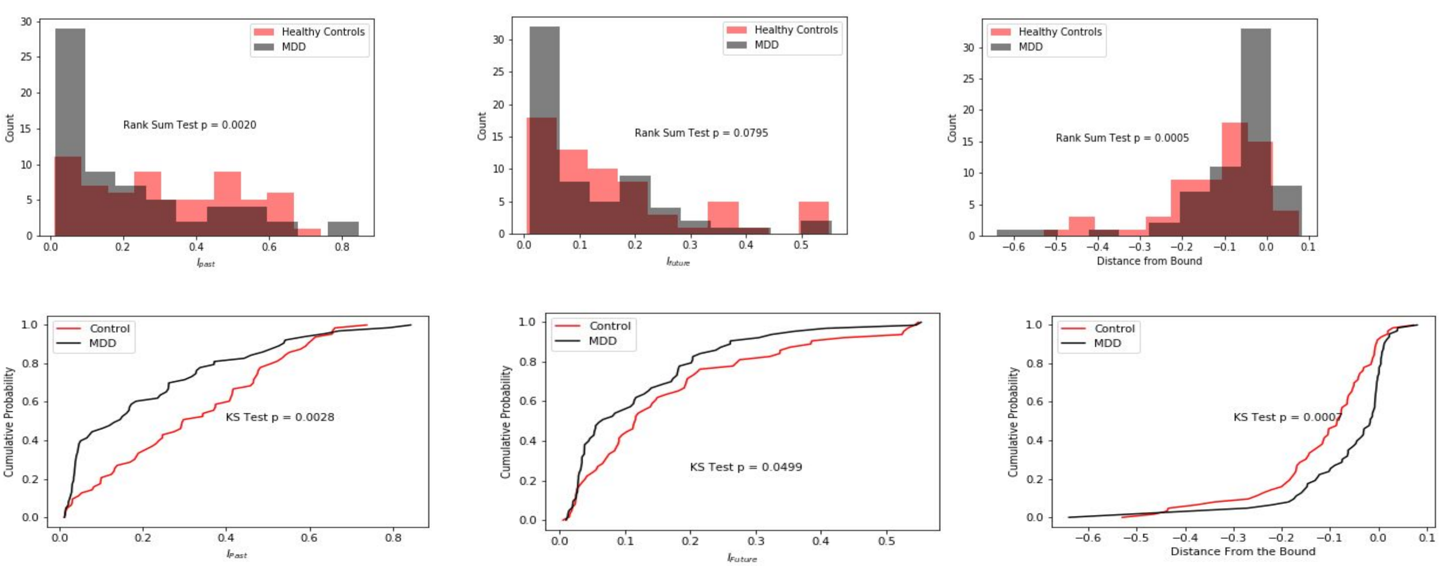


Fig. 8 | Decision Complexity and Optimality in Depressed and Healthy Individuals in an RL task

We calculated predictive information and empirical bounds using the IB method for 64 depressed and 64 healthy individuals in an RL task. We tested the relationship between complexity and optimality: **Top**: Information-based measures of complexity showed a clear inverse relationship with optimality, highlighting a tradeoff between effort and accuracy.. **Bottom**: Histograms and cumulative distributions show that depressed patients were significantly different from healthy patients in both complexity (Left) and optimality (Right).



Discussion

We used a model-independent measure of cognitive effort, defined by the principles of information theory, to analyze mental models of decision-making in the triangles task and the Daw two-step task. We confirmed the findings of Glaze et al. that computational complexity in the triangles task was related to the size of the prior hypothesis space of possible values of the hazard rate. Using an information based metric of model optimality, we demonstrated that more complex models performed a less optimal compression of task information than simpler models, highlighting a tradeoff between cognitive effort and accuracy. This trend was corroborated using human data, where model complexity was significantly correlated with decreasing optimality in the Information Bottleneck space. More complex models can account for more data, but also are more reliant on noisy data sources and thus their compression of the data is less optimal.

To explore this trend further, we examined the Daw two-step task. In recent years, the Daw two-step task has become the method of choice for describing the tradeoff between accuracy and effort in sequential decision-making. Yet, contrary to the prevailing hypotheses in the field, we found that a data-driven definition of model complexity shows the more model-free agents as possessing more complex strategies. When applying the IB optimality metrics to the this data set, no clear trend emerged. While the most complex model, corresponding to the most model-free simulation, fell farthest from the bound, there was not a clean gradient of optimality as a function of complexity. Further analysis of the model-based/model-free strategies in the Information Bottleneck will be crucial in distinguishing between models with complexity that exhibit this optimality tradeoff and those that do not.

These principles are likely to have broad relevance to other tasks that require multiple, sequential inferences. To test the applicability of these methods to clinical populations, we adapted the calculations for a reinforcement learning task that was completed by 64 patients with major depressive disorder and 64 healthy controls. These results highlight that in several different task environments, individual differences in strategy complexity can be measured without knowledge of the underlying model. These differences can then be placed along an Information Bottleneck to evaluate their optimality. These metrics of complexity and optimality are generalizable, and can be compared across models and across task structure. Future work should seek to standardize the window sizes and dimensionality of the feature spaces considered in the information calculations, as well as work to optimize the algorithms for high dimensional data. A larger study with more diverse environments might also prove helpful in evaluating the usefulness of a complexity vs. optimality tradeoff more generally.

Contributions

J.A.L. and A.L.S.F. designed research; J.A.L. and A.L.S.F. analyzed data; J.A.L. wrote the paper.

References

Akam, T., Costa, R., & Dayan, P. (2015). Simple plans or sophisticated habits? State, transition and learning interactions in the two-step task. *PLoS computational biology*, *11*(12), e1004648.

Apps, M. A., Grima, L. L., Manohar, S., & Husain, M. (2015). The role of cognitive effort in subjective reward devaluation and risky decision-making. *Scientific reports*, *5*, 16880.

Bishop, C. M. (2006). Pattern recognition and machine learning (information science and statistics) springer-verlag new york. *Inc. Secaucus, NJ, USA*.

Chater, N., & Oaksford, M. (Eds.). (2008). *The probabilistic mind: Prospects for Bayesian cognitive science*. OUP Oxford.

Glaze, C. M., Filipowicz, A. L., Kable, J. W., Balasubramanian, V., & Gold, J. I. (2018). A bias–variance trade-off governs individual differences in on-line learning in an unpredictable environment. *Nature Human Behaviour*, *2*(3), 213.

Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When does model-based control pay off?. *PLoS computational biology*, *12*(8), e1005090.

Legenstein, R., & Maass, W. (2014). Ensembles of spiking neurons with noise support optimal probabilistic inference in a dynamically changing environment. *PLoS computational biology*, *10*(10), e1003859.

Mukherjee et. al. (Submitted for publication, 2019). Reward? What Reward? Probabilistic reinforcement reversal learning in Major Depressive Disorder.

Ozcimder, K., Dey, B., Musslick, S., Petri, G., Ahmed, N. K., Willke, T. L., & Cohen, J. D. (2017). A Formal Approach to Modeling the Cost of Cognitive Control. *arXiv preprint arXiv:1706.00085*.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual review of neuroscience*, *40*, 99-124.

Wyart, V., & Koechlin, E. (2016). Choice variability and suboptimality in uncertain environments. *Current Opinion in Behavioral Sciences*, *11*, 109-115.