Emotion prediction errors guide socially adaptive behaviour

BSE662A – ENDTERM REPORT – TEAM HOOMANS

Aman D. (190103) | Kunal S. (190449) | Pranjal S. (190627) | Raghavan G. (190660)| Siddharth G. (190838) | Tanishq G. (190894)

Indian Institute of Technology, Kanpur ***May 03 , 2022***

Introduction

Decision-making plays a crucial role in the life of all individuals and a topic of great interest is the factors that affect decision-making. The expectation of reward greatly influences an agent's choice during decision-making. While its violations, which are termed Prediction Errors (PEs), help an agent upgrade his knowledge of his habitat to facilitate his survival. In the former, one's behavioral reaction increases while the latter maintains well-established appetitive behavior. The standard decision-making models embedded reinforcement learning framework, encapsulating these insights. The violation of expected outcomes is supposed to encourage even sophisticated behavioral patterns.[1].

Need for the original study- The paper is a crucial study in the field of decision making as:

1. The background for the study arises from the literature demonstrating a significant influence of emotion on choice apart from reward in decision-making. The transition between distinct emotional states and expected aversive emotions shaping social interactions has been predicted in earlier works using sophisticated mental models of emotion.
2. Previous studies and models like Pearce-Hall reinforcement learning have focused only on reward prediction errors and considered them the sole contributing factor to adaptation and decision making.
3. Emotions have always been treated as a nuisance variable or an internal proxy for value. Past studies have ignored the potential role of emotions, and the effect of violation of emotions on decision-making is largely an open question.
4. In social scenarios, where two or more people are interacting and exchanging resources, money, etc. there is a good possibility that emotions could be factoring into the decisions being made. Also, emotions tend to evolve with time as the situation changes, which the authors have studied using the Justice Game.
5. Finally, the paper also studies how depressed/ in-risk individuals make decisions. This is necessary as such people have different mindsets and emotions, which may again factor into decision making. Further, showing that depression-risk individuals make decisions differently from normal ones would be a good indicator that emotions and violation of emotions do indeed factor in decision making and social behaviour.

Limitations of Original Study

According to us, the limitations of the study are:

1. Using the dARM measure to find emotion prediction errors is not very reliable. Participants of the experiment are expected to gauge their feelings and mark them on dARM, which may not give accurate results as people are generally not good at judging their emotions.
2. In real-life social interactions, the parties involved generally know each other. However, in the Ultimatum and Justice games played by the participants in the study, the offers were given by an anonymous source. Therefore, we believe that the role of emotions may have been suppressed.

Our Aim

Apart from replicating the results of the Ultimatum game from the original paper, our project attempts to overcome the limitations stated above:

1. To remove anonymity and make interactions more realistic, we have conducted trials where the proposer and responder in an Ultimatum game know each other (more specifically, they are friends living in the same hostel wing)
2. We have used emotion recognition software that reads emotions off images. This is in theory a more accurate way to determine emotion prediction errors than dARM.

Methods

The Ultimatum game:

Anonymous- An anonymous proposer gives an offer to the participant for which he has to decide whether to accept or decline. If accepted, the money is divided as proposed by the proposer. But if declined, both proposer and participant get nothing.

Non-Anonymous- Same concept but this time the proposer is not anonymous, but someone who the participant knows.

Method:

Participants: The experiment was conducted in three wings parallelly. 10 participants and 5 proposers from each wing. 10 trials were conducted for each participant, 5 for anonymous proposers, and 5 for the non-anonymous proposers. So overall 30 participants recorded 300 data points(150 - anonymous; 150 - non-anonymous).

Pre-Processing: Before conducting the experiment, 5 proposers were chosen from each wing (a total of 15) to list down offers for 10 of their wingmates who were going to play the Ultimatum Game.

Process: An experiment setup was set to collect the data conveniently. One laptop had the game running on it which the participant used. Parallelly the screen was mirrored to a monitor where a member of our team is monitoring the emotions of the participant when he was marking on dARM. A webcam was kept in front of the participant and as he marks the emotions on dARM, the picture is being taken and saved in the format <name_trialNum_expected/actual>. The participant is first given 5 anonymous offers followed by offers from his wingmates. After completing 10 trials, the next participant was called.

