

A Go at Classic Cognitive Modeling in a Two-Alternative Forced-Choice Bandit Task: Examining Task Influences on Human and Model Performance

Pulkit Goyal, Joel Vasama

October 4, 2022

1 Abstract

We use two classic learning models from cognitive science to model learning in 19 participants who completed a two-alternative forced-choice task. The task consisted of two aliens (alternatives) and participants were instructed to select which option would win based on 2-4 features (n of features condition). Additionally, the task was completed under a time pressure condition with a set time limit of 3 or 6s per trial. In our behavioral analyses we find indication of learning with improvement in performance across trials. We then continue to fit the parameters of our respective models including an ϵ value to capture noisiness in decision making. Here we find that Rescorla-Wagner fits the human data the best while Kalman filter provides a limited fit. After analyzing parameters we conclude that time pressure influences noisiness in decision making but does not influence learning rate. Similarly, we notice that the number of features influences decision noise but not learning rate.

2 Introduction

Learning and decision making under uncertainty requires a good balance between exploration and exploitation. Additionally, other factors like the complexity of the task, the degrees of freedom, and the time pressure can highly influence learning and decision making.

Here, we investigate the effects of such complexities on the learning and decision making.

Existing literature suggests that with increased complexity the agent prefers exploitation over exploration. We expect to see this as lower decision noise and learning rate under the effects of time constraints, and the number of the features.

H_1 - Increasing number of features increases decision noise and decreases learning rate.

H_2 - Time pressure decreases decision noise and learning rate.

3 Methods

3.1 Participant and Data

Anonymous data collected from 19 participants was provided without demographic variables other than source (Computational Cognitive Science class at University of Tübingen). Data from 1 of the 19 participants, participant number 18, was excluded from our analyses after being identified

as an outlier, performing significantly better than the rest of the sample with both notably higher accuracy ($M = 89\%$ vs. 63.3%) and faster reaction time ($M = 675.16$ ms vs. 2131.19 ms). The reaction time of participant number 11 was also identified as an outlier ($M = 3072.21$ ms vs. 2131.19 ms) but as their accuracy was normal we decided to keep this participant (see fig. 1).

