

Modeling Generalized Reward Expectations Using Bayesian Hierarchical Drift- Diffusion Modeling

Itzel Martinez

Angela Radulescu (Capstone Advisor)

Mount Sinai School of Medicine

Master's in Biomedical Data Science

Summary 1 Introduction 2

Datasets 3 Participants 3 Data Characteristics 4

Task Design 4 Learning Phase 5 Compound Generalization Phase 6

Methods 7 Value Difference 7 Model Fitting 9 Results 11 Discussion Limitations

11 Future Directions 13 Resources 14 Acknowledgements 16 References 18

Summary

Generalized reward expectations are thought to be formed by how people allocate attention to different features. If attention is focused on high-value features, reward expectations are higher. If attention is focused on low-value features, reward expectations are lower. How does mood impact people's attention allocation when they form reward expectations? To answer this question, authors of "Affect-congruent attention drives changes in reward expectations" induced affect in 120 participants and assessed the effects on attention choice and eye-tracking data from a compound generalization task. In this project, we used dockerHDDM, to model the reaction time data from the task using hierarchical drift-diffusion modeling, a computational modeling approach that simulates the evidence accumulation process that goes into human decision-making. We were interested in whether this evidence accumulation process is influenced by affect. The analysis showed that the process by which people form reward expectations is modulated by affect, such that drift-rates are higher when people are either in a positive or negative mood state.

Introduction

Affect (or mood) refers to the persistent experience of a positive or negative emotional state [5]. In a previous study, Bennett et al found that people's reward expectations for novel compound

stimuli were modulated by affect in an affect-congruent way. Positive affect increased reward expectations for compounds and negative affect decreased reward expectations for compounds. . Computational modeling and eye-tracking revealed that the effects were driven by affect-congruent changes in participants' allocation of attention to positive and negative-valence stimulus features. Understanding these mechanisms is important because they give us insight into optimistic and pessimistic expectations that influence an individual's psychological well being. Mood disorders have been viewed as changes in expectation. For instance, mania is associated with highly positive reward expectations, and depression is associated with pessimistic reward expectations.

Feature-based attention[cite] is a process by which animals, including humans, form reward expectations for previously unknown stimuli. For example, in determining a vacation destination, a traveler may use distance, season, expenses, and available tourist attractions to form a generalized reward expectation for that particular destination before they decide to go. Understanding how people form generalized reward expectations for new stimuli is dependent on how people allocate attention to different features. If attention for a stimulus is biased towards high-value features, reward expectations increase, and the opposite is true if attention is biased towards low-value stimulus features. Therefore, human visual attention can be influenced in an affect-congruent manner, depending on the subject's affective state.

Bennett et al, have experimentally induced positive and negative affect on feature-based attention in 120 participants. The participants completed a compound generalization eye-tracking task that assessed the effects of affect on the distribution of attention over different cues within a compound stimuli. In this project, I will analyze data from this study to model affect-congruent attention driven changes in reward expectations using drift-diffusion modeling. This approach will enable us to gain more insight into the role of affect in attention mechanisms that influence how people form reward expectations.

The drift-diffusion model (DDM) [cite] is a computational model often used to quantify processes of decision-making in neuroscience and psychology. This model relies on experimentally observed reaction times, and behavioral choices to examine individuals' differences across cognitive tasks. The four essential parameters of a DDM are the drift rate, decision boundary, and non-decision time. I will fit the hierarchical drift diffusion model to the experimental data by using Markov chain Monte Carlo (MCMC) convergence. The hierarchical drift-diffusion model (HDDM) is an especially useful tool for modeling psychological data, due to its ability to capture group level variability. As the data in this project is split into three distinct clusters groups (participants who experienced a happy mood induction, those who experienced a negative mood induction and a neutral group), HDDM can help us gain insight into the population-level distribution of decision-making parameters for each group.

Datasets

Participants

There were a total of 120 participants used to generate the dataset. 77 participants were female and 43 were male, with ages ranging from 18 to 55. 28 participants were excluded from my analysis because they did not learn the cue-reward associations or encountered technical mishaps when completing the task. The task took approximately 90 minutes for each participant to complete. Each participant completed a behavior task with concurrent eye-tracking recording, self-reported their measures of trait depression, hypomania, and current affect (positive and negative). Participants were randomly assigned to one of the three affect groups (positive, neutral, or negative), and each group included 40 participants.