Fig 1*: Experiment Setup - Front View*

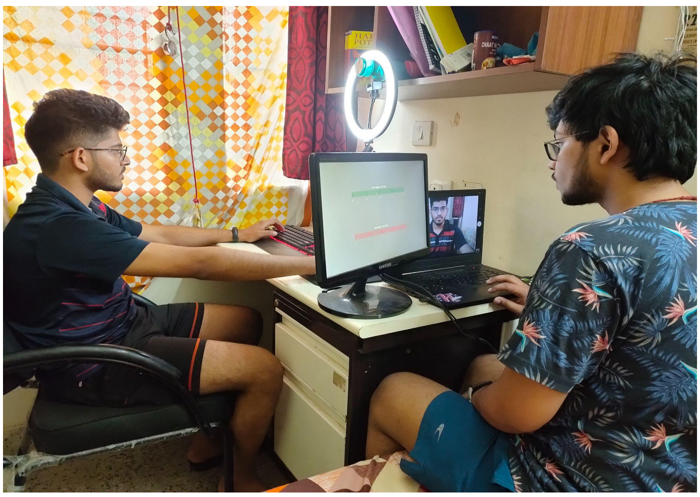

Fig 2*: Experiment Setup - Side View*

dARM: Instead of following the normal convention of the 2D plane for marking emotions, we transformed 2D into the 1D plane by using two sliders, one for each valence and arousal. This made understanding dARM easier for participants and more convenient.

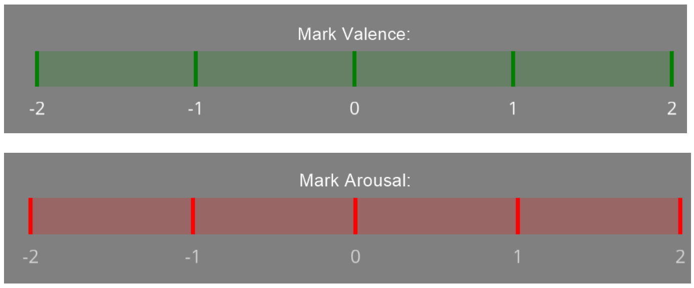

Fig 3*: Parallel dArm used in our game*

Reward: Instead of providing the preset amount of money as participating compensation, we rewarded participants based on their decisions. We kept 35 Rs as the maximum amount the participant can win. Money Won = (Total Accepted Money / Total Offer Money) x 35 Rs.

Emotions: The pictures taken of participants were analyzed using emotion recognition software, Noldus FaceReader Software. The software gave the pictorial representation of the emotions detected in 7 parameters (Neutral, Happy, Sad, Angry, Surprised, Scared, Disgusted). The pictorial representation was converted into numerical value by measuring the line dimensions in pixels. The total number of photos analyzed was 320 (only 16 participants gave consent).

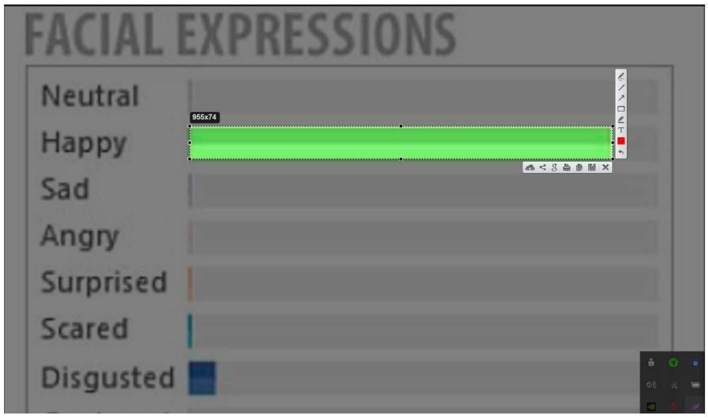

Fig 4*: Graph for The Ultimate game - Non-Anonymous*

Analysis:

We plotted the graph of Prediction Error vs Probability of rejection. After compiling the data for each participant, it was segregated in terms of anonymous and non-anonymous and the corresponding graph was plotted using python script. To remove the variations, we smoothened the curve by taking the mean of every consecutive 3 points and constructed the graph for the same. For each mean, the maximum and minimum of the three points are used to plot the upper bound and lower bound of the error bar for each plot.