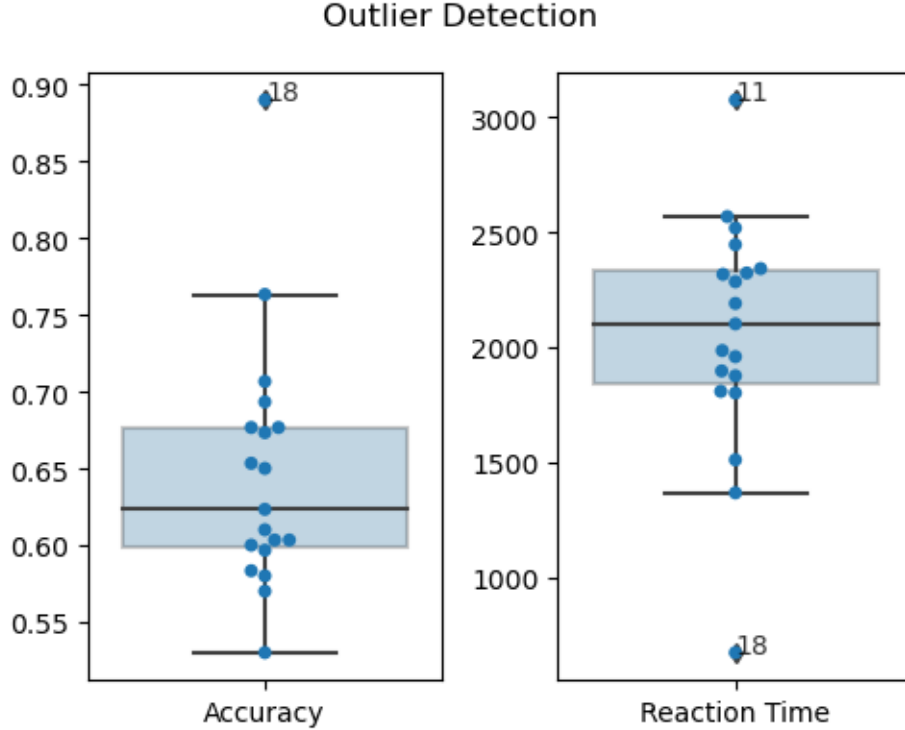


Figure 1

3.2 Task

A two-alternative forced choice task was used to investigate learning in the participants. Participants went through 30 blocks of 10 trials. In each trial the goal was to determine which of two aliens win an abstract, non-visualized game based on 2, 3 or 4 features (number of features condition) provided to the participant under either a 3s (fast) or 6s (slow) time limit (time-pressure condition). The first 15 blocks were either in the slow or fast time pressure condition while the last 15 would be in the other, with this order of time-pressure conditions being counter-balanced between subjects. The number of features for each block were set constant but randomized within these 15 set blocks resulting in 5 sets of each 2, 3, or 4 features per time-pressure condition. Each of the features of each alien (2 to 4D vectors $x_{t,F}$ and $x_{t,J}$) were generated using a Gaussian distribution with a mean (M) of 0 and a standard deviation (sd) of 1. These were further weighted w using a Gaussian distribution of $M = 0$ and $sd = 1$, defining which of the aliens won based on $y_{t,F} = w^\top x_{t,F}$ and $y_{t,J} = w^\top x_{t,J}$.

3.3 Modeling

3.3.1 Learning Models

We compared two classic learning models of cognition against a random policy to model learning in participants.

1. Rescorla-Wagner
2. Kalman Filter

3.3.2 Decision-Making

For more accurate likelihood calculations, we wanted a decision algorithm that outputs a non-binary choice probability and wanted to be learning model agnostic, so we chose Softmax. Additionally, to model errors, we added an ϵ -greedy error rule on top of the decision making process. The epsilon value was added as a search parameter during parameter fitting/optimisation.

3.3.3 Scoring Criterion

We choose to use the Bayesian Information Criterion (BIC) as our performance measure as our goal is to find the best fitting model or maximise the posterior probability.

3.3.4 Model Fitting

We used grid search with sufficient precision for optimisation because of its ease of implementation, availability in standard Python libraries, and potential for caching (leading to faster computations when re-running analysis on different subsets of tasks).

For accuracy validation, we compared it with other more accurate optimization techniques. The comparisons resulted in minimal differences in BIC scores to the Grid Search optimization. Moreover, the run-times were comparable.

4 Results

4.1 Behavioral

A one-sample t-test against a 50% chance level (participants had two alternatives) confirmed that participants performed on average above chance level ($t(17) = 9.75$, $p < 0.05$) with an total average accuracy of 63.3% across all participants.

