Effect of Valence & Uncertainty on Reinforcement Learning

A thesis
Submitted towards the partial fullfilment of
BS-MS dual degree programme
by

Swarag. T(20181013)

Date:1st April 2023

under the guidance of

Dr. K.M. Sharika

Assistant Professor, Department of Cognitive Science, Indian Institute of Technology Kanpur

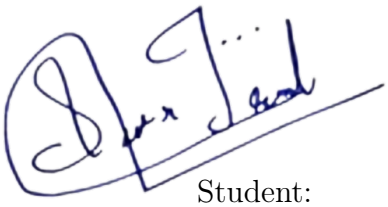
from May 2022 to Mar 2023

Indian Institute of Science Education and Research
Pune

Certificate

This is to certify that this dissertation entitled 'Effect of Valence & Uncertainty on Reinforcement Learning' submitted towards the partial fulfillment of the BS-MS degree at the Indian Institute of Science Education and Research, Pune represents original research carried out by Swarag T. at Indian Institute of Technology Kanpur under the supervision of Dr. K. M. Sharika during academic year May 2022 to March 2023.

Supervisor:
Dr. K.M. Sharika
Assistant
Professor
Department of
Cognitive Science,
IIT Kanpur
Date: 01/04/2023



Student:
Swarag. T
Roll No. 20181013
BS-MS
IISER Pune
Date: 01/04/2023

Committee:
Dr. K.M. Sharika
Dr. Raghav Rajan

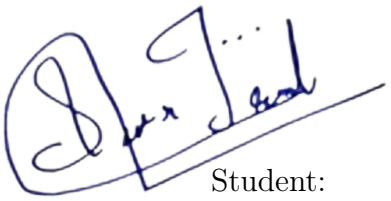
This thesis is dedicated to all living and non-living entities in the perceivable universe, which give me, at least, if not absolute, an illusion of my life being "real," motivating me to do things, including this one. Even though I am not sure, this is real. These may be illusions against which we still have to figure out how to test them...

> *"Reality is merely an illusion, albeit a very persistent one."* - ***Albert Einstein***

Declaration

I, hereby declare that the matter embodied in the report titled "Effect of Valence & Uncertainty on Reinforcement Learning" is the result of the investigations carried out by me at the Indian Institute of Technology Kanpur from the period 01-05-2022 to 01-04-2023 under the supervision of Dr. K. M. Sharika and the same has not been submitted elsewhere for any other degree.

Supervisor:
Dr. K.M. Sharika
Assistant
Professor
Department of
Cognitive Science,
IIT Kanpur
Date: 01/04/2023



Student:
Swarag. T
Roll No. 20181013
BS-MS
IISER Pune
Date: 01/04/2023

Acknowledgements

I would like to express my deepest gratitude to my supervisor Dr. K.M. Sharika, for her guidance and support for my thesis project, Dr. Raghav Rajan for his constructive criticism during my mid-year review, as an expert member in my thesis advisory committee(TAC), which helped me a lot.
I want to thank my lab members for their discussions and constructive suggestions. And other students and faculties in the department.
I would especially like to thank the Department of Cognitive Science, IIT Kanpur, and Computer Center, IIT Kanpur, for providing computational resources for running the models and analysis. I would specially like to thank Kishore Vaigyanik Protsahan Yojana(KVPY) for providing me scholarship for the duration of my thesis work.

Contributions

Contributor name	Contributor role
Swarag. T, Dr. K.M. Sharika	Conceptualization
Swarag. T, Dr. K.M. Sharika	Methodology
Swarag. T	Software
–	Validation
Swarag. T	Formal analysis
Swarag. T, Ansh Bharadwaj	Investigation
Dr. K.M. Sharika	Resources
Swarag. T	Data Curation
Swarag. T	Writing – original draft preparation
Swarag. T, Dr. K.M. Sharika	Writing – review and editing
Swarag. T	Visualization
Dr. K.M. Sharika	Supervision
Dr. K.M. Sharika	Project administration
Dr. K.M. Sharika	Funding acquisition

This contributor syntax is based on the Journal of Cell Science CRediT Taxonomy[1]

[1]https://journals.biologists.com/jcs/pages/author-contributions

Abstract

In prior work, a bias towards negative valence had been observed in humans and animals. This bias was explained through attention, memory, and response biases. Previous studies also indicated that uncertainty could interact with these behavioral biases, but the underlying mechanisms were not well understood. In this study, we approached this multifaceted problem by considering how valence, levels of expected uncertainty, and traits modulate the output bias using a modified version of the cue association learning task. We used a newly derived axis and reinforcement learning drift-diffusion model (RL-DDM) to assess the bias in behavioral data. Data collected includes 70 participants(40 in 70% and 30 in 80% correct feedback blocks) . Supporting our hypothesis, we found that 1) uncertainty modulated the negative bias and 2) this modulatory effect was captured in the prior bias parameter of the RL-DDM($p < 0.05$ for 70% but not in 80% correct feedback blocks). Additionally, we found that the bias coded in a new normalized variable was explained by trait variables(STAI)(β=-0.03, p=0.02), indicating that the bias is not invariant and depends on the population. This provided a more generalizable approach for addressing such biases and their potential implications in clinical studies.

Contents

List of figures

List of tables

(b) 6.2- Forward model selection table for RL-DDM model in mixed and symbolic feedback blocks of(70% correct feedback)

(c) 6.3- Forward model selection table for RT models in 70% correct feedback blocks

(d) 6.4- Forward model selection table for RT models in 70% and 80% correct feedback blocks, for symbolic and mixed feedback.

Chapter 1

Introduction

Decision making is a complex process, an earliest account of which could be seen in the review "Theories of Decision-Making in Economics and Behavioral Science" by Herbert A. Simon (Simon, 1966). In this article, the author raises the argument that economics and psychology can have intervening interests, and if economists are able to define a theory that is well generalizable about economic behaviour of humans, it is natural to expect that the same could be adapted to psychology and human behaviour. Microeconomics could be thought of as the branch that deals with the interaction between the agents. But even there it is interested in how agents or individuals ought to behave and not how they behave in a certain way, contrary to in the behavioural sciences. On the other hand, when it comes to Macroeconomics, where individual actors aren't considered anymore, they rely on strong assumptions such as the rationality of the individuals in the population (Kurz, 2016). von Neumann and Morgenstern extended the question about choices with uncertain prospects by suggesting that behavior could be explained in terms of cardinal utility, which is the satisfaction obtained after consuming any goods or services that can be scaled in terms of countable numbers. Individuals who are following von Neuman's axioms would try to maximize the utility, which is the weighted sum of the outcome, where the weights are the probabilities in an uncertain condition. The probabilities that these subjects assigned were assumed to be identical to the objective probability that the experimenter knew. However, when utility theory was tested using experimental data the difference between the "objective" and "subjective" probabilities became more apparent (Kurz, 2016): This resulted in a concept proposed by Frank Ramsey where he showed via careful experimentation that this internal probability could be estimated(Ramsey, 2016).

Among the many paradigms used, the two alternative forced choice tasks are very popular. An example is the two-armed bandit task where a correct

and an incorrect choice would be assigned randomly with certain probability to two given symbols. That is if an option is correct for p fraction of the total number of trials, it will be incorrect for the remaining 1-p fraction of trials. The participants in such games can stick to one option as well. But, the classical utility theory would tend to fail here, because the maximum reward on offer can be received only via event matching (Simon, 1966), i.e., when an individual chooses alternatives in proportion to that option getting rewarded. Hence, as per sequential sampling theory (Simon, 1966) individuals are hypothesized to compute their expected outcomes while sampling an option which is associated with either reward (if the choice was correct) or information about that option (if the choice was incorrect). When behavior is found to evolve in such an adaptive manner it is said to constitute a learning process where after choosing an option, the probability to go for the same would increase if that choice is followed by a reward. Reinforcement learning (Minsky, 1961) models are a strong tool for investigating such associative learning (Schultz *et al.*, 1997;Sutton *et al.*, 1998) known to be modulated by prediction errors (Schultz *et al.*, 1997), i. e., the internal representation of the mismatch between expected and obtained outcome.

Sequential sampling models are widely used to describe the decision making process in cognitive neuroscience. These models assume that people make time constrained decisions by sequentially accumulating evidence over time(Forstmann *et al.*, 2016). Earliest explanation of such models can be seen from the 1960s(Stone, 1960). From the modeling perspective, the paradigms are typically two alternate choice tasks where the decision making scenario is repeated across multiple trials, resulting in large number of repetitions to estimate parameters (which is required to reduce standard error associated with the parameters). And the response time associated with correct and incorrect choices are recorded(Simon, 1966;Forstmann *et al.*, 2016). Empirical studies collecting response time and choices were found to capture some properties of decision making, namely, shorter response time for easy compared to hard stimuli and an emphasis on speed generating shorter response times, but, more errors (Forstmann *et al.*, 2016). At the broader level, these models have two boundaries depicting alternate choices, and at the finer level these models differ in terms of whether they are having a single accumulator or multiple of them. Even though over time the statistical methods and the models became sophisticated, the underlying principle of sequential sampling models remains the same today (Forstmann *et al.*, 2016).

1.1 Biases in decision-making

Animals make decisions by accumulating evidence, but, this can be subjected to biases (Rescorla, 1988) which can either be, at the perception, decision or execution level. Pavlovian bias, in which the animals would approach appetitive stimuli even when it is a sub-optimal choice was reported by Hershberger et al(Hershberger, 1986). In an experiment done on chicks, he showed how chicks when had to retract from food in order to obtain it, they could not learn the optimal strategy. Pavlovian bias is hypothesized to result from the dopaminergic neurons signaling for reward in the environment(Ereira *et al.*, 2021) also being responsible for invigorating approach responses(Mohebi *et al.*, 2019;Hamid *et al.*, 2016). And the dominant argument for the same signaling pathway carrying out two functions is that both functions are likely to be highly correlated in naturalistic conditions (Berke, 2018;Rangel *et al.*, 2008).

Biases in decision making have been proposed to be useful in assessing the emotional or affective state the animals are in(Harding *et al.*, 2004;Mendl *et al.*, 2009). According to Russel et al(Russell, 2003), an affective state could be described as "a neurophysiological state that is consciously accessible as a simple, non reflective feeling that is an integral blend of hedonic valence (pleasure–displeasure) and arousal (sleepy–activated) values". The first appearance of the term "valence" could be traced back to 1935, when the German-American psychologist Kurt Lewin introduced the term (Lewin, 1936)borrowing it from physical sciences, especially the valence-bond theory in chemistry. Lewin believed that the psychological phenomenon could be better explained by something conceptually similar to that in the physical sciences, where the attraction and repulsion was explained in terms of angular momentum and spin. More importantly, the point of view was that the valence could be acting as a force applied on an individual that in turn determines their behavior. The idea was that certain objects or experiences may exert attractive or repulsive force on individuals, which can either be motivating or inhibiting. Although, Lewin emphasized that the same object could cause different valenced experiences depending on place and time, with each encounter changing how things are perceived and subsequent responses, towards the later part of the 20th century, the concept of valence got refined as a single and bipolar dimension. And valence along with arousal became the important dimensions of core affect(Russell, 2003).

Harding et al (Harding *et al.*, 2004) reported how the emotional state of an individual affects decision making and learning and suggested that the cognitive bias could be an indicator of affective state in an animal. For example, an individual in a negative emotional state will make negative (pes-

simistic) judgments about an ambiguous stimulus in comparison to someone in a positive emotional state (Lagisz *et al.*, 2020). In Harding et al(Harding *et al.*, 2004), two tones having different frequencies were used as cues for a positive and a negative outcome. Participants were required to respond in a particular way to the positive cue for obtaining reward (food) and differently to the negative tone in order to avoid aversive outcome. After cue learning, ambiguous cues (ones having intermediate frequency) were presented and participants' response to them recorded to assess "optimism" (making more positive responses) vs. "pessimism"(making more negative responses), and it was found that there is a negative bias.

However, an important point to note in these studies is the use of ambiguous stimuli which by their very nature could be perceived as negative(Anderson *et al.*, 2019). This is because ambiguity is a type of uncertainty which can be broadly classified based on probability or the risk associated with the future outcome. It is said that any agent would try to reduce the uncertainty to accurately estimate the occurrence of motivating events, which is to reduce the amount of resources spent (Hirsh *et al.*, 2012;Peters *et al.*, 2017). Therefore theoretical models suggest that the uncertainty is aversive in itself and it activates behavioral inhibition systems(BIS) (Hirsh *et al.*, 2012;Morriss *et al.*, 2022b). A set of literature, in line with this argument, has shown that uncertainty dampens positive emotional state and intensifies negative emotional state (Morriss *et al.*, 2022b).