Data Characteristics

The datasets used to perform my research are two datasets collected from anonymous participants by Bennett, et al.(2021), at Princeton University. The datasets are labeled as “all_data.csv” and “relative_looking_time.csv.” The two datasets were combined to create “merged_df.” The merged dataset contains 10336 rows and 20 columns. The columns represent subject id, emotional conditions (neutral, happy, sad), trial types (1-13), cue placements (left or right), cue types (simple or compound), chosen cues(1-6), looking time, response, and reaction times. Each row represents one trial. There were a total of 89 participants included in the dataset. There were 128 trials for 56 participants and 96 corresponding trials and for the remaining 33 participants. Participants in the positive and negative affect-induction groups completed an additional block of compound generalization choice trials prior to the first affect induction. This enabled the authors to quantify the effects of the affect induction as a within-subjects change from baseline for these participants.

Eye-tracking data were acquired using an infrared eye-tracker(SR research EyeLink 1000 Plus) at a sampling rate of 500 Hz. Fixation points were calibrated prior to the task , and automatic drift-correction was used after each block to re-calibrate fixations to the center of the screen. Participants were required to maintain fixation within 200 pixels of the center of the screen for 1 second before stimuli appeared on each choice trial. Raw output files were analyzed using NivLink, an open-source Python package for preprocessing EyeLink eye-tracking data (available at <https://github.com/nivlab/NivLink>).

Task Design

The behavioral task participants completed was a “space mining game.” Participants were told that they would be mining minerals from different “planets”, and that each planet stimulus could be made from a valuable mineral or a worthless mineral. Each planet's stimuli were marked with different cues (rune symbols). The participants were instructed that each cue or cue combination provides information about the probability that a given planet could yield a valuable mineral

when mined.

This task consisted of three phases: an initial cue learning phase, a learning testing phase, and a compound generalization phase. Two planet stimuli were presented in each trial. One above and to the left of a central fixation cross, and the other above and to the right of the central fixation cross. The participants chose one of the two planets to mine using a computer keyboard. After a 0.5 to 1.5 second delay, the chosen planet became available to mine, and participants were able to extract the mineral by pressing a key. Each trial yielded 3 measures, stimulus choice, choice response time, and reaction time for mining the planet. Mining reaction time was taken as an index of approach motivation for the chosen stimulus.

To induce a positive, neutral, or negative affect during the compound generalization phase of the task, participants were shown three 90-second videos that played positive, negative, or neutral emotional content. Then each participant self-reported ratings of affective arousal and valence before and after each video. The videos were selected from publicly available sources or videos used in previous studies.

Learning phases

In this phase of the task, participants learned to associate six cues with different reward probabilities, and 24 exposures per cue that were randomly given across 3 blocks of 48 trials. Two cues were associated with a probability 25% reward, two with a 50% reward, and two with a high reward of 75% were randomly given to the participants.

The purpose of the second part of the learning phase was to verify if participants have learned the cue-reward contingencies. To ensure that all learning was done in the first half of the learning phase, participants did not receive any feedback on the outcome of their choices. Here, participants completed one block (30 trials; two repetitions of each pair) that consisted of the six initial learned cues in phase 1.

Compound generalization phase

In the third and final phase of the experiment, the researchers introduced compound stimuli that had two rune cues, rather than only one that was observed in the learning phase. Researchers assessed compound generalization in three blocks consisting of 32 trials each. The participants did not receive any additional instructions on how the compound stimuli should be interpreted. This was done to ensure that the participants' generalization was from previous learning and not new learning of the value of the compounds themselves.

The trials in this phase were split into simple probe trials, compound probe trials, and non-probe trials. In the simple probe trials, participants chose between a medium-valued simple stimulus (a single 50% reward cue) and a low/high compound stimulus (a 25% reward cue and one 75%

reward cue). The combination of the low/high compound stimulus averages the value to 50%, making the two options equal in terms of reward probabilities. The compound probe trials consisted of two low/high compound cues. The non-probe trials were a mixture of simple-versus-compound trials and compound-versus-compound trials. The non-probe trials were designed to yield a rich behavioral dataset that can be used to fit computational models of the distinct psychological processes that result in compound generalization. There were a total of 96 trials in the compound generalization phase: 16 trials were simple probe trials, 24 trials were compound probe trials, and 56 trials were non-probe trials. The data used in my project derives from the compound generalization phase of the study mentioned previously.