Original Result

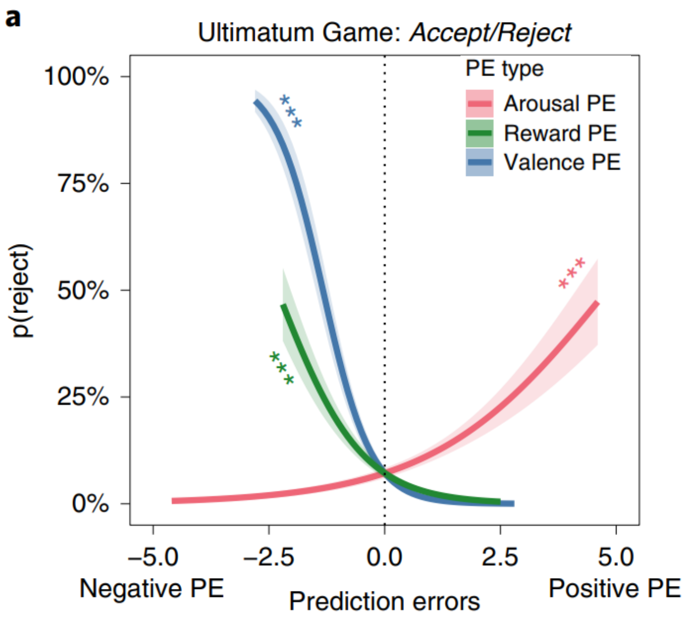

Fig 5*: Graph for The Ultimate game - Original*

After conducting the experiments, in the original paper, it was found that both reward and emotional PEs make distinguishable contributions toward making a choice. It was also found that participants who experienced more arousal PE or less valence PE or less reward PE were more punishing.

There are mainly three major implications from the results obtained. Firstly, both reward and emotional PEs make distinguishable contributions toward making a punitive decision. Secondly, the beta values indicate that valence PEs have the highest impact on punitive decisions. Thirdly, when the direct contributions of experienced reward and emotion are evaluated, reward appears to bias behavior more strongly than emotion, demonstrating that emotion only surpasses reward if PEs are taken into account.

Upon replication of the experiment, we found our results similar to those of the original research. The conclusion made by the original paper that reward Prediction error and Valence Prediction error graphs following the same trend can also be noticed in our graph. We denoted the error bar using dotted lines, which is a little bit more diverged from the original plot due to less number of datasets. Taking more observations will increase the accuracy and further solidify our result.

Our Result

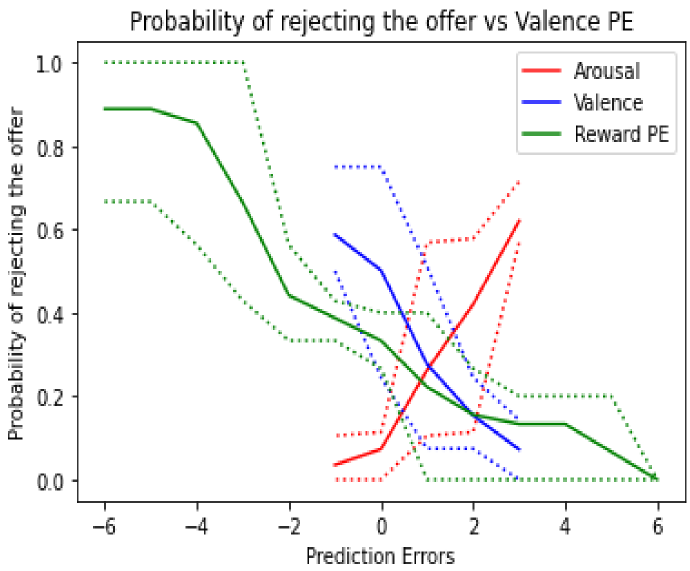

Fig 6*: Graph for The Ultimate game - Anonymous*

When we repeated the experiment for the non-anonymous proposer, we get a similar trend following the graph but we could observe a significant increase number of rejections for a particular prediction error.