The appraisal theory(Scherer, 1999) treats emotions as an adaptive process which reflects the future state of the environment, much like the concept of momentum, thus perceived uncertainty could in turn determine what emotion is getting elicited. For example, when considering an animal whose major food source is a specific fruit that is abundant in spring season as the timeline approaches spring the uncertainty about getting that fruit decreases as a result its emotional state may also become positive. But, an agent can always use cognitive processes to modify the computational processes and actions. One can even see affect and emotion as a way to make quick and frugal decisions when the outcome is uncertain or complex where deliberate reasoning and evaluation would be impractical(Slovic *et al.*, 2007). Importantly the affect heuristic theory proposes that mental representations in general can carry an affective tag(Slovic *et al.*, 2007). Risk as a feeling theory(Loewenstein *et al.*, 2001) suggests that people use emotion that contains information about uncertainty when making decisions. It has been proposed that the fear of the unknown is the most prominent and fundamental fear of all human beings(Carleton, 2016). However, interestingly enough, uncertainty can also bring positive effects as in case of a pleasant surprise, and thus may not be generalizable to a particular valence(Kurtz *et al.*, 2007).

While uncertainty has been reported to both intensify and dampen affect, it may induce an attentional bias towards negative information given that from an evolutionary point of view, missing a potential negative stimulus could be costlier to an organism than missing a positive one. However, how uncertainty explicitly interacts with valence has not been systematically tested(Anderson *et al.*, 2019) and remains to be further explored.

1.2 Negativity Bias

Negative bias is the phenomenon of negative valence having stronger effects on processes such as effects on attention, perception, memory, decision making and other related behaviour, compared to a positive stimulus of equal strength (in terms of normative values) and arousal. Previous neurophysiological studies suggested that the negativity bias occurs relatively early in the decision making process, driving more attention towards the negative valence in comparison to positive valence(Smith *et al.*, 2003), but can persist in later stages of the decision making process also. It is largely argued that theoretically negativity bias can have evolutionary advantage as avoiding potentially harmful stimuli is more critical(Norris, 2021).

Norris et al(Norris *et al.*, 2010) suggest that the negativity and positivity are separable, that is, they need not be part of a continuum and because of that they do not have to be equal in terms of the way they operate or in terms of the consequences. This is in contrast to the concept of core affect, and the suggestion that they are irreducible. The Evaluative Space Model(ESM)(Cacioppo and Berntson,1994;Cacioppo *et al.*, 1997)(see Figure 1.1) was developed on this idea which further suggests that the separability of positive and negative valence would allow them to have different activation functions, and in turn different outputs. This difference in strengths of response was conceptualized in terms of differential gain, that is, for a unit increase in input, negativity increases at a faster rate in comparison to positivity. Further, the rate of negative decreases slowly in comparison to that of positive, implying that negative evidence could be weighed higher when compared to positive. This could be interpreted as the model weighing survival over opportunities(Norris, 2021).

A rather classical experiment on negativity bias was done by Brown et al (Brown, 1948), which was done in hungry albino mice where they have to run through a short ally to either secure food or to avoid a shock. A harness device was used to record the strength of pull when the rats are stopping at the alley. It was observed that those rats that stopped near food pulled stronger towards the food compared to those rats which stopped away

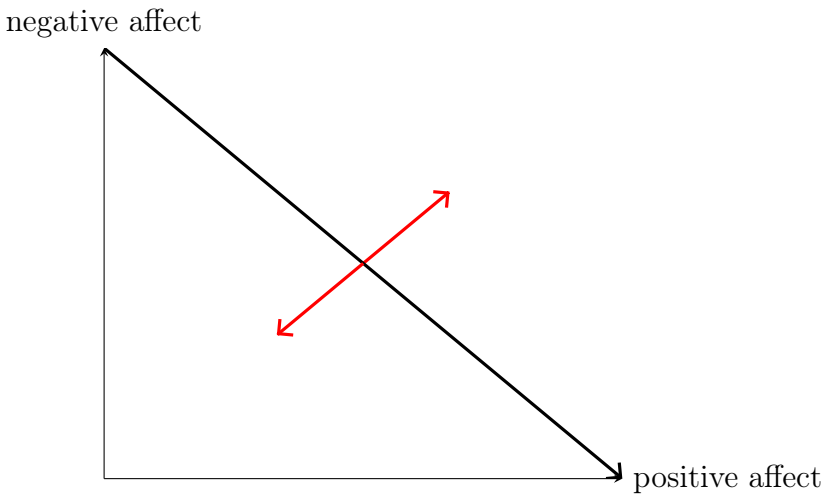


Figure 1.1: *Seperable negative and positive valence as suggested by ESM(Cacioppo and Berntson,1994;Cacioppo* et al.*, 1997). In this, the x axis represents the positive valence and y axis represent the negative valence. If the bipolar valence aaxis exist, that would be represented by the black diagonal line, where one valence is linearly decreasing function of the other. The red arrows describes the direction in which the deviation would happend from this bipolar concept.*

from the food. Similarly, rats who stopped near the shock point pulled away stronger in comparison to those who stopped farther. And when plotting the strength of pulling against the distance from the respective points at which they have stopped, it was observed that there exists a steeper slope for shock when compared to the appetitive stimulus. Hours of food deprivation didn't change the slope but shifted the curve upward or downward. Parsi and Kurtz (Parisi and Katz, 1986) conducted a survey on posthumus organ donation and with the bivariate system found that the positive and negative subscales were uncorrelated, which imply that, it is less likely to have a linear relationship between the two, which would be the case if it were a continuum. Green and Goldfried(Green and Goldfried, 1965) also pointed out that the bipolar scale may constraint the seperable manifestation of the positive and negative scales. Another experiment using the PANAS scale on student population found out changes in negative and positive affect upon receiving exam results, which were uncorrelated (Miller, 1959) again suggesting that the positive and negative components may have non-overlapping operative components. However, imaging studies aiming to untangle the bipolarity vs. bivalent debate about valence via its neural basis have found it to be supported by a distributed set of brain regions that are flexibly involved during positive & negative affective states (Lindquist *et al.*, 2016).

1.3 Effect of outcome valence on learning

Differential effect of positive and negative feedback on learning has been reported in the literature (see Reinen *et al.*, 2021 for a recent review). For example, there has been experiments in which within one learning block stimuli association has to be learnt from negative-blank feedback (that is negative feedback for incorrect responses and no feedback for correct responses) or from a positive-blank feedback (that is positive feedback for correct response and no feedback for incorrect responses) (van de Vijver *et al.*, 2015). It was observed that in some tasks adults performed better in negative-blank but not in the positive-blank feedback, which was not consistent with other tasks(Eppinger *et al.*, 2010;Eppinger and Kray, 2011;Eppinger *et al.*, 2013), where such a differential effect was not found.

There is a wealth of literature suggesting that positive reinforcement is accompanied by positive emotions and vice versa(Sidman and Center, 1991;Starling *et al.*, 2013;Rakos *et al.*, 2008). It is also possible that the affective factors could affect feedback processing(Paul and Pourtois, 2017). And in the same line it was found that the feedback related negativity(FRN) event related potential(ERP) amplitude was higher during positive emotional state tha during negative emotional state(Paul and Pourtois, 2017;Bandyopadhyay *et al.*, 2019), and it is consistent with the fact that the origin of FRN signals, anterior cingulate cortex(ACC)(Allman *et al.*, 2001) has extensive connections to orbito-frontal cortex(OFC)(Carmichael and Price, 1996), amygdala and anterior insula(AI)(Koban *et al.*, 2013) which are regions that are mainly involved in emotion processing.

Xu et al.(Xu *et al.*,2021) specifically investigated the effect of reinforcement type(positive or negative) on FRN and learning biases. They found that FRN responses were more profound following wrong responses than following correct responses. Which can be seen in line with the concept of loss-aversion(Kahneman and Tversky,1984). One idea was that the dopamine system itself may be differentially involved in the case of positive versus negative reinforcement type.

1.4 Limitations so far and motivation for the current study

Arguments that the negativity bias is conferred due to evolutionary advantage is mostly post-hoc, making it hard to test or falsify it. However, it suggests the following.

- Average population behavior would show a strong response to aversive stimuli (compared to appetitive ones)
- Stable individual differences should exist in the population for the natural selection to act on, which would lead to a difference in fitness.
- Negativity bias could have a developmental aspect as it is less observed in infants compared to adults(Verburg *et al.*,2019).

There is also contradicting evidence to such a bias, that is present in the literature, which found null or opposite results, but many of these failed to meet the criterion of equivalency (Norris,2021). That is, the criticism in studies reporting negative bias are also about the calibration and the strength of the negative stimuli in comparison to the positive one. And some studies failed to match the strength of positive and negative stimuli. Because, if we are using stimuli having different strengths in the first place the difference in results could just be attributed to that. Ideally the negative and positive stimuli chosen should be equidistant from the neutral point in terms of valence and should match in all other properties. Lagisz et al(Lagisz *et al.*,2020) pointed out that given that there is heterogeneity present in the studies investigating these biases across different species, one could expect stable individual differences in these traits to be assessed via administration of questionnaires(Kim *et al.*,2003).

In Sir Arthur Eddington's The Philosophy of Physical Science(Broad,1940) there is a hypothetical example of a scientist trying to sample fish populations in the ocean for determining the size of fish. He sampled the oceans using a two inch mesh, repeatedly and exhaustively and ended up concluding that there is no fish in the ocean, having size less than two inches. One important thing to note here is that the scientific instruments or the paradigms themselves are formed by the researcher's implicit assumption of the phenomena to be investigated. Because in this example, the scientist began with the assumption that any fish would be of considerably larger size and began sampling from 2 inches, or the instrument that he used itself had that bias so that nothing below two inches would be sampled. In a similar vein, it is important to examine the role of uncertainty (including ambiguity) which itself is known to be aversive, as a probe for negativity bias. This is because, as discussed above, the integration of emotion and decision making could be seen as a dynamic iterative process that adjusts the internal state of the agent about the environment(Paulus and Angela,2012). For example, Paulus et al.(Paulus and Angela,2012) suggested that in the Bayesian framework these affective states could be conceptualized as a factor that would modulate the representation of the beliefs and therefore, the resulting

computations. However, while phenomenologically emotions have been considered introspective in nature, decision making has been quantified primarily using external variables. Hence, their conceptualizations in computational models need to be considered in light of these theoretical differences (Paulus and Angela,2012) such that emotions may affect the way the probabilities of loss or gain are translated into weights or values. For example, Andersen et al.(Anderson *et al.*,2019), proposed that uncertainty would affect effect via mental simulations such that uncertainty is no longer a feature of the objective world (as it has been typically examined in the learning & decision making literature), but something integral to an individual's subjective experience. However, if it is expected uncertainty, the difference need not show up as a change in learning rate since as we can't be sure about it is changing the perception about change of probability as well, in comparison to unexpected uncertainty(Angela and Dayan,2005;Behrens *et al.*,2007;Nassar *et al.*,2010). And in a constant environment, the optimal learning strategy would be rapid updation at the beginning and then decaying the learning rate as an inverse function of number of samples(Doya,2008), thus it is possible and we expect that, even when there is a difference initially over learning blocks, because of this inverse function nature, the difference may die out, which is also possible even when there are differences in expected uncertainty. And in previous studies, that looked at the changes in learning rate in response to changes in expected uncertainty or valence, we coudln't find a result suggesting that the learning rate would be a good candidate for quantifying either valence or changes in expected uncertainty(Nassar *et al.*,2012;Duff *et al.*,2013).

Moreover, learning rate is only one parameter of interest, and response time which could potentially explain factors such as conflict involved in the decision process, has been kept outside of the traditional RL models(Pedersen *et al.*,2017), we propose that including it could provide new insights to understanding negativity bias in learning. Specifically, by keeping the value or monetary outcome associated with the feedback constant, we expect to see a valence based difference in learning in the mixed vs. symbolic block, that is hypothesized to be captured by parameters in the more sensitive RL-DDM model.

Secondly, if the bias doesn't depends on positive and negative getting presented together in the same block, we expect to see the same amount of bias reproduced in the positive and negative feedback block, if presenting the, where only one valence is used as feedback. Thirdly, we expect the uncertainty to have a positive modulatory effect on bias such that when the outcome is more certain less bias would be observed. Lastly, we would expect the trait scores to explain the observed bias along with the RL-DDM parameter coding for it.

Chapter 2

Models And Theory

2.1 Reinforcement Learning- Drift Diffusion Model

RL-DDM was presented in 2017 by Pedersen and Frank(Pedersen *et al.*,2017)(and implemented in HDDM package(Wiecki *et al.*,2013)) to combine sequential sampling models with the reinforcement learning framework, which would enable characterizing much richer dynamics of the learning process. It is implemented in a hierarchical framework, where group and individual-level parameters are mutually constrained. In addition, Bayesian parameter estimation allows parameter estimation using a relatively more minor data set.