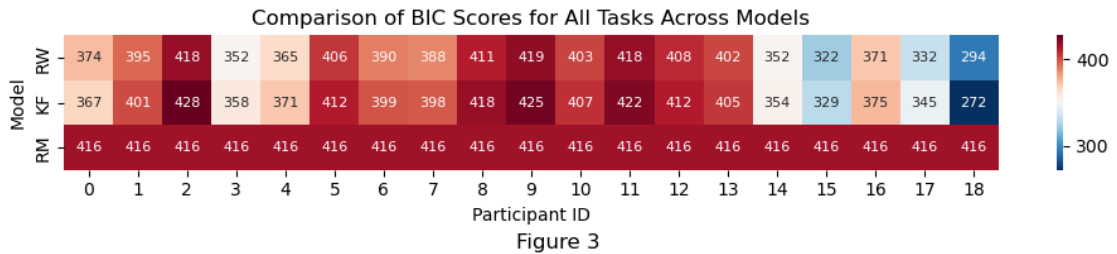
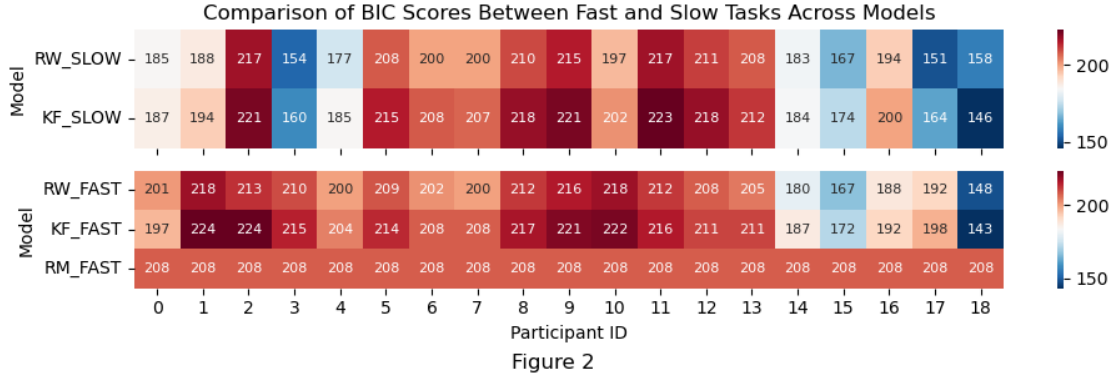
We then proceeded to complete five linear regression analyses to examine performance in trials and tasks, the influence of the number of features on performance, and the relationship between reaction time and performance. Regressing performance on trial within block, the first analysis indicated that participants perform better with each trial within blocks ($\beta = 0.014$, $R^2 = 0.007$, $p < 0.001$). On average, participants perform 1.4% better each trial after the first, suggesting participants indeed learn the task. In our second analysis we regressed performance on block, finding statistically significant relationship ($p = 0.329$) suggesting participant performance did not vary within blocks. However, including the starting condition as a covariate resulted in a significant relationship ($\beta_{block} = 0.0007$, $\beta_{start} = 0.0475$, $R^2 = 0.002$, $p = 0.003$), though only finding the starting condition as a significant regressand ($p_{block} = 0.329$, $p_{start} = 0.001$). Our third analysis regressing performance on number of features found a statistically significant relationship ($\beta = -0.043$, $R^2 = 0.005$, $p < 0.001$), indicating that the addition of each feature decreases accuracy on

average by 4.3%. In other words, this suggests learning became more difficult with the addition of each feature. Finally, in the fourth analysis regressing performance on reaction time we found a that higher reaction time generally resulted in better accuracy ($\beta = 0.00005$, $R = 0.011$, $p < 0.001$) with a 0.5% increase in accuracy for every 100ms. However, an additional Pearson's correlation test suggested no statistically significant relationship ($r = -0.43$, $p = 0.07$).

We also performed two additional t-tests to examine the differences in performance due to time-pressure condition and starting time-pressure condition. Using a related samples t-test we found that participant perform of average 4.5% better ($M_{fast} = 61\%$ vs. $M_{slow} = 65.5\%$) in slow trials ($t(17) = -3.5$, $p < 0.001$), possibly due to having the ability to better consider the features. In contrast, an independent samples t-test showed no differences in participant performance resulting from first encountered the slow condition instead of the fast condition ($t(17) = -1.69$, $p = 0.121$). We also tested for a difference in reaction time between time-pressure conditions, and found that participants took almost 1s less to respond in the fast vs. slow condition ($t(17) = -37.087$, $p < 0.001$, $M_{fast} = 1630.56$ ms vs. $M_{slow} = 2631.83$ ms).

4.2 Model Comparison

In the heatmaps below (fig. 2 and 3), we include participant 18 due to the interesting performance they exhibited.