Methods

The first step to cleaning the data was to combine the data sets into one dataset where the appropriate rows and columns indicate the appropriate participant and features. Then cue value difference was computed by giving the cues options a numerical value: simple cues were given a value of 0.5, and compound cues were calculated using the average of the two values in the compound cue. Once cues had a numerical value, cue value difference was done by subtracting the right cue value from the left cue value (See figure 1.).

Value Difference

Value difference (VD) is the difference between subjective values assigned to different options when it comes to decision making and plays an integral role in decision-making processes and cognitive tasks. In previous literature using DDMs, it has been shown that value difference has influence on the decision threshold, with a larger VD leading to a lower threshold. When VD is small, decision accuracy decreases, as it is more challenging to make a precise decision as evidence accumulates more slowly. Lastly, VD also has an impact on reaction time. A larger value difference results in faster decisions due to the evidence accumulation towards one option being more distinct[cite]. This makes value difference an important measure to quantify in fitting the diffusion model to the data in this project.

In the compound probe trials, participants had to choose between a medium/medium compound cue and a low/high compound cue. The averages of both compound cues vary between -0.25 and 0.12. This range is due to the exclusive left compound stimulus values of 25/75 and 25/50, while the right compound stimulus had probabilities of 50/50, 25/50, and 50/75.

Figure 1 suggests larger reaction times for trials that were completed under the “Happy” and “Sad” conditions when the value difference was 0.0. This suggests that mood has an effect on the evidence accumulation process during decision-making.

Figure 1. Value difference of all compound trials by mood condition. Trials consisted of simple versus compound cues, and compound versus compound cues. The cue types allowed for value differences ranging from -0.25 to 0.12. Trails were grouped by affect: neutral(blue), happy(orange), and sad(green).

Results

Fitting the models

The final preprocessing step to fitting the HDDM model was to remove the outliers. We removed values that were two standard deviations away from the reaction time mean. We also removed reaction times shorter than 250 ms. 3.7% of the data was lost once the outliers were removed. The new dataframe had 9932 rows of participants' trials.

Model	parameters	Parameters modulated by condition
m1	a,v,t	-
m2	a,v,t,sz,st,sv	-
m3	a,t,v	t
m4	a,t,v	v
m5	a,t,v	a
m6	a,t,v	a,t,v

Table 1. Table two shows the 6 models discussed in this paper. The parameters column shows us the parameters that were fit in the model. The last column shows the parameters that were modulated by condition in that model (neutral, happy, sad).

To fit the models, we use Markov Chain Monte Carlo(MCMC) algorithms to obtain samples that approximate the target distributions for the parameters [citation?]. We created 5000 samples and burned 1500 of those samples for all of the models in this project. A total of two chains was used for each model in this project. Two chains are sufficient to capture the variability between chains and to measure the convergence of our target distribution for each of our models because of the large sample size used. We also used a parallel mapping function to run multiple chains efficiently.

We initially created a simple HDDM model (m1). This model was used as a baseline model to determine if the experimental data can be fit using HDDM. The simple m1 model has three parameters: a (threshold value), v (drift-rate) and t (non-decision time). The threshold parameter “a” represents the boundary separation in the decision process. That is, the amount of evidence

required before one can make a decision between two stimuli. A higher value of “a” would indicate that people require more evidence before making a decision. The drift rate characterizes how quickly evidence accumulates in favor of one option over the other. Higher drift rate values indicate faster evidence accumulation. Non-decision time is the time not used for evidence accumulation. Non-decision time encompasses factors such as sensory encoding, motor preparation and other processes. Non-decision time is typically constant across trials and participants.

In our second model (m2), we added additional parameters of “sz”, “st”, and “sv” to capture trial-to-trial variability in the drift-rate, boundary, and non-decision time. “Sz” is the drift-rate trial variability parameter and captures the variability in evidence accumulation process across different decision instances. “St” accounts for the trial-to-trial variability in the non-decision time. Lastly, “sv” represents the trial variability in the boundary separation in the diffusion process. “Sv” captures the variability of the amount of evidence needed to make a decision on each trial.

In our last three diffusion models, we estimated the non-decision time (m3), drift-rate (m4), and boundary separation (m5) by the different condition groups (neutral, happy, sad). This was done to learn what effect, if any, mood can have on the parameters of the DDM.