In the RL model the value for a given choice is given by

$$V_i(t) = V_i(t-1) + \alpha[R_i(t-1) - V_i(t-1)] \tag{2.1}$$

$V_i(t)$ is the value of option i at time t, and $R_i(t)$ is the feedback about the option i, that they have recieved at time point t. α is the learning rate. The probability that the decision maker will choose option i will be governed by a softmax function, given by:

$$p_i(t) = \frac{e^{(\beta(t)V_i(t))}}{\sum_{j=1}^{n} e^{(\beta(t)V_j(t))}} \tag{2.2}$$

When we have n alternate options. But, in our case, it is just two options. Parameter β depicts the sensitivity or exploration-exploitation trade-off.

Now if we consider the drift-diffusion model, it is a noisy accumulation of evidence, which can be implemented using a Weiner first pass in time function(WFPT), the likelihood of response time of choice i will be given by

$$RT(i) = WFPT(a, z, ndt, v) \tag{2.3}$$

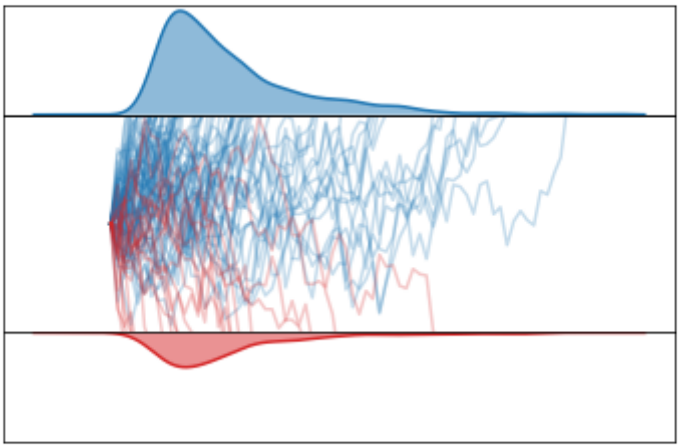


Figure 2.1: *Depiction of the Diffusion-Decision model(DDM). Integration of evidance is represented by the traces, which are stochastically moving towards either of the decision boundaries. It starts from a point in between the boundary, specified by prior bias(z). In the picture, starting point is the midpoint as z=0.5. Once these traces crosses either of the boundary, that choice is considered to be the decision and the probability distributions of response time for the two choices are plotted on each of the boundaries. Red traces denote the decision processes, which crossed the lower boundary and blue traces denote that crossed the upper boundary.*

Where a is the separation between the two boundaries, z is the relative starting point with respect to the two boundaries(prior bias), and t is the non-decision time, which consists of encoding and motor time. v is the drift rate, ndt is the non-decision time. In the RL-DDM model drift rate was presented as a function of time, and it gets updated as shown below:

$$v(t) = m[V_{upper}(t) - V_{lower}(t)] \tag{2.4}$$

Values of upper and lower functions were updated according to the value updation rule of RL. m is a scaling parameter. The same paper empirically showed that the RL-DDM model better captured the data than the RL model since it could also incorporate response time.

2.1.1 Hierarchical Structure of the Model

The model's hierarchical structure estimates parameters at the group and individual levels, which mutually restricts the estimation process. The following distributions would give the distributions of parameters at the group level.

$$a_{\mu} = G(1.5, 0.75) \tag{2.5}$$

$$a_\sigma = HN(0.1) \tag{2.6}$$

$$v_\mu = N(2, 3) \tag{2.7}$$

$$v_\sigma = HN(2) \tag{2.8}$$

$$z_\mu = N(0.5, 0.5) \tag{2.9}$$

$$z_\sigma = HN(0.05) \tag{2.10}$$

$$t_\mu = G(0.4, 0.2) \tag{2.11}$$

$$t_\sigma = HN(1) \tag{2.12}$$

$$\alpha_\mu = N(0.5, 0.5) \tag{2.13}$$

$$\alpha_\sigma = HN(0.05) \tag{2.14}$$

N is normal distribution, G is Gamma distribution, and HN is half-normal distribution. Given parameter distributions at the group level and at the individual level, the distributions are given by:

$$a = G(a_\mu, a_\sigma^2) \tag{2.15}$$

$$v = N(v_\mu, v_\sigma^2) \tag{2.16}$$

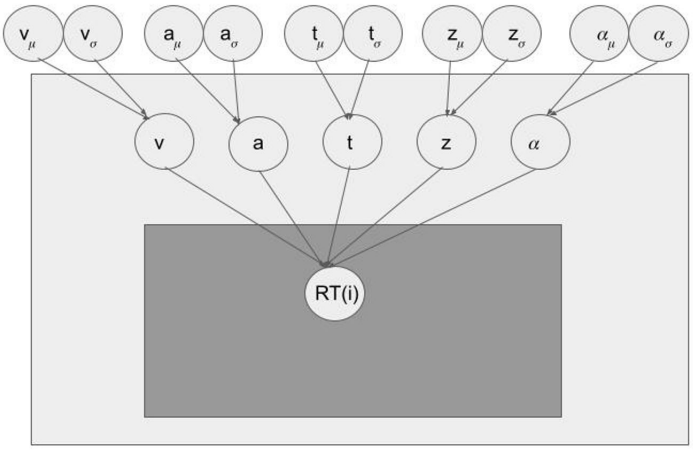


Figure 2.2: *The hierarchical structure of the RL-DDM model. The outer white region depicts the group level, and at this group level, we have parameters describing the mean and standard deviation of parameters, which gives the parameters at the individual level(light grey region). Ad the parameters at the individual level, ultimately determine the likelihood for each RT and response combination*

$$z = invlogit(N(z_\mu, z_\sigma^2)) \tag{2.17}$$

$$\alpha = invlogit(N(\alpha_\mu, \alpha_\sigma^2)) \tag{2.18}$$

$$t = N(t_\mu, t_\sigma^2) \tag{2.19}$$

(For pictorial representation of the heirarchical structure, see figure 2.2)

2.2 Win-Stay and Lose-Switch Behaviour- A new coordinate system

We wanted to see how participants' decisions change after encountering a win and a loss. $p(stay_{t+1}|win_t)$ is the probability of choosing the same op-

tion after getting matching feedback for the same condition in the previous trial. $p(Switch_{t+1}|lose_t)$ is the probability of .Switching given a mismatch between response and feedback in the previous trial. This is a heuristic way of looking at the learning(Worthy *et al.*,2013), and several studies(Mellers *et al.*,1999;Johnson and Tversky,1983;Sokol-Hessner *et al.*,2013) had been shown that the negative valence would induce more conservative decision making strategies, in which case behavior such as win-stay would be a potential means to explain the bias that we are observing.

Mathematically, the probability of staying with the same choice after getting a win in the previous trial is given by:

$$p(stay_{t+1}|win_t) = \frac{p(stay_{t+1} \cap win_t)}{p(win_t)} \tag{2.20}$$

If the events are independent, that is, if the probability of staying with the same option is independent of that of the probability of a win in the previous trial, the expression can be written as:

$$p(stay_{t+1} \cap win_t) = p(stay_{t+1}) \cdot p(win_t) \tag{2.21}$$

Then the conditional probability would become:

$$p(stay_{t+1}|win_t) = \frac{p(stay_{t+1} \cap win_t)}{p(win_t)} = \frac{p(stay_{t+1}) \cdot p(win_t)}{p(win_t)} = p(stay_{t+1}) \tag{2.22}$$

If we divide by the probability of staying with the same option, we will get, let's call this, the "win-stay factor" and denote it as WS.

$$WS = \frac{p(stay_{t+1}|win_t)}{p(stay_{t+1})} = \frac{p(stay_{t+1} \cap win_t)}{p(win_t) \cdot p(stay_{t+1})} \tag{2.23}$$

This factor can have values in two possible regions, and it will be precisely one if the events are independent. Possibilities are as follows:

- If staying is preferred more after a win(maybe we can call this as an exploitation strategy), more number of staying with the same option will occur after the win:

$$p(stay_{t+1} \cap win_t) > p(win_t) \cdot p(stay_{t+1}) \implies \frac{p(stay_{t+1} \cap win_t)}{p(win_t) \cdot p(stay_{t+1})} = WS > 1 \tag{2.24}$$

- If staying is less preferred after a win:

$$p(stay_{t+1} \cap win_t) < p(win_t) \cdot p(stay_{t+1}) \implies \frac{p(stay_{t+1} \cap win_t)}{p(win_t) \cdot p(stay_{t+1})} = WS < 1 \tag{2.25}$$

- If the events are independent:

$$p(stay_{t+1} \cap win_t) = p(win_t) \cdot p(stay_{t+1}) \implies \frac{p(stay_{t+1} \cap win_t)}{p(win_t) \cdot p(stay_{t+1})} = WS = 1 \tag{2.26}$$

Our conditions are based on what the symbols are associated with, whether pleasant or unpleasant. Let's now consider WS as a function of condition c; by estimating the probabilities separately in each condition, the probability of staying given win in the previous trial and given condition is given by

$$WS(c) = \frac{p(stay_{t+1}|win_t, c)}{p(stay_{t+1}, c)} = \frac{p(stay_{t+1} \cap (win_t \cap c))}{p(stay_{t+1} \cap c) \cdot p(win_t \cap c)} = \frac{p((stay_{t+1} \cap c) \cap (win_t \cap c))}{p(stay_{t+1} \cap c) \cdot p(win_t \cap c)} \tag{2.27}$$

Now we can define a plane whose axes are $WS(c_1)$ or $WS(positive)$ and $WS(c_2)$ or $WS(negative)$(see Figure 2.3). If the condition has no effect, we are likely to see all subject's values distributed along the diagonal if there is no bias in one condition compared to the other. The plane would tell us about the relative exploitatory behavior of the participants.

But, this variable can range from zero to some powers of 10, which we can approximate to a normal distribution, or zero centered distribution, by applying log, which results in a new plane(see Figure 2.4).

$$LWS(c) = log(WS(c)) \tag{2.28}$$

If we are applying a 45-degree rotation in the clockwise direction, then the y-axis would represent the bias, and x would suggest learning or adaptation. For this, we need a transformation matrix, which is given by

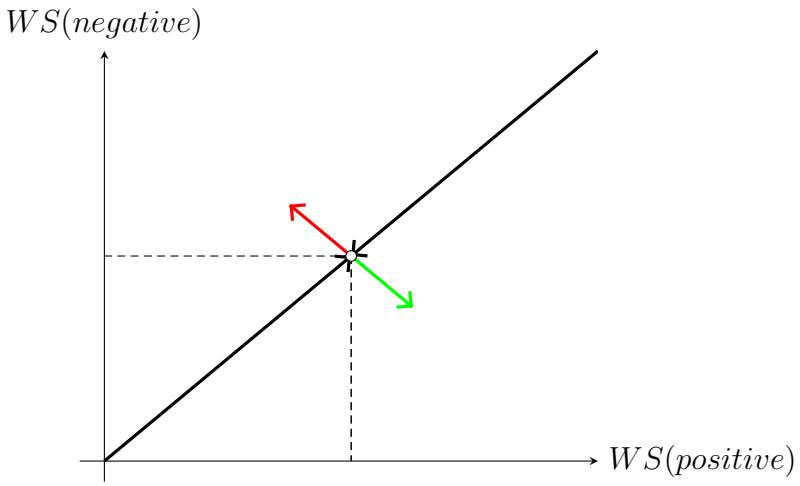


Figure 2.3: *Pictorial representation of the WS(positive)-WS(negative) plane. Here, individual WS being greater than one denotes the probability of staying getting enhanced after a win. The dotted vertical line represents the point where the value of WS(positive) will become equal to one, where the events are independent. On the right side of it, the win would enhance the probability of staying after a win in positive condition would get enhanced, and on the left side, it will be the other way around. The dotted horizontal line denotes the point where the event of staying with the same choice in the subsequent trial is independent of the event of getting a win in the current trial in a negative condition. Above that, the probability of staying with the same choice given a win in the previous trial is enhanced, and below that horizontal like, it is the other way around. The white dot, having coordinate-(1,1), denotes an "indifference" point in this coordinate, where all stay events and win events are independent and equal between the two conditions. On the diagonal line, only the effect of a win on the decision to stay with the same choice would be varied without any bias. The red line denotes a perpendicular displacement from the diagonal line above it, which shows a negative bias as it would increase the effect of the win on the decision to stay in the negative condition while decreasing that in the positive. The green arrow shows the opposite effect*