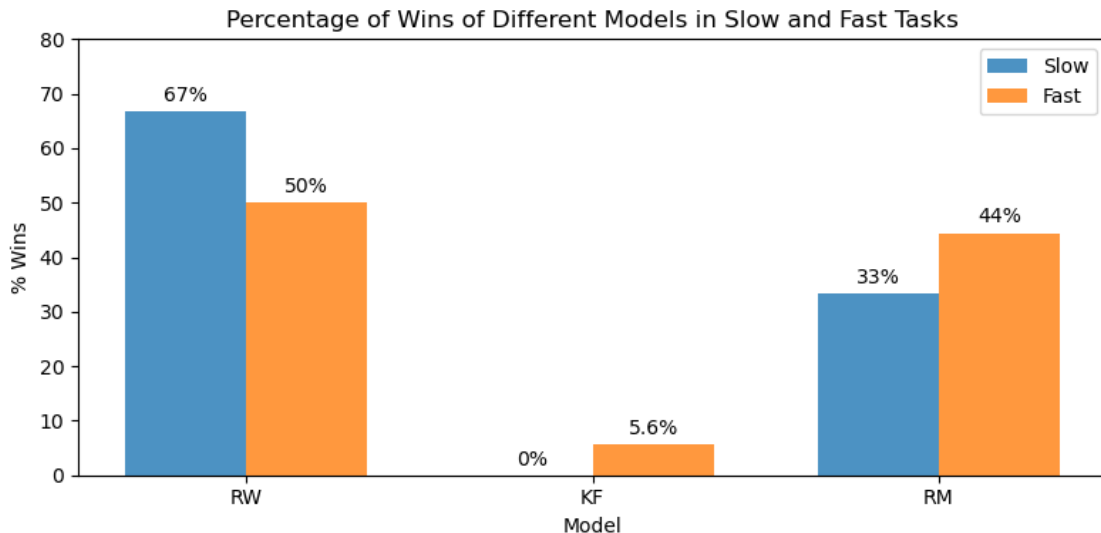


Figure 4

Comparison of the BIC scores of the best performing models (fig. 4) reveal that Rescorla Wagner model explains the human data best, with and without time pressure (~67% without time pressure and ~50% with time pressure), followed by the random model (~33% without time pressure and ~44% with time pressure), and finally by the Kalman Filter (~0% without time pressure and ~6% with time pressure).

However, for participant 18 in particular, the Kalman Filter model atypically performs best. This is expected because this participant was an ideal participant (a bot) that indeed employed Kalman Filter to learn, but the BIC score is not low enough (for an ideal learner) because there was some artificial randomness added for realism, and thus the ϵ error value for the best fitting Kalman Filter model for participant 18 is also not zero.

The parameter recovery resulted in almost the same parameters ($\sigma_y = 0.1$, $\sigma_w = 0.9$) of the Kalman Filter model that the bot employed ($\sigma_y = 0.1$, $\sigma_w = 1$). (σ_w is not the same because $\sigma_w > 0.9$ was not included in the search space $[0, 0.9]$).

Based on these results we conclude that participants tend use a simpler learning strategy, such as those formulated in the Rescorla-Wagner and Random Policy models, in general and especially under time-pressure.

4.3 Parameter Analysis

4.3.1 Rescorla-Wagner

We continue to analyze the fitted parameters of our models to gain further insight into the behaviour of the models.

We first explored the influence of the time pressure condition on the learning rates of the Rescorla-Wagner model, finding no statistically significant difference between the slow or fast conditions (paired-samples, $t(17) = -0.176$, $p = 0.862$). Considering the learning rate of Rescorla-Wagner to

represent participant learning, this suggests that participants did not learn slower (or faster) under or without time pressure.

Similarly, we note no difference in the fitted epsilon values between the time pressure conditions (paired-samples, $t(17) = -1.350$, $p = 0.195$), suggesting that time pressure does not lead to noisier decisions. But if the number of features are two, we found that time pressure significantly affects the noisiness of decisions (paired-samples, $t(17) = -2.667$, $p = 0.016$).

We also tested whether the number of features lead to differences in learning or noisiness of decisions. Using linear mixed effects models, regressing learning rate on number of features we find no statistically significant relationship for slow ($p = 0.073$) or fast ($p = 0.632$) tasks, suggesting that the number of features, like time pressure, does not influence the rate of (especially fast) learning.

However, we did find a positive relationship between number of features and epsilon ($\beta = 0.119$, Marginal $R^2 = 0.146$, $p = 0.005$), suggesting that increasing the number of features led to noisier decisions, possibly indicating more exploration. After performing a further Tukey's HSD test, we note that there are no differences in epsilon values between three and four features ($M_{diff} = -0.0184$, $p = 0.9555$), indicating that any features beyond two may not influence decision errors (see fig. 5).

Also, although not statistically significant, increasing features might also lead to noisier decisions in fast tasks ($\beta = 0.086$, Marginal $R^2 = 0.075$, $p = 0.057$) (see fig. 6).