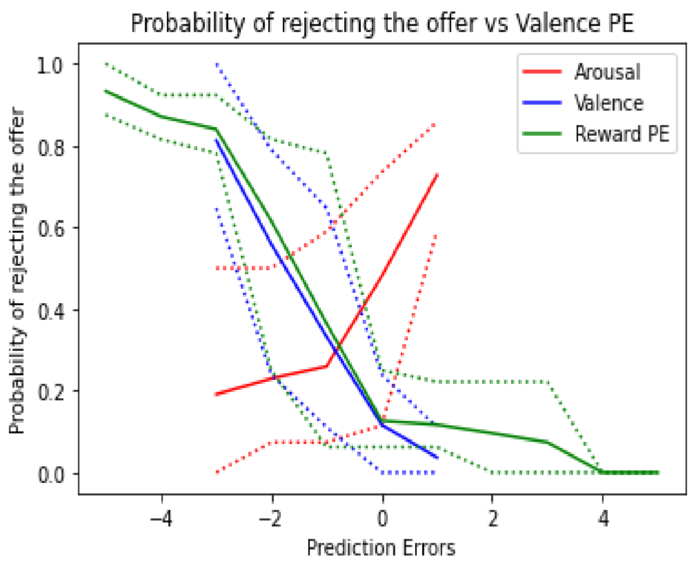

Fig 7*: Graph for The Ultimate game - Non-Anonymous*

On comparing anonymous and non-anonymous graphs, we can notice that the non-anonymous graph is more abrupt as opposed to the gradual slope of the anonymous graph. The probability of rejections is greater for negative prediction errors in the later graph.

To compare the graphs of the two results more accurately, we plotted the deviation graph: (Non-Anonymous – Anonymous)

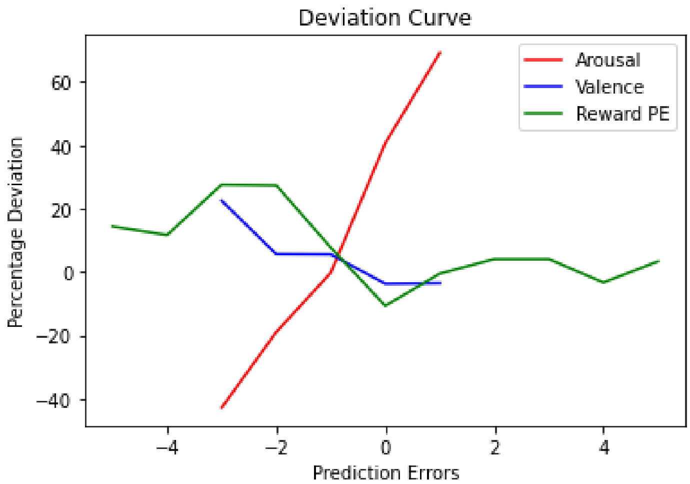

***Fig 8**: Deviation graph between Non-anonymous and anonymous*

We can observe the positive deviation for reward prediction error and valence prediction error in the graph for negative prediction error. And accordingly positive deviation for arousal prediction error for positive prediction errors. As we tend towards zero, the deviation also converges towards zero. Reward prediction error of non-anonymous graph has as much as 30% deviation from the anonymous graph.

We also made a claim about expecting more from our friends than from an anonymous proposer. So to support it, we plotted a bar graph (Fig 9) depicting the average of expected rewards by each participant separately for anonymous and non-anonymous trails.

We can observe that for most of the participants, reward expectations from a non-anonymous proposer is greater than reward expectation from an anonymous proposer. After compiling the results of each participant we get:

Average Expected Reward - Anonymous = 48 Rs

Average Expected Reward - Non-Anonymous = 53 Rs

Significance and Discussion

As seen from the results highlighted above, arousal and valence prediction errors do play a role in decision making. Both the original study and our replication have shown that the probability of rejection increases with the prediction error in arousal (feelings of anger, surprise). The authors have also pointed out a noteworthy unique contribution of valence PEs for decisions to punish, which can be seen in our study as well.