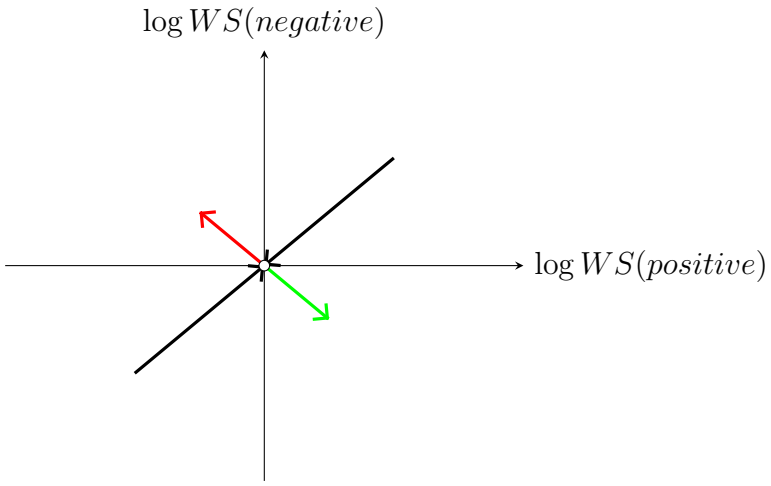


Figure 2.4: *Pictorial representation of the* $\log WS(positive)$-$\log WS(negative)$ *plane. Here, the origin denotes the point where, on each condition(positive and negative), the event of staying with the same choice in the next trial is independent of getting a win(correct) in the current trial. On the x-axis, as we go from left to right, the win in a trial would enhance the decision to stay with the same choice in a positive condition. The same is true when we go from bottom to top in the y-axis in the case of a negative condition. A shift along the red arrow would suggest an increase in the enhancement a win has on the probability of staying with the same choice in the next trial in a negative comparison compared to a positive one, hence showing a negative bias. On the other hand, if there is a shift along the green arrow, it shows that there is a bias in the opposite direction(positive bias)*

$$Tr = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix} \tag{2.29}$$

When the rotation is by 45 degrees clockwise, the values would be

$$Tr = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{2.30}$$

Now, the transformed coordinates would be given by a cross product, after multiplying with square root of 2

$$(L\bar{W}S_1, L\bar{W}S_2) = (LWS(c_1), LWS(c_2)) \times \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{2.31}$$

$$(L\bar{W}S_1, L\bar{W}S_2) = (LWS(c_1) + LWS(c_2), LWS(c_2) - LWS(c_1)) \tag{2.32}$$

In this new coordinate(see Figure 2.5), the first coordinate tells us the impact of a win on their stay probability, regardless of the condition. If the value is zero, even if getting a match in the previous trial has nothing to do with the probability of staying with the same option in the current trial. If it is less than zero, getting a match has a detrimental effect on the probability of staying, and if it's positive, it promotes the probability of staying. We can consider this property, given that it is the geometric mean of the two variables. The second coordinate tells us about the bias. If it is zero, there is no bias towards either condition regarding staying, given that the previous trial had matching feedback. But, if it's higher than zero, that means the participant had a higher probability of sticking to the same option in condition c_2(negative) compared to c_1(positive). Using this 2D planar depiction, we can better explain the behavior. This is because what the second coordinate is giving us is the log of ratios between two factors, when it is greater than one and its log is greater than zero implies one factor is higher than that of the other.

Now we need to show the rationale for calling the new axes as learning or adaptation and bias learning and bias. Let's consider the first coordinate:

$$\frac{LWS(c_1) + LWS(c_2)}{2} = \log(\sqrt{WS(c_1)} \cdot \sqrt{WS(c_2)}) = \log(\sqrt{WS(c_1) \cdot WS(c_2)}) \tag{2.33}$$

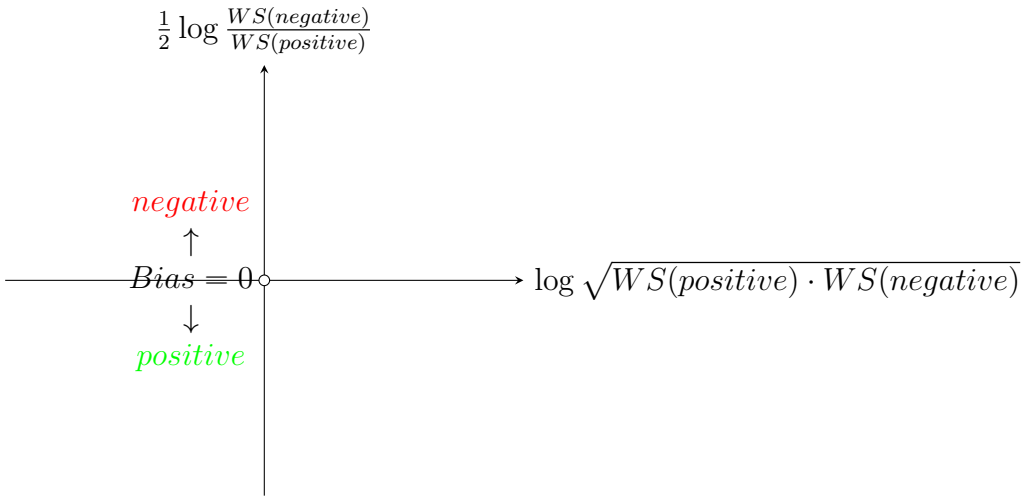


Figure 2.5: *Pictorial representation of the rotated* $\log WS(positive)$-$\log WS(negative)$ *plane, by 45 degrees. Here, the origin denotes the point where, on each condition(positive and negative), the event of staying with the same choice in the next trial is independent of getting a win(correct) in the current trial. On the x-axis, as we go from left to right, the win in a trial would enhance the decision to stay with the same choice across conditions as it is the log of the geometric means of WS(positive) and WS(negative). But, along the y-axis, if we go from bottom to top, the ratio between WS(negative) and WS(positive) increases, and it becomes one at the origin, meaning that the behavior is the same between the two conditions. Below the x-axis is the region showing positive bias, and above the x-axis is the region showing negative bias.*

Which is log of geometric mean across conditions. And the second coordinate:

$$\frac{LWS(c_2) - LWS(c_1)}{2} = \log \frac{\sqrt{WS(c_2)}}{\sqrt{WS(c_1)}} = \frac{1}{2} \log(\frac{WS(c_2)}{WS(c_1)}) \tag{2.34}$$

The same was repeated for the lose-switch behavior, but we used the probability of switching there, given a loss in the previous trial. Thus, we expect this new coordinate system to capture the bias at the population level when individuals are represented in the new coordinates.

Chapter 3

Materials and methods

We used a modified version of a standard cue association learning task(Frank *et al.*,2004) to investigate the role of outcome valence on learning. In this paradigm, the participants were required to learn the association of two Japanese alphabets with either a positive or negative valence in a session. In each trial, a given Japanese alphabet was presented along with two symbols or shapes: a square & a triangle indicating a positive & a negative outcome, respectively. Participants had to choose between the two shapes to indicate which valence is a particular Japanese alphabet likely to be associated with. Feedback was presented at the end of the trial in the form of a positively or negatively rated image correctly associated with the corresponding Japanese alphabet. Participants were asked to choose the target (square for pleasant and triangle for unpleasant or square for square and triangle for triangle that is most likely to be associated with that particular Japanese character). Selection of images for being used as feedback was done via a standardization procedure as described later.

3.1 Participants

We collected data from 36 participants (age = 22.89 $\pm$ 0.650), 26 males, 7 females) on the first learning task, based on the power analysis described below. All participants were students in the Indian Institute of Technology Kanpur, and gave their informed consent in accordance with the Institutional Ethics Committee of IIT Kanpur and Declaration of Helsinki(Goodyear *et al.*,2007). Participants were monetarily compensated for their participation in the study. For the 80% correct feedback blocks, data were collected from 30 participants(20.93 $\pm$ 0.398), 18 males, 12 females).

3.2 Standardizing The Stimuli

The primary criteria for the images to be selected as feedback stimuli was that they should be matched for arousal, as well as be equidistant from the neutral point on the valence axis to ensure that both positive and negative valence are compared when presented at equal strength . To get standardized rating data of non-social, emotional images, we first picked 60 numbers of arousal matched images with no human body parts in them, from the following emotional picture datasets: IAPS(Lang *et al.*,1997a), DIRTI(Haberkamp *et al.*,2017), OASIS(Kurdi *et al.*,2017). Rating data was collected for each image on three dimensions: valence, arousal and motivation using a 9 point Likert scale where 1 and 9 corresponded most unpleasant/least arousing/least motivating and most pleasant/most arousing/most motivating respectively. Images were presented on the screen with a black background with size 400×500 pixels for four seconds while participants were asked to look at it. After that it was presented with a size of 240×300 pixels with the rating scales below it, in black background. Responses were collected online using cognition run platform (`https://www.cognition.run/`) and jsPsych version 6.3.1(`https://www.jspsych.org/6.3/`) from a sample of 50 participants (recommended for high rating reliability; DeBruine and Jones,2018).

Data was analyzed with and without including raters whose responses in the PHQ-9(Kroenke and Spitzer,2002) and GAD-7(Williams,2014) questionnaires fell in the severe or moderately severe anxiety or depression categories (specifically those scoring >11 in either scale were removed, where we had to remove 33 participants out of 83, and remaining analysis was done on 50 participants), since individuals with anxiety and depression have been reported to differ in their rating of emotional stimuli(Gray *et al.*,2021). Finally, six images were shortlisted (three from each valence category) matching for mean arousal values (5.4714) and distance from neutral on the valence axis of the mean valence ratings (mean for pleasant images: 7.5143, mean for unpleasant images: 2.5143). In the remaining sample of raters (32 were male, and (64.0%)(binomial test for equal proportion, p value =0.0649), age (mean $\pm$ std) = 22 $\pm$3.7 yrs) Exact instructions for respective ratings are as follows:

> ***Valence Rating:*** *"You will be asked to rate how each picture makes you feel on a scale from unpleasant to pleasant (bad to good). If you feel completely unpleasant (bad) while viewing the picture, please click the left end of the scale. If you feel completely pleasant (good) while viewing the picture, please click the right end of the scale. If you feel completely neutral while viewing the*

picture, you will be required to select neutral on scale which is 5. You are rating how you feel in reaction to the pictures. The slider also allows you to indicate intermediate feelings of the strength of your urge to avoid (move away from) or approach (move towards), by choosing an intermediate value."

Arousal Rating: *"You will be asked to rate how each picture makes you feel on a scale from calm to aroused. If you feel completely calm (sleepy, relaxed, soothed) while viewing the picture, please select the left end of the scale. If you feel completely aroused (awake, excited, agitated, jittery) while viewing the picture, please select the right end of the scale. You are rating how you feel in reaction to the pictures. The slider also allows you to indicate intermediate feelings of the strength of your urge to avoid (move away from) or approach (move towards), by choosing an intermediate value."*

Motivation Rating: *"You will be rating the strength of your urge to avoid (move away from) or approach (move towards) each picture on a scale from avoid to approach. If you feel that the strength of your urge to avoid (move away from) is strong while viewing the picture, please select values at the left end of the scale. If you feel that the strength of your urge to approach (move towards) is strong while viewing the picture, please select values at the right end of the scale. If you feel absolutely no urge to avoid or approach the picture, then please select neutral on the scale by clicking at 5: The slider also allows you to indicate intermediate feelings of the strength of your urge to avoid (move away from) or approach (move towards), by choosing an intermediate value. You are rating how you feel in reaction to the pictures."*

3.3 Reinforcement Learning Task

The experimental task was designed in NIMH MonkeyLogic(`https://monkeylogic.nimh.nih.gov/`). Participant were seated at a distance of 73cm from a 52inch LCD monitor, and used a chin rest for keeping their face steady throughout the task. Eye positions were calibrated using a five point calibration routine and recorded during the task using an infrared camera from SR Research Eyelink 1000 Plus(`https://www.sr-research.com/eyelink-1000-plus/`). Bioplux EDA sensors were attached to the tip of index and middle fingers of the non-dominant hand (`https://www.pluxbiosignals.com/products/`

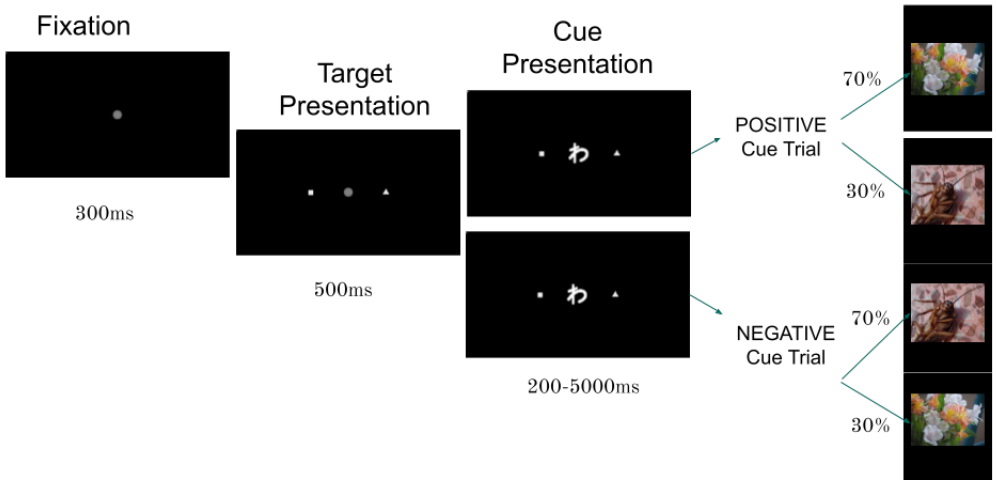


Figure 3.1: *Diagram of the reinforcement learning task- mixed valence feedback block*