We also fit a HDD linear regression model (m6) to learn about the relationship between the HDDM parameters (“a”, “v”, and “t”) and value difference. The linear regression model additionally assumes the parameters to vary on each trial according to cue value difference. Here, drift rate will be modeled as a function of value difference and mood condition.

Convergence is essential to obtain accurate and reliable results from Markov Chain Monte Carlo (MCMC) samples. To ensure the Markov chain has sufficiently explored the target distribution we assessed convergence, which is the exploration of the posterior distribution. We used Gelman-Rubin statistics and convergence plots. The Gelman-Rubin (R-hat) statistic is a method of convergence assessment in MCMC that is typically used in Bayesian statistics to evaluate the convergence of multiple chains. The R-hat statistic value is indicative of how well the chains have mixed and converged. Table 2 shows the model chains have converged and sampled from the target distribution effectively.

Model	R-hat Value
m1	1.0001
m2	1.0001
m3	1.0000
m4	1.0011
m5	1.0001
m6	0.9985

Table 2. Average of Gelman Rubin value of each parameter in all models. All values are ~1,

which is the standard for convergence

To verify that the models converged for each parameter individually, we assessed the convergence plots. Convergence plots provide insight into the efficiency and reliability of the sampling procedure and estimated posterior distribution. In figure 2, we see that the simple model parameters of a , v , t , sv , and st show low autocorrelation, and no drifts in distributions or trace plots. This suggests that the models did converge properly, indicating that the chains adequately explored the parameter space and produced stable estimates of the posterior distribution for all parameters. However, the convergence plots for sz , the inter-trial variability in boundary decision parameter, shows a positively skewed distribution and the trace plot shows that the sz parameter is constrained to positive values. This suggests that the sampled values are concentrated toward the lower end of the distribution. This means that the inter-trial variability for boundary decision is low and extreme differences in boundary decision are unlikely.

Convergence Plots

Figure 2. Convergence plots for each parameter in m2. Parameters a (top left), v (top middle), t (top right), sv (bottom left), sz (bottom middle), and st (bottom right) Posterior plots show the traces, autocorrelations, and marginal posteriors of parameters a (threshold), v (drift-rate), t (non-decision time), sz , st , and sv (inter-trial variability parameters for boundary decision, non-decision time, and drift rate respectively) in our simple model. It is key to note that we see no drifts in the trace, the central tendency is symmetric and a low autocorrelation shows us that the model has converged.

To determine whether the mood conditions in models m3, m4, and m5 had an effect on the HDDM parameters, we performed bayesian t-testing to compare the varying drift rates, non-decision times, and boundary decision of the three affect groups. To do this, we compared the probabilities of the diffusion model parameters that the parameters for the happy and sad groups would be larger than the parameters for the neutral group.

Model	Parameter of interest	P(Happy > Neutral)	P(Sad > Neutral)
m3	Non-decision time (t)	0.604	0.751
m4	Drift-rate (v)	0.925	0.877
m5	Threshold (a)	0.666	0.766

Table 5. This measure is statistically significant if the probability is 0.9 or larger.

The Deviance Information Criterion (DIC) is commonly used to measure model fitting and complexity used in Bayesian model comparison. This comparison method quantifies the balance between goodness of fit and model complexity. Lower DIC values are indicative of a better model fit. In the table below, we see the DIC value of all the models in this project. The best fitting model is m2.

Model	DIC score
m1	38665.118
m2	38500.870
m3	38662.575
m4	38663.928
m5	38664.943
m6	38609.051

Table 4. In the table above we see the best performing model, according to the DIC, is m2.

Discussion

Figure 3. Posterior plot of m2 (left), our best fitting model and posterior of drift-rate (right). Here we see the distribution of all parameters values from the posterior samples on the left, and on the right we see the distribution of drift rate by condition.

In regards to diffusion parameters ‘a’ and ‘t’, the posterior distribution is positive for all non-decision time and threshold values. However, the estimated group means differed in drift-rate. Both happy and sad groups had a positive posterior distribution, but the neutral group had a negative posterior distribution (figure 3). This indicates that the neutral group had a preference in the left sided cues, where there was a medium/medium valued cue while the happy and sad groups preferred the right cue value, where the stimuli had a low/high value. As mentioned earlier, a larger value difference results in faster decisions due to the evidence accumulation towards one option being more distinct. Drift rate is estimated from observed reaction time. In the probe trials where the value difference is zero, reaction time should be similar between groups. However, they are not similar in our analysis of the experimental data. The difference in reaction time suggests that when aroused by affect, participants were valuing the cues in an affect congruent manner.