Figure 5

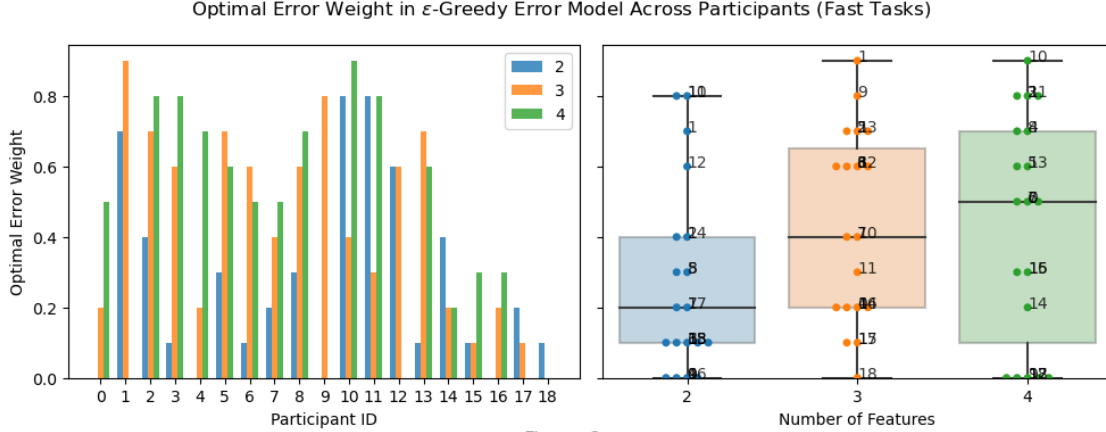


Figure 6

4.3.2 Kalman Filter

We completed similar analyses for the parameters obtained from the Kalman filter as we did for the Rescorla-Wagner model.

We used σ_y (score confidence) to represent learning and tested whether it would be influenced by the time pressure conditions. Similar to the Rescorla-Wagner model analysis, a paired-samples t-test revealed no statistically significant differences in the mean value of σ_y between normal and slow conditions ($t(17) = -0.244$, $p = 0.810$).

Furthermore, we explored noisiness in decision making using ϵ (error) to represent the stability, as before. No significant differences in time pressure conditions here we found here either (paired-samples, $t(17) = -1.160$, $p = 0.129$).

5 Discussion

We found that participants seem to learn the weights as they perform above chance level. Supporting this notion is the observed gradual improvement during blocks. Additionally we observe an effect of time-pressure condition with participants performing better in slow tasks over fast tasks. This suggests that time indeed makes learning harder. From modelling, we see that the Rescorla-Wagner model approximated human behaviour best in both slow and fast tasks. Also, from insights gained from modeling we infer that participants prefer simpler learning strategies under time pressure making more random decisions. Particularly, for two features, time pressure significantly increases noisiness of decisions. Finally, we note that addition of features decreases performance possibly due to increased difficulty, since participants need to take more information into account in their learning. This can also be seen by the increased noisiness in decisions with addition of features in regular (slow) tasks.

Also, we found indication but could not statistically confirm that increasing features might lead to slower learning in regular (slow) tasks and noisier decisions under time pressure (in fast tasks) .

Wu et al. (2019) found that participants seek out uncertain options when time is unlimited, but avoid uncertainty under time pressure (for a four-arm bandit task), but in our case, given two

features, we infer that under time pressure the uncertainty increased. While the task paradigms are different because of the different number of choices (4 vs 2), we think that the contradiction might be because the participants weren't able to learn the association under time pressure (because there just wasn't enough time) and had to resort to random decisions (Fig. 6).

Further analysis can answer questions such as; does doing slow tasks first lead to better performance in subsequent fast tasks? Or, whether there's a significant correlation between reaction time and decision noisiness.

Building upon the work presented in this paper, potential improvements could be, comparison of more advanced learning models, like Heuristics, Neural Networks, Exemplar-Based Models, Gaussian Process etc.. Other decision making rules like Accuracy Maximisation, Probability Matching, Sample Based Rules, Drift Diffusion Models, Resource-rational Models can also be compared.

6 References

Binz, M., Gershman, S.J., Schulz, E. and Endres, D., 2022. Heuristics from bounded meta-learned inference. *Psychological review*.

Schulz, E., Speekenbrink, M. and Krause, A., 2018. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, 85, pp.1-16.

Wu, C.M., Schulz, E., Gerbaulet, K., Pleskac, T.J. & Speekenbrink, M. (2019). Under pressure: The influence of time limits on human exploration. *Proceedings of the 41st Annual Conference of the Cognitive Science Society*.

[GitHub](#) | [Website](#)