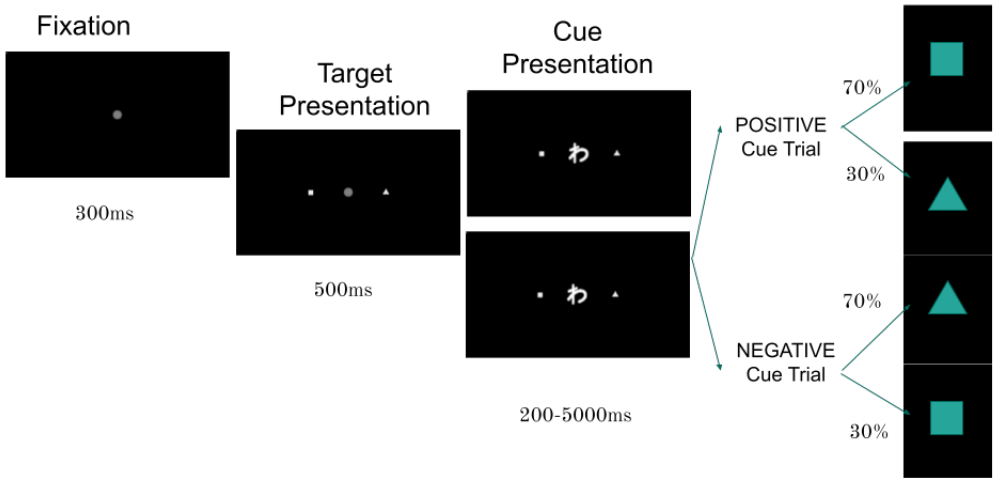


Figure 3.2: *Diagram of the reinforcement learning task- symbolic feedback block*

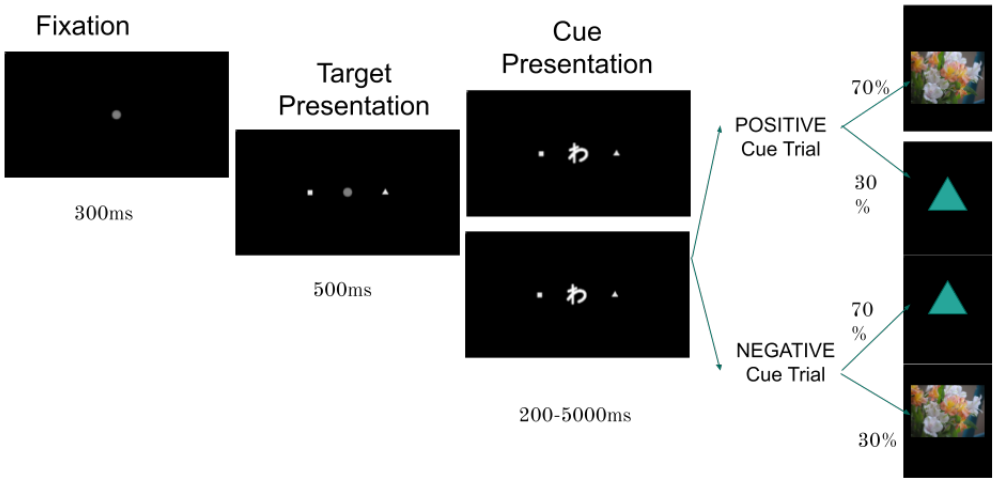


Figure 3.3: *Diagram of the reinforcement learning task- positive feedback block*

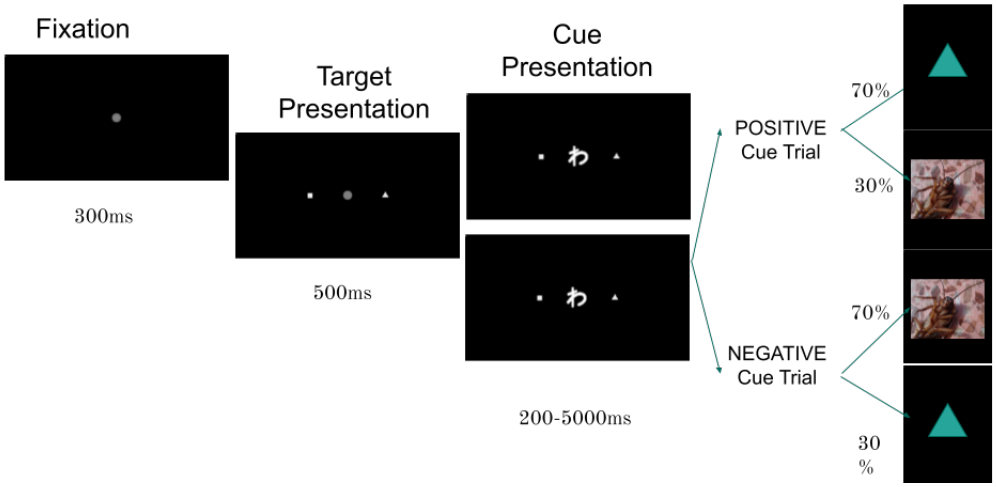


Figure 3.4: *Diagram of the reinforcement learning task- negative feedback block*

`electrodermal-activity-eda-sensor-1`) to acquire galvanic skin response(GSR) data wirelessly using Bluetooth 4.0 technology while data was streamed through Lab Streaming Layer(LSL, `https://labstreaminglayer.org/#/`) to the MonkeyLogic. Participants were then asked to rate the six images used in the experiment on three dimensions: valence, arousal and motivation as described above for stimuli standardization.

Before each learning block, participants were provided with a practise block, which was intended to provide better understanding about performing the task and was terminated only when the participant reported that they are confident to proceed with the task, on average, they took 20 trials in practise, however more trials were given if they are not clear about the paradigm.

In each block, each trial started with a fixation circle, which would be blinking, and the participant is required to look at it >=300ms for beginning the task, and two targets (triangle and square) appeared after. Upon fixation, fixation circle was presented for 500ms, and then the Japanese alphabet was shown. Participants were required to respond as soon as possible. Trials with responses later than 5000ms ± 250ms were aborted. Experiment consisted of a conditions file in which 20 items were listed. Items 1-7 and 11-17 were provided with the correct feedback type and for items 7-10 and 17-20 incorrect feedback was provided by inverting the feedback. These items were provided randomly without repitition This ensured that an equal number of both trial types were presented. only attempted trials were counted, upon an incomplete attempt the trial was repeated at a delayed time point to avoid same symbol being repeated consecutively , in which case the participant could potentially use that time to make decisions. There was an intertrial interval of 2000ms with 5% jitter. Experiment terminated upon completing 100 trials for each participant which lasted around for 17mins per block.

In the first block [called as the symbolic block or SYM block], the outcomes or feedback image was that of the square or triangle, which were luminosity matched and the participants were asked to predict whether the Japanese alphabet was associated with a square or a triangle. In the MIXED VALENCE block [VAL], the outcome was any of the six shortlisted valenced images presented equally randomly in a block. To assess the impact of pleasant and unpleasant outcomes independently, In third and fourth blocks, the outcome was either square/pleasant image or unpleasant image/triangle. The symbols were intended to serve the role of a 'no valenced' feedback. To test if the task is inducing an affective response, we collected responses on the PANAS(positive and negative affect schedule)(Watson *et al.*,1988) questionnaire after the 3rd and 4th blocks. And there was a break of 7-10mins between blocks. Each block is also divided into 20 trial bins for analysis.

Pain Belief questionnaire(Edwards *et al.*,1992), pain sensitivity question-

naire(Ruscheweyh *et al.*,2009), Disgust sensitivity scale(Olatunji *et al.*,2007;van Overveld *et al.*,2011), Beck's Depression Inventory(Beck *et al.*,1961;Beck *et al.*,1988;Hojat *et al.*,1986;Steer *et al.*,2000), State and Trait Anxiety Inventory(Spielberger *et al.*,1971;Spielberger *et al.*,1983) questions, and Pain Catastrophizing Scale (PCS)(McNeil and Rainwater,1998;Osman *et al.*,1997;Sullivan *et al.*,2001) were implemented for each participant after the task.

3.4 Experiment 2

Experiment 2 was of the same as experiment 1, except that the probability of receiving accurate feedback at the end of the trial was changed from 70% to 80%. This consisted of only first two blocks and was a single session experiment. This experiment was done on an independent set of 30 participants mainly to compare between the block 1 and 2.

3.5 Behavioural Pilot

We collected pilot data for valenced feedback blocks for 70-30 probability condition and 80-20 probability condition. Pilot sample was relatively small, N=10 (5 in 70-30 and 5 in 80-20, and these participants were not included in the study). However, it was found that there exist a difference between 70% correct feedback and 80% correct feedback in terms of parameter z(prior bias) of the RL-DDM model(see section 2.1 for detailed explanation of the model), but, not in terms of learning rate. And no other parameters shown a credible difference. Overall bias towards the correct association was higher in 80% correct outcome blocks, compared to 70% correct outcome blocks as expected.

Model has a heirarchial structure for extimating the group level and individual level parameter distributions(see section 2.1.1 for details). Based on this, we used the group level posterior distributions of the best model to simulate the data sets, with different numbers of subjects. We drew samples from the group level distributions and constructed individual level distributions as explained in section 2.1.1, resulting in a varying number of subjects. Repeated this procedure for N number of subjects 30 times and found the correlation between simulated and recovered parameters. We found that for N=30 and above all correlations were above r=0.75($p<0.001$),. Hence we decided to collect N=40, in anticipation of some exclusions.

3.6 Data Analysis

3.6.1 Exclusion Criteria For RL experiments

We excluded participants if they are either having too few incorrect responses, or if they are choosing the same options regardless of the true underlying probability(ie, choosing pleasant/ unpleasant always) or if they are consistently responding within 200ms for a majority of trials.

3.6.2 Determining the number of bins for analysis of the RL task data

From the five blocks we estimated a performance threshold threshold(fraction of responses, which are showing the true association) such that it won't result in too few bins of trials, in which case participants would be able to cross that by chance or won't result in participants talking up almost the entire learning block for a majority of the trials, which is also less than the proportion of correct feedback given. Thus for 70% correct feedback blocks, we set 65% as the threshold at the bin level(data till first bin of 20 trials that reaches this threshold will be taken separately for each participant), and for 80% correct feedback blocks, we set 75% as the threshold at bin level.

3.6.3 Simulation Based Sample Size Calculation

Main idea behind the power analysis was to get a number of participants to reliably get the separation in parameter estimates.To capture the true parameters underlying the behavior to explain negativity bias, we used the pilot data from the MIXED or VAL block of the 70-30 RL task (n= 5 individuals) to simulate 30 datasets of varying sample sizes and tested how the number of participants improved the correlation between actual parameters used to simulate the data and the recovered parameters. We relied on the hierarchical structure of RL-DDM to do this(see Figure 2.2 and section 2.1.1, for more details).

We used the group level posterior distributions of the best model to simulate the data sets, with different numbers of subjects. We drew samples from the group level distributions and constructed individual level distributions as above resulting in a varying number of subjects. Repeated this procedure for N number of subjects 30 times and found the correlation between simulated

and recovered parameters. We found that for N=30 and above all correlations were above r=0.75($p<0.001$),. Hence we decided to collect N=40, in anticipation of some exclusions.

3.6.4 Response Time Analysis

We analyzed how response time after a win and after a loss changes depending on condition, block and type of feedback. This is to see whether any is getting reflected in the response time. As RT follows an ex-Gaussian distribution, we log transformed it before applying any linear models. Our attempt to analyse RT is mainly motivated by previous studies. In some of the previous studies, such reliable differences were reported for diffence in valence(Simpson *et al.*,2000;Estes and Verges,2008;Fox *et al.*,2001;Algom *et al.*,2004). Hence we are expecting the negative valence to increase the response time given our task context. We implemented a forward selection method for selecting potential variables that would explain the data using Akaike Information Criteria(AIC)(Akaike,1974;Wagenmakers and Farrell,2004;Konishi and Kitagawa,2008), for explaining the effects. A decrease in the AIC score would suggest that the model is improving in the forward modeling scheme. Forward modeling was implemented to explain the effect using the best model, as otherwise, when including all factors, without searching for a potential best model, would inflate the standard error(refer Gelman and Hill,2006, chapters: 3, 10, 12) for coefficients, which would reduce the effects, if any is present in the data. Thus, it is important to explain the effects using the best model.