Our studies with non-anonymous participants have also yielded some interesting results that indicate the following:

Emotions take the front seat

When the participant knows the proposer, emotions take high priority, while reward-motivated decisions are less commonplace. People tend to have different expectations from people they know as compared to anonymous proposers, which is reflected in their decisions and our results. The abrupt nature of the graph also supports this observation.

Reward motivated decisions are reduced

As the original aim of the paper, is to separate reward from emotions and observe the result, the non-anonymous experiment achieves the objective more successfully. By introducing non-anonymous proposers, emotions are amplified and start having more intense effects on the decisions. This is the reason for the increase in the number of rejections for negative prediction errors, even a slightly off offer from expectations leads to rejection of the offer more than half of the time in the case of a non-anonymous proposer.

Smaller Buffer Region for Rejection of Offer

In the standard UG with unknown proposers, participants tend to accept unfair offers to a large extent. However, when

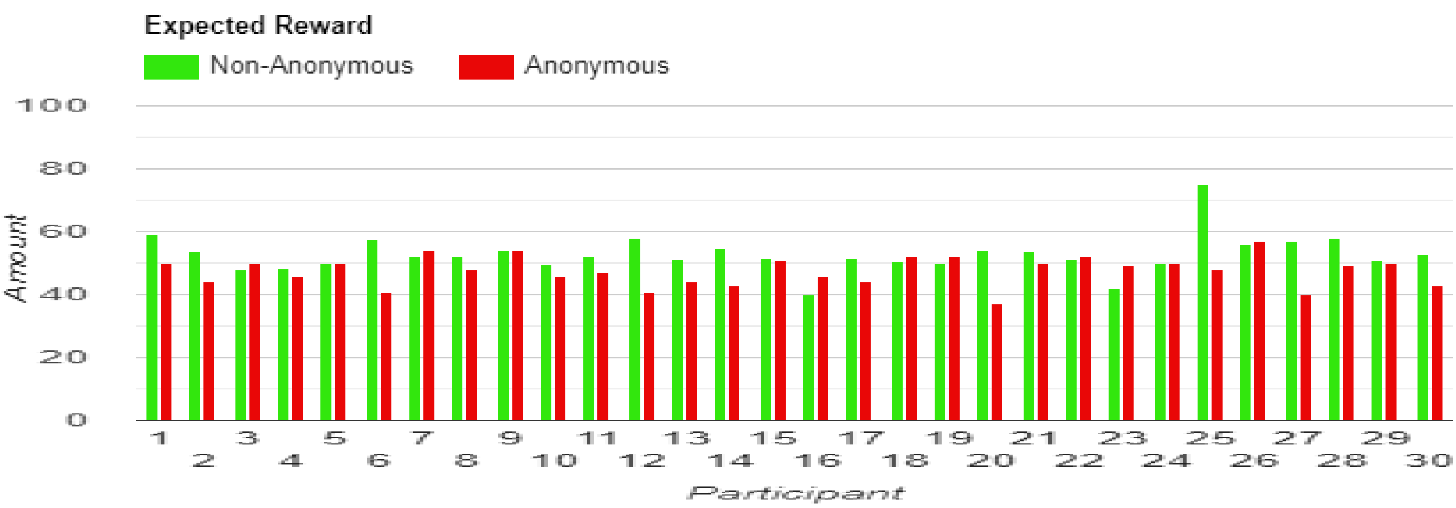

***Fig 9**: Expected Reward from Anonymous and Non-Anonymous*

proposers are known, a small prediction error can lead to the participant rejecting the offer.

Expectations Bar is Set Higher

People tend to expect fairer/ higher offers from people they know i.e. the initial expectation of reward is higher than in the standard UG where people tend to expect unfair monetary offers.

Limitations of Our Study

Emotional Analysis using Facial Recognition:

The results of the emotional analysis from imaging were not at all like we expected. The results were not enough to draw any conclusion. Most of the time, the expression of the participant remained neutral or may have been very subtle to catch noticeable change by the Noldus software, though sometimes when the value given was very high or very low then we did get some expressions. But to infer any claim from such variable data was not intuitive.

One theory may be that as the experiment was not able to exactly replicate and give them the feeling of a real-life situation, it may have not provoked any considerable emotions.