According to DIC scoring, our best fitting model is model m2, which is the only model that included the additional trial-by-trial variability parameters. This suggests the model best captured fluctuations in the data and produced the most accurate representation of the underlying decision-making mechanisms. Although it is important to note that DIC is not the best metric for Bayesian model comparison. In table 5, we see that the probability for a happy affect participant to have a larger drift rate than a neutral affect participant is 0.93, which is statistically significant and verifies our belief that drift rate as a function of value difference can be used to model affect-congruency in generalized reward expectations.

In Bayesian statistics, the posterior distribution represents the updated knowledge about our parameters of interest after taking the data from the experiment into account.

Future Directions

The models were compared using the deviance information criterion (DIC). DIC measures the fit of the model to the data, and it penalizes complexity due to addition of degrees of freedom. DIC is not the best model comparison method, although it is supported by the HDDM package. Other model comparison methods may be more accurate. LOO calculates the log-likelihood of each data point to evaluate predictive accuracy. The Widely Applicable Information Criterion (WAIC) is based on expected log pointwise predictive density (ELPD) to estimate out-of-sample prediction accuracy. These model comparison methods can provide more insight from the models in this project. However, due to the complexity of the HDDM likelihood function, it is computationally costly to convert HDDM objects into inference data.

In continuing the research, we would suggest to model the linear regression models with a larger sample size to improve the model performance, and include the trial-by-trial variability parameters sz , st , and zv . Secondly, modeling the variability parameters by condition may provide more detailed insight into the nuances of decision-making. Due to computational complexity, it was not favorable to run locally. Lastly, it would be interesting to model non-decision time and boundary decision as a function of value difference and analyze the findings from those regression models.

Materials/Tools

dockerHDDM

dockerHDDM was used to fit all drift diffusion models (DDM) to experimental data, and to complete their analysis. dockerHDDM is a docker image used to run the HDDM library. ArviZ, a Bayesian modeling Python package, has been integrated into HDDM and is available through Docker.

Google Colab

Google Colab was used to complete all preprocessing of the data.

Resources

All computation was done in the Python programming language, with Jupyter Notebooks running on a Macbook Air. I used Google Colab to run preliminary analysis with a hosted run time. All HDDM analysis used several Python packages such as pymc, hddm, kabuki, and arviz. Parallel processing was performed using p_tqdm, and functools. Lastly, numpy, pandas, xarray, matplotlib, and seaborn packages were used for scientific computing and visualization.

Acknowledgements

The authors of this paper wish to acknowledge the Icahn School of Medicine and the DCA Lab at the Center for Computational Psychiatry at Mount Sinai for providing the guidance for the work done on this project. The authors would also wish to acknowledge Bennett D, Radulescu A, et al. for providing the data source for this project and the participants who contributed to the data. This work was supported by Angela Radulescu and Blair Shevlin as advisors.

References

- [1] Bennett, D., Radulescu, A., Zorowitz, S., Felson, V., & Niv, Y. (2021, November 4). Affect-congruent attention drives changes in reward expectations. <https://doi.org/10.31234/osf.io/vu2cw> - <https://psyarxiv.com/vu2cw/>
- [2] Hu, C., Geng, H., Zhang, L., Fengler, A., Frank, M., & Zhang, R. (2022). A Hitchhiker's Guide to Bayesian Hierarchical Drift-Diffusion Modeling with dockerHDDM. *PsyArXiv*. <https://doi.org/10.31234/osf.io/6uzga>
- [3] Basic HDDM Tutorial — HDDM 0.9.8 documentation. (n.d.). https://hddm.readthedocs.io/en/latest/tutorial_basic_hddm.html
- [4] Wiecki TV, Sofer I and Frank MJ (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Front. Neuroinform.* 7:14. doi: 10.3389/fninf.2013.00014
- [5] <https://study.com/learn/lesson/affect-psychology-overview-types-examples.html>
- [6] Bringing the drift-diffusion model (DDM) to PyMC3. (2021, March 22). PyMC Discourse. <https://discourse.pymc.io/t/bringing-the-drift-diffusion-model-ddm-to-pymc3/7057>