3.6.5 Win-Stay and Lose-Switch behaviour analysis using GLMs

We fitted separate models for win-stay and lose switch behaviour. In the case of the win-stay model, we selected trials for which the previous trial was a win and then coded the response as 1 if they are deciding to stay and zero otherwise. Negative valence is shown to promote more conservative decision making strategies(Mellers *et al.*,1999;Johnson and Tversky,1983;Sokol-Hessner *et al.*,2013), and win-stay behaviour could be one such measure, that can potentially quantify the bias. For this, we implemented a forward selection method for selecting potential variables that would explain the data using Akaike Information Criteria(AIC). Factors included in this modeling scheme are condition(whether the Japanese symbol is associated with positive or negative outcome), type for feedback block(whether it is a mixed("VAL"),

positive("POS"), negative("NEG") or symbolic("SYM"-control)), bin number of trials(which also represent time in a session).

For lose-switch behaviour, the way in which the response was coded was the only difference. We picked trials after a lose and coded the response as 1 if they decided to switch and as 0 otherwise. Model was fitted using lme4 package in R(Bates *et al.*,2015).

3.6.6 Quantifying Bias in the New Coordinate System

Our attempt was to show the existence of bias using the new coordinate system we defined(see Section: 2.2) We were attempting to compare different types of feedback loops against the symbolic block, which is the control. We are interested in seeing whether the means at the population level are showing a bias. For this we implemented a bootstrapping method to assess the bias. Our null hypothesis was that, there is no negative bias, meaning that the bias can either be zero or positive. In terms of the bias dimension that we defined, the null hypothesis can be stated as follows:

- H_0: there is no negative bias, difference between means for mixed block and control(symbolic) block would be either less than or equal to zero
- H_A: there is a negative bias, the difference between means would be greater than zero.

For testing this, we sampled the individual coordinates computed from the behavioural data, 1000 times with replacement, and computed the difference between means(between VAL block and SYM block). This created a distribution of such differences, and we checked how often it would be less than or equal to zero, in support of the null hypothesis. If this probability is less than the significance level of 0.05, we will reject the null hypothesis, and would consider that there is a bias that is observable from the sample. Same is repeated for negative and positive feedback blocks. And was also repeated for adaptation/learning variables from the new coordinate system.

3.6.7 RL-DDM analysis

We tried to fit RL-DDM(Pedersen *et al.*,2017) to the behavioural data, and tried to see in what all parameters, we need to fit it separately for positive and negative conditions, in a forward modeling scheme using Deviance Information Criteria(DIC)(David *et al.*,2002;Gelman *et al.*,1995, Ch: 7, 12, 14),

using data from the mixed block. We found that when fitting parameters separately for scaling factor(m), prior bias(z), and learning rate(α) of the RL-DDM model, the best possible model is obtained. Model was fitted separately for each type of feedback block and then within each of those models the parameters are compared, using the posterior distribution. The hull and alternate hypothesis are stated below.

- H_0: There is no negative bias in terms of the parameter, $p(negative \leq positive)$ is the probability in support of this null hypothesis.
- H_A: There is a negative bias in terms of the parameter, $p(negative \leq positive) \leq \alpha^{'}$, where $\alpha^{'}$ is the alpha corrected for multiple comparisons($\alpha^{'} = \alpha/4$,$\alpha = 0.05$, based on Bonferroni correction(Hochberg,1988)

And if this null hypothesis is getting rejected for any of the parameters, that suggests, the parameter is capturing the bias observed. Importantly, we are not expecting the null hypothesis to get rejected in the case of learning rate based on previous studies(see Appendix section 6.3 for details of model fitting and evaluation).

3.6.8 Checking Explainability of behaviour using trait

We are also interested in the explainability of shift in prior bias variable using the traits. This is because, we are not expecting the bias to be a phenomenon, which is present at the individual level, but, we are expecting it at the population level. Hence, we are expecting the traits to explain it. Hence, our first attempt was to fit an RL-DDM regression model, separately for each individual, where the prior bias(z) is regressed against conditions, which are category coded as 1 for negative and zero for positive. And the mean of posterior distributions of coefficient from each individual model is regressed against their standardized trait scores, after running a forward modeling scheme to see the best model that would explain the data.

Initially we tried to see how each individuals across dimensions of the new coordinates(see section 2.2) defined would be explained using their trait. This would help us to make more ecologically valid claims about such behavioral biases. Here also we implemented a forward modeling scheme to explain the effects using the best model.

3.6.9 Effect of uncertainty

We repeated the same set of analysis, but this time only using mixed("VAL") and symbolic("SYM"-control) blocks of 70%and 80% correct feedback blocks.

The uncertainty was category coded using the percentage of incorrect trials(ie, as "$prob - incorr[20p]$" for 80% correct feedback blocks and "$prob - incorr[30p]$" for 70% correct feedback blocks). In RT analysis and win-stay and lose-switch behaviour analysis, this extra variable($\imath prob - incorr\jmath$) was included in the forward model selection scheme.

For quantifying bias in the new coordinate system, the same procedure is repeated, but, this time between the mixed and control block of 70% correct feedback blocks and 80% correct feedback blocks, to see whether the difference is varying in the later. For the RL-DDM model also we mainly compared between the mixed blocks of 70% and 80% correct feedback blocks(here alpha correction is applied by dividing alpha by 2, as the data sets for different uncertainty conditions are independent). For the trait level analysis we were interested in how the probability is interacting with the trait-behaviour relationship. To quantify this, we repeated the same forward modelling scheme, this time, included variable coding for different uncertainty. In this we checked how the relation between the new coordinates defined and traits would vary with uncertainty.

Chapter 4

Results

4.1 Response Time Analysis

To examine the behavioural responses to valenced feedback we first looked at reaction times across the four block types which differed in the way valence was introduced as part of the performance feedback: mixed valence, positive valence, negative valence or no valence block. In previous work, tasks where valence information was not relevant for optimal performance, successful trials have been associated with slower response times in negative vs. positive valence condition,(Simpson *et al.*,2000;Estes and Verges,2008;Fox *et al.*,2001;Algom *et al.*,2004). We ran a generalized linear mixed model through the forward model selection procedure using AIC as a criterion. The following factors were considered for the selection of the best model: type of feedback block (mixed, positive, negative or no valence/ symbolic), trial condition (pleasant or unpleasant probable outcome, depending on type of feedback block), temporal bin number in the session. Based on past literature, we expected to see a valence based modulation in reaction times such that RTs will be longer following a negative vs. positive valence feedback. In our task, a similar effect could be expected as a valence is associated with each feedback condition, resulting in a negative valenced in one condition on average over the other. In the case of negative and positive blocks, this is also true at the block level.

We found the best model to include the type of block and temporal bin in the session as factors(The best model was having the form $rt \sim 1 + bin + type + (1|subj_i dx)$,see table 4.1 for details) suggesting that there is overall no difference in RT based on unpleasant vs. pleasant outcome condition. The temporal bins showed a negative relation with RT (β= -0.08 , p <0.001), such that as the session progressed, responses became overall faster regardless of

Best GLM for RT in 70% correct feedback blocks			
Predictors	Estimate(β)	CI	p
(Intercept)	0.02	-0.07 – 0.11	0.631
bin	-0.08	-0.09 – -0.06	**<0.001**
type[NEG]	0.14	0.10 – 0.18	**<0.001**
type [POS]	-0.06	-0.10 – -0.02	**0.005**
type [VAL]	0.00	-0.03 – 0.03	0.945
Random Effects			
σ^2	0.20		
τ_{00subj_idx}	0.12		
ICC	0.37		
N_{subj_idx}	30		
Observations	6324		
Marginal R^2 /Conditional R^2	0.030 / 0.387		

Table 4.1: *Table showing results from the best model selected for response time data. Significant effects are shown in* ***BOLD***

which block we looked into. However, when we compare the coefficients for each block with respect to the symbol block, we found that the negative valence block showed an overall significant increase in response time (β=0.14, p <0.001), when compared to the no valence or symbolic block. This trend was reversed for the block with only positive valence (β= -0.06, p =0.005). Interestingly, the mixed or valence block showed no overall influence on RTs (β= 0.0, p =0.945) which might be because of the individual positive and negative trials' effect on RT got averaged out in the mixed valence block making it similar to that of the control block.

In experiment 2, we manipulated the expected level of uncertainty in the task by introducing a 80% probability of participants receiving correct feedback. To check whether expected uncertainty had an effect on reaction times across the four types of feedback blocks , we again used a generalized linear mixed effects model, this time with the level of uncertainty as additional factor for the forward selection approach to select the best model using AIC as the selection criterion. In this new best model, only the temporal bin in the session got selected as a factor which also showed a significant effect on RT such that, as before, RTs decreased as the session proceeded consistent with the idea of learning. Overall, this result (Table 4.2) suggests that in-

Best GLM for RT data in Uncertainty Comparison			
Predictors	Estimate(β)	CI	p
(Intercept)	0.03	-0.06 – 0.12	0.540
bin	-0.06	-0.08 – -0.05	<**0.001**
Random Effects			
σ^2	0.19		
τ_{00subj_idx}	0.13		
ICC	0.40		
N_{subj_idx}	57		
Observations	4504		
Marginal R^2 /Conditional R^2	0.012 / 0.403		

Table 4.2: *Table showing results from the best model selected for response time data, across different type of percentage correct feedback blocks, for "VAL"(mixed valence) and "SYM"(symbolic) feedback blocks. Significant effects are shown in* ***BOLD***

cluding the level of uncertainty as a factor in the regression analysis did not account for any RT differences implying that uncertainty did not play a role in valence based differences in reaction times.

4.2 Response strategy Analysis: Win-Stay and Lose-Switch Behaviour

In addition to slowing response times, negative valence has been shown to promote more conservative decision-making strategies(Mellers *et al.*,1999;Johnson and Tversky,1983;Sokol-Hessner *et al.*,2013), including staying with the same choice that resulted in a win previously (i.e. win/stay behavior). Though a trend of enhanced switching following an incorrect choice especially in a trial with negative vs. positive outcome could also be expected, it has not been commonly reported in the literature to the best of our knowledge. Hence, we first examined if the win-stay probability was significantly more in a negative vs. positive valence trial in the mixed valence block. To test this, we again ran a generalized linear mixed model through the forward model selection procedure using AIC as a criterion. The following factors were considered for the selection of the best model to separately explain the win-stay and lose-switch choice probabilities: type of feedback block (mixed, positive, negative or no valence/ symbolic), trial condition (pleasant or unpleasant outcome),

Best GLM model for Win-Stay behaviour(70% correct feedback blocks)			
Predictors	Estimate(β)	SE	p
(Intercept)	1.77777	0.18592	<**0.001**
condition [G]	-0.60883	0.10675	<**0.001**
bin	-0.03338	0.09497	0.725
type[NEG]	0.41361	0.16410	0**0.012**
type[POS]	0.10529	0.16362	0.520
type[VAL]	0.01407	0.14917	0.925
bin × type[NEG]	-0.12012	0.16259	0.460
bin × type[POS]	-0.05518	0.14114	0.696
bin × type[VAL]	-0.02399	0.13581	0.860
Random Effects			
σ^2	3.29		
τ_{00subj_idx}	0.52		
ICC	0.14		
N_{subj_idx}	30		
Observations	2236		
Marginal R^2 /Conditional R^2	0.033 / 0.166		

Table 4.3: *Table showing results from the best model selected for win-stay behaviour, across different type of* 70% *correct feedback blocks. Significant effects are shown in* ***BOLD***

temporal bin or block no. in a session. Consistent with the longer RTs in the negative valence block observed in the above analysis, we found trial condition to have a significant effect on the win-stay probability (regardless of the type of block) such that participants stayed with the same choice after a correct response significantly less in case of a pleasant vs. unpleasant outcome trial (β= -0.609, p<0.001). Likewise, when compared to the no valence or positive block, negative block had a significant main effect on win-stay behaviour (β= 0.414, p<0.012) suggesting that the tendency to stay after a win was more overall in the negative valence block than in the control block(Table 4.3). We repeated the same procedure for lose-switch behaviour, but, found no significant interaction or main effect (Table 4.4). Overall, these results point towards the modulatory effect of negative valence feedback on the win

stay behavior such that participants tended to choose a conservative strategy following an unpleasant vs. pleasant image presented as part of feedback.