Also, it was observed that the expression of emotions is very subjective in human beings, it depends upon one's personality if he/she is likely to show it or be subtle about it.

Also, one potential reason for the same could be that the participants didn't take the experiment seriously, and even if they did then, chances are that the monetary value was not large enough for them to provoke any emotions.

Sources of Error:

No experiment is free of errors. The primary causes for the errors that have affected our study are:

1. **Small dataset**: The original study had around 364 participants, each playing 20 rounds of the Ultimatum Game. Due to time and resource constraints, our study had only 30 participants, each playing 5 games with an anonymous proposer and 5 with a known one. Due to this, our results are not very accurate and the graphs obtained are not smooth.
2. **Approximation in data**: To obtain the curves, we have done some approximations to the data. For example, there was a need to approximate some points so that the probability of rejection corresponding to a prediction error could be calculated as taking exact values would result in a 0 or 1 probability.
3. **Participant error:** Some of the errors can be related to the participants- some may have not understood dARM completely, some may have been disinterested in the study and carried it out half heartedly, stress due to upcoming endsems may have factored in, etc. In addition to this, the environment where the study was conducted may have induced participant errors due to background noise and other factors.

Finally, another limitation of our study is that all participants were of the same age (20-22 years) and were mostly male. Therefore, the diversity that was present in the original study, and that is desired in any general study of human decision-making is absent.

Follow up Studies and Conclusion

The estimation of valence and arousal correctness is quite uncertain. A good follow-up study will be to do the above experiments using a more modern and novel technique like a Study Mapping brain waves can perform more comprehensive research to determine how emotion and prediction errors are produced inside our brains. This technique is not only easily implementable and will provide real-time results as well; hence we can use it in all of the experiments and get even more robust results.

A simple suggestion will be to increase the sample size and even precise results. However, a good idea will be to add people from different cultures and backgrounds. For example, in the games, only around 35 Rs were at stake. Losing or winning that amount to many participants may not provoke any emotion but to others (underprivileged) it will be a great deal. Getting data from different cultures will also give us more insight into to what extent the reward and emotions are separable. Dissociating reward and emotion may more generalize the research by accurately depicting emotions in decision-making. Are their pathways separated from influencing an individual's decision or work collectively to determine behavior as a whole? As we further dive into the processes involved in the brain due to prediction errors, we can analyze how neural networks associated with emotion PEs evolve with time as we develop more complex emotions like trust, envy, love, etc.

As we observed deviation in the non-anonymous curve from the anonymous curve, we can infer that some other theory is responsible for influencing people's decisions. This study can further be conducted on a large scale professionally to check the robustness of the observation presented by us.

Also, we can take into account social discounting while conducting the non-anonymous study because for us, the value of every person is different and that can be reflected in our result as noise.

The experiment provided a different perspective on the decision-making process and challenged the belief that only reward PEs are the sole factor influencing our decisions. The results of the social experiment conducted by us answer the vital question of how many of our decisions are motivated by emotions, and among reward-oriented and emotion-oriented decisions, which one is more dominant? Emotions are a critical part of our life, shaping us as unique individuals and navigating us through the ever-evolving social world. From our actions to our learning processes, every domain can be determined by our emotional experiences. By understanding how they are encoded in our brain, we can find the vital key to solving the mystery of making decisions!

Appendix

The data, game, code, and results can be accessed from: https://drive.google.com/drive/folders/1xLEX23c7Yhw-KwrTvZc9fYb9M68nLF8l?usp=sharing

References

1. Heffner, J., Son, JY. & FeldmanHall, O. Emotion prediction errors guide socially adaptive behaviour. Nat Hum Behav 5, 1391–1401 (2021). https://doi.org/10.1038/s41562-021-01213-6

Contribution by Team Members

All team members have contributed equally to this project and report. We have all written different parts of the report and examined each other’s contributions to ensure that we are submitting the best work possible.

Aman Dixit - 16.67%
Kunal Singh - 16.67%
Pranjal Sharma - 16.67%
Raghavan Gopalan - 16.67%
Siddharth Gupta- 16.67%
Tanishq Gupta- 16.67%