To examine the effect of expected uncertainty on win-stay and lose-switch behavior, we next compared the data from 70% vs. 80% correct feedback sessions for the mixed and control blocks. We found an overall significant main effect of unpleasant vs. pleasant outcome condition (β -0.643, p $<$ 0.001) as well as mixed vs. no valence block (β=-0.364, p=0.010) on win-stay behavior regardless of the level of uncertainty suggesting higher win-stay choice for unpleasant vs. pleasant outcome valence in the mixed valence block. However, a significant interaction effect between the type of block and level of uncertainty (β=-0.436, p=0.038) suggests that win-stay bias towards the unpleasant vs. pleasant outcomes in the mixed valence block increased with higher uncertainty(Table 4.5)

In the best model for lose-switch behaviour, however, we did not find any effect of expected uncertainty. Only the main effect of the type of block was significant (β= 0.2, p=0.046) suggesting higher switching behavior following an incorrect response in the unpleasant vs. pleasant outcome trial in the mixed vs. no valence block. (Table 4.6)

4.3 Normalized measure for quantifying bias in response strategy

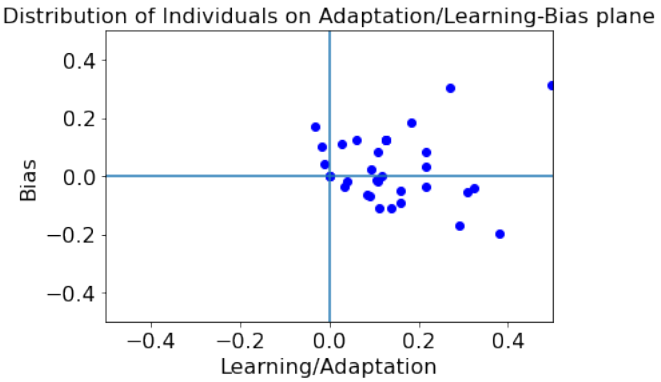


Figure 4.1: *Figure showing individual distributed in the learning/adaptation-bias plane defined in section. Each dot represent an individual 2.2*

To get at a normalized index of an individual's choice strategy and learning performance, we used conditional probability to derive at two independent measures that quantify 1) bias in opting for a previously chosen option

(or the alternative option) after a correct (or incorrect) trial of the same type (unpleasant vs. pleasant outcome valence) and 2) average adaptive response across each session, . As described in the Methods section, the bias in win-stay policy was quantified as the log of the ratio between win-stay choice factor (WS, a term derived by dividing the conditional probability of staying and winning with the probability of staying, see equation 2.23) in unpleasant and pleasant outcome conditions. Hence, if the ratio is greater than one it is suggestive of the probability of staying after a win being higher in unpleasant outcome condition than that in the pleasant outcome condition, and the log of this ratio would be a positive value. On the other hand, if the ratio between win-stay choice factors in unpleasant and pleasant outcome conditions is less than one, it is indicative of the probability of staying after a win being higher in the pleasant outcome condition than that in the unpleasant outcome condition and the log of this ratio would be a negative value. We averaged values of bias across individuals (Figure 4.1 showing bias for win-stay across individuals plotted against their respective average adaptive response values, estimated as the log of geometric mean of WS across the two trial conditions) in a sample to derive a population estimate which for each experimental block of interest was then compared against the no valence control block. To quantify the effect, we bootstrapped the differences between means sampled for the same set of individuals (repeated 1000 times to construct the distribution) which was then compared against zero, with the null hypothesis of no win stay bias for the unpleasant vs. pleasant outcome condition. A bootstrap distribution of greater that zero for $\geq 95\%$ of the time indicates a significant difference. Figure 4.2 plots the the bootstrap distribution of difference between means for bias in win-stay behaviour. We found a significant bias in case of mixed valence block compared to symbolic feedback block. This was not true for any other experimental blocks suggestive of a win-stay bias emerging in a context when the valenced outcomes of equal strength are presented with equal frequency. Even though an effect of valenced outcome was observed on RT and win-stay choice behavior in the separate valence blocks suggestive of the modulatory role of outcome valence on behavior, a bias in win-stay policy for negative vs. positive outcome valence using the new normalized measure could be seen only in the mixed vs. no valence block. We attribute this to the use of conditional probabilities in deriving the new bias measure which (unlike GLM) does not rely on any underlying assumptions (e.g. independence between events in sequential trials) and therefore is likely more sensitive to strategic biases in choices occurring over trials. We also compared the average adaptive response measure across blocks but did not find any differences when comparing to symbolic feedback block for any other feedback block(Figure 4.3).

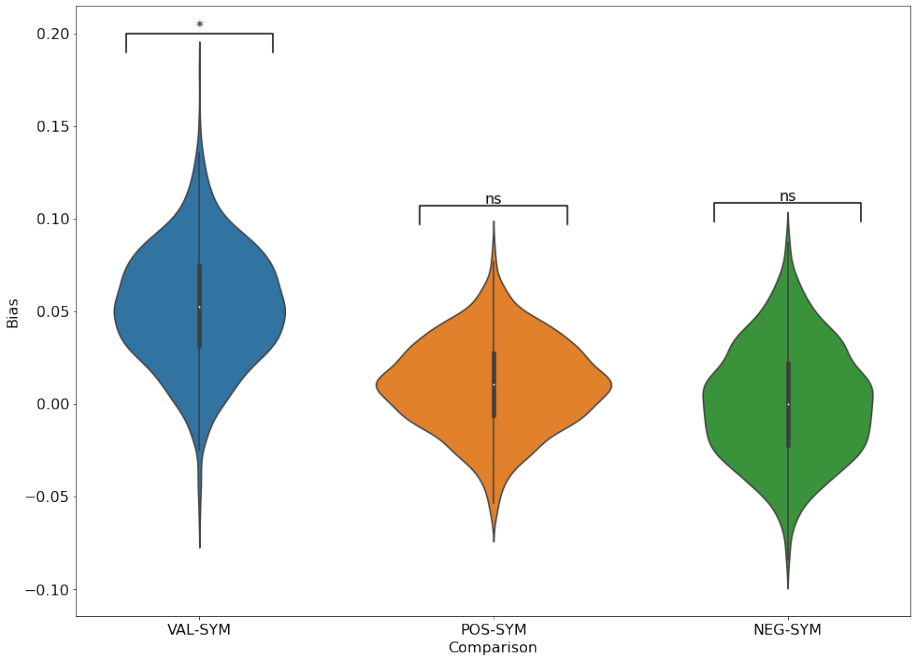


Figure 4.2: *Bootstrap distribution of the difference between the mean of bias for different types of feedback blocks against control block(SYM). For each block, individual data points were simultaneously sampled from the control block and corresponding blocks(VAL/POS/NEG) with replacement. The mean value of bias was then computed for them, and calculated the difference between the means for each type of block separately. This procedure was repeated 1000 times to construct the bootstrap distribution. In this distribution, in the case of mixed valence outcome("VAL")(blue), the difference between was less than or equal to zero, only for < 0.05 fraction of the total bootstrap samples, thus leaving a very little evidence in support of the null hypothesis, which leads to a rejection of the null hypothesis. "*" denote significance, and "ns" indicate non-significant results.*

While comparing bias in win stay policy across the two levels of uncertainty, we found that the high bias that was observed in 70% correct feedback blocks almost vanished in 80% correct feedback block, in line with our hypothesis that the uncertainty had a significant modulatory effect on response bias towards unpleasant vs. pleasant outcome condition and tends to amplify it(Figure 4.4).

In case of lose switch behaviour, we found a significant effect of uncertainty on the normalized adaptive learning measure: compared to the no valence control block, there was an overall decrease in average learning in the more certain 80% correct feedback mixed block(Figure 4.5) .

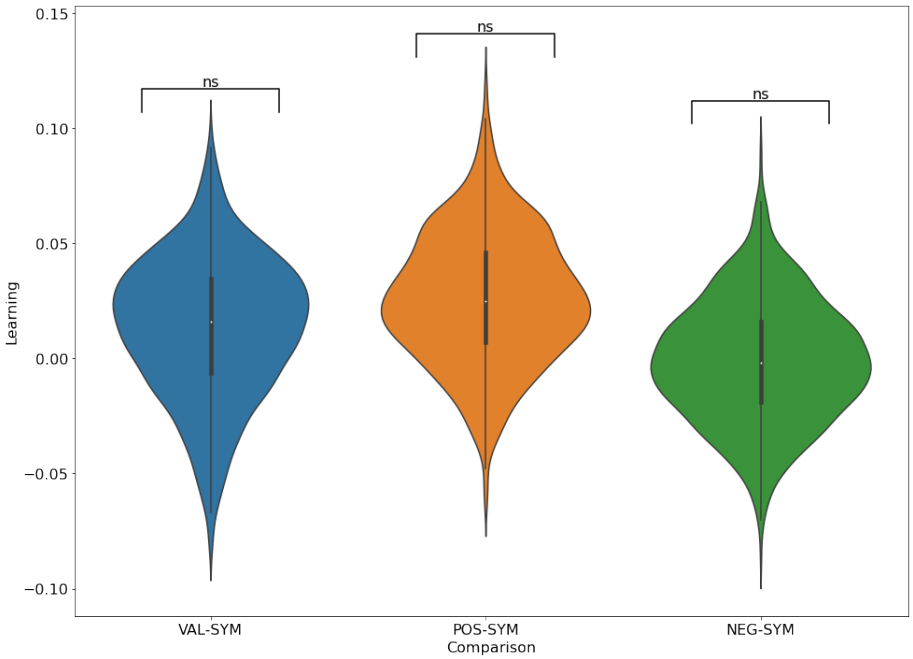


Figure 4.3: *Bootstrap distribution of the difference between the mean of adaptation/learning for different types of feedback blocks against control block(SYM). For each block, individual data points were simultaneously sampled from the control block and corresponding blocks(VAL/POS/NEG) with replacement. The mean value of bias was then computed for them, and calculated the difference between the means for each type of block separately. This procedure was repeated 1000 times to construct the bootstrap distribution. In this distribution, in the case of all kinds of feedback blocks, the mean difference was less than or equal to zero for < 0.05 fraction of the total bootstrap samples, thus leaving evidence in support of the null hypothesis, Thus failing to reject the null hypothesis. "*" denote significance, and "ns" indicate non-significant results.*

4.4 RL-DDM results

Next, we modeled the decision-making process using the RL-DDM approach to understand if, and how, compared to the control, no valence block, the model parameters are modulated by pleasant vs. unpleasant outcome valence across the experimental blocks. For this, we first used the forward modeling approach to derive at the best model by independently estimating the model parameters separately for each of the conditions (negative and positive valenced conditions) and evaluating the improvement in model using Deviance information Criteria (DIC), which is typically used for evaluating performance of models, mostly in Bayesian statistics. The best model was the one having parameters separately estimated for prior bias (z), learning rate (α) and, scaling factor (m). Across experimental blocks, a credible dif-

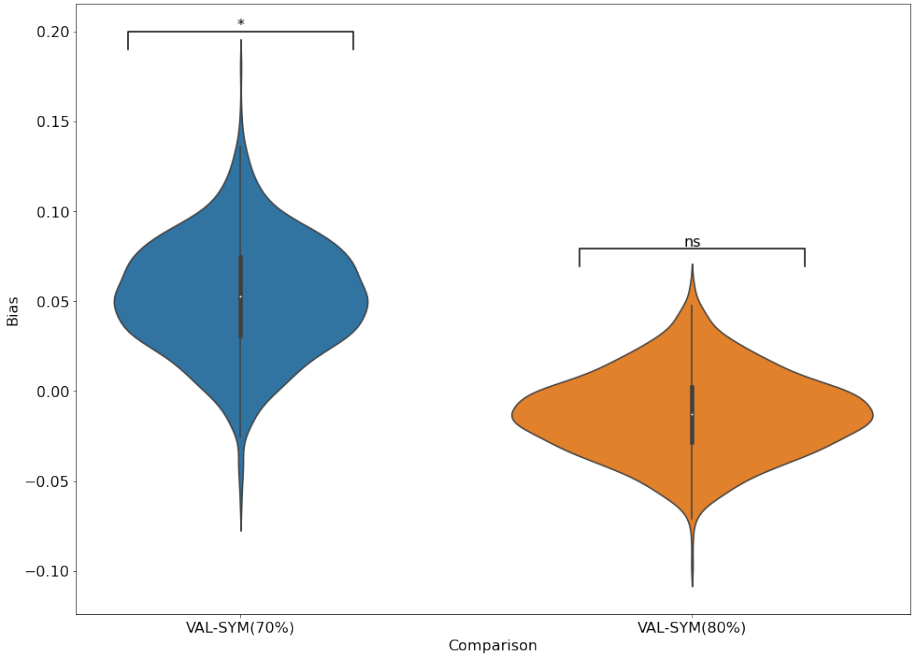


Figure 4.4: *Bootstrap distribution of the difference between the mean of bias for mixed feedback block(VAL) against control block(SYM) for different uncertainty conditions for win-stay behavior. Individual data points were simultaneously sampled from the control block and mixed block corresponding to the same uncertainty for each uncertainty block, with replacement. The mean value of bias was then computed for them, and calculated the difference between the means for each mixed valence block and the corresponding control(SYM) block was separated. This procedure was repeated 1000 times to construct the bootstrap distribution. In this distribution, in the case of a high uncertainty block having 70 percent correct feedback(blue), the mean difference was less than or equal to zero, only for < 0.05 fraction of the total bootstrap samples, thus leaving a very little evidence in support of the null hypothesis, which leads to a rejection of the null hypothesis. "*" denote significance, and "ns" indicate non-significant results.*

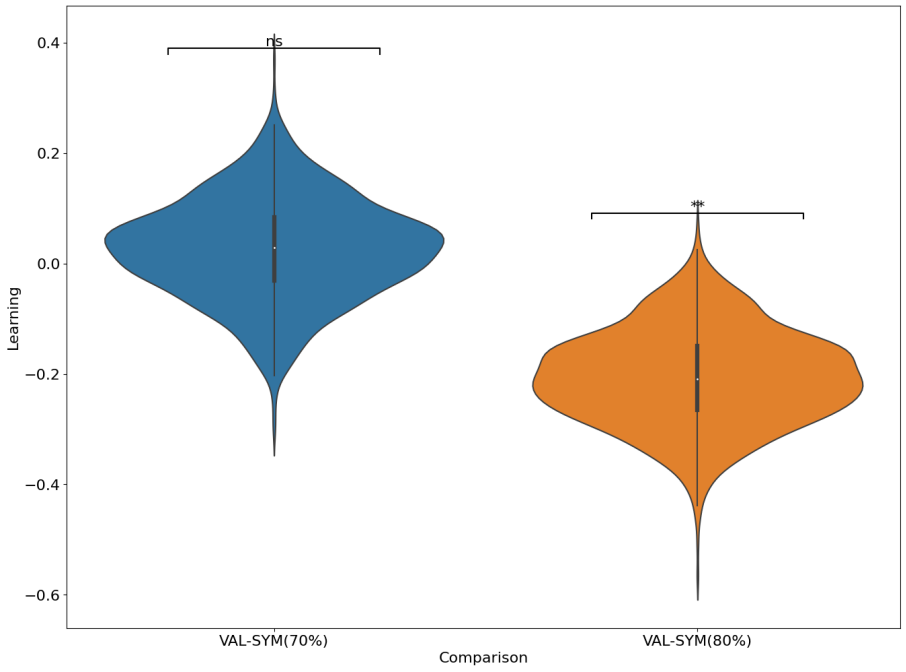


Figure 4.5: *Bootstrap distribution of the difference between the mean of learning/adaptation for mixed feedback block(VAL) against control block(SYM) for different uncertainty conditions for lose-switch behavior. Individual data points were simultaneously sampled from the control block and mixed block corresponding to the same uncertainty for each uncertainty block, with replacement. The mean value of bias was then computed for them, and the difference between the means for each mixed valence block and corresponding control(SYM) block was calculated separately. This procedure was repeated 1000 times to construct the bootstrap distribution. In this distribution, in the case of both types of uncertainty blocks, the difference between the mean was less than or equal to zero for < 0.05 fraction of the total bootstrap samples, thus leaving evidence supporting the null hypothesis, hence failing to reject the null hypothesis. "*" denote significance, and "ns" indicate non-significant results.*

Figure 4.6: *The difference in prior bias(z) being controlled across types of feedback ("VAL"= mixed valenced images, "SYM"= symbols, "POS"/"NEG" = a valenced image and a symbol). p(bad>good) was computed on the posterior distributions and corrected for multiple comparisons using Bonferroni correction. "*" denote a credible difference between the posterior distributions*

ference was found between posterior distributions of only parameter prior bias(z)(Figure 4.6) in the mixed valence block. This was of special interest as we were expecting the bias to be showing up in some parameter other than the learning rate. This supports the use of the RL-DDM framework for modeling our data since it could capture the bias present in the behavior even though there was no observable difference in the learning rate.

Consistent with the idea of uncertainty amplifying the bias towards an unpleasant vs. pleasant outcome condition, compared to what we found in Expt. 1 with 70% correct feedback probability, we found no difference in the prior bias parameter of the RL-DDM in the mixed vs. no valence block in Expt. 2 with 80% correct feedback probability. Again, no changes were observed for other RL-DDM parameters including learning rate(Figure 4.8).

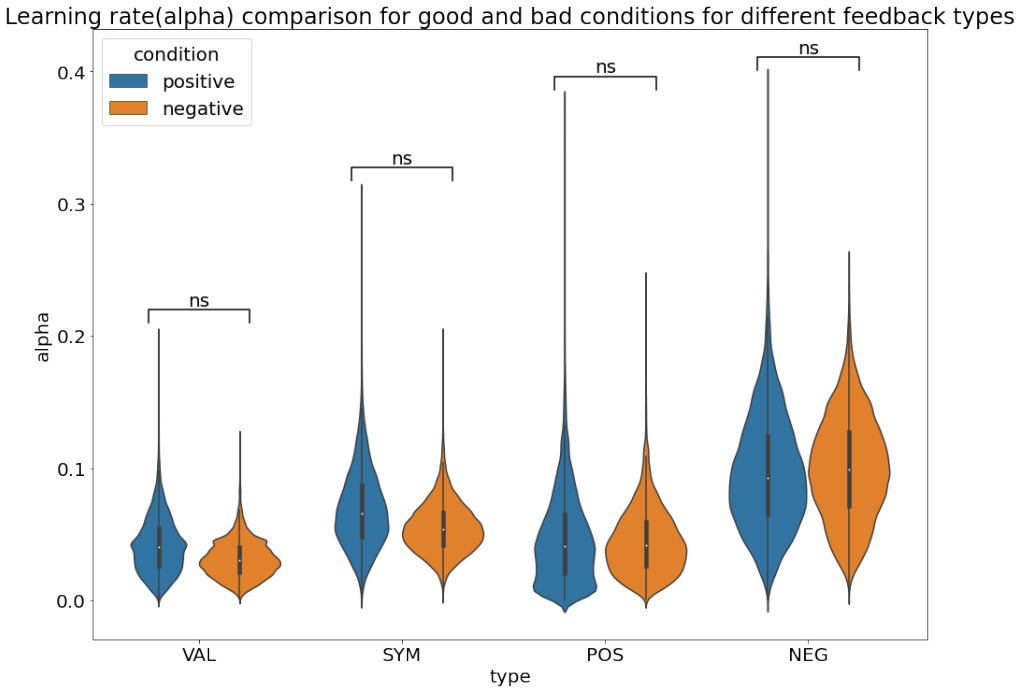


Figure 4.7: *The difference in learning rate(α) is controlled across types of feedback ("VAL"= mixed valenced images, "SYM"= symbols, "POS"/"NEG" = a valenced image and a symbol). p(bad>good) was computed on the posterior distributions and corrected for multiple comparisons using Bonferroni correction. "*" denote a credible difference between the posterior distributions*

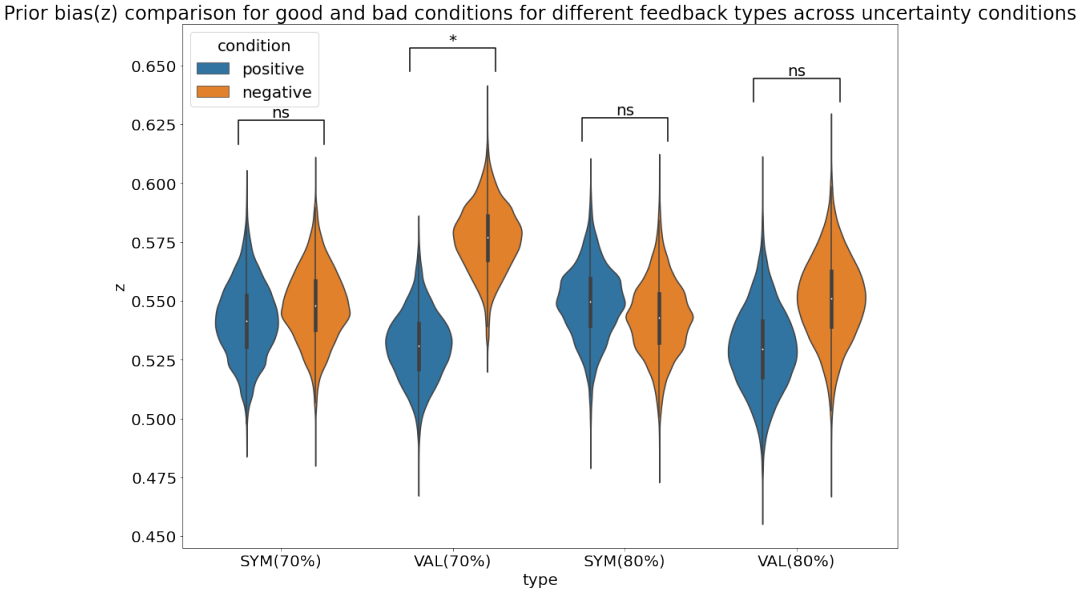


Figure 4.8: *Difference in prior bias(z) being compared across types of feedbacks, and uncertainties("VAL(80%)"= mixed valenced images in 80% correct feedback block, "SYM(80%)"= symbolic feedback in 80% correct feedback block,"VAL(70%)"= mixed valenced images in 70% correct feedback block and, "SYM(70%)"= symbolic feedback in 70% correct feedback block,). p(bad>good) was computed on the posterior distributions and corrected for multiple comparisons using Bonferroni correction. "*" denote a credible difference between the posterior distributions.*

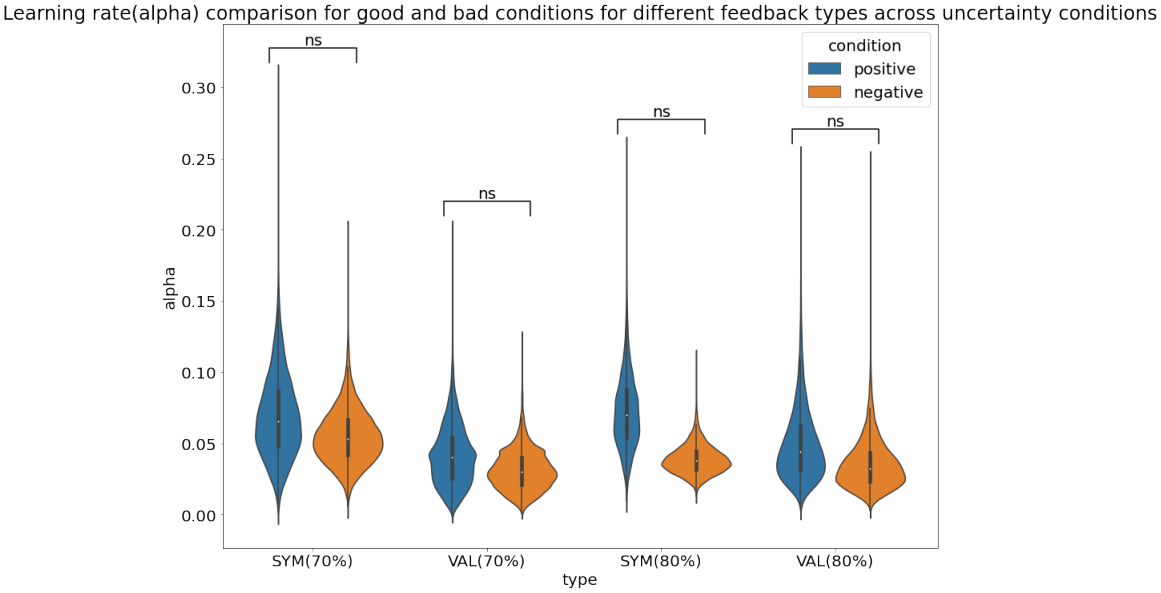


Figure 4.9: *Difference in learning rate(α) being compared across types of feedbacks, and uncertainties("VAL(80%)"= mixed valenced images in 80% correct feedback block, "SYM(80%)"= symbolic feedback in 80% correct feedback block,"VAL(70%)"= mixed valenced images in 70% correct feedback block and, "SYM(70%)"= symbolic feedback in 70% correct feedback block,). p(bad>good) was computed on the posterior distributions and corrected for multiple comparisons using Bonferroni correction. "*" denote a credible difference between the posterior distributions.*

4.5 Trait based explanations of bias

Next, we were interested to investigate if the variability in the bias measure bias shown by individuals in our sample for win stay and lose switch behavior is explained by any of their questionnaire scores indexing relevant psychological traits like anxiety, depression and sensitivity to negative stimuli like pain and disgust. We incorporated participants' individual questionnaire scores as potential factors in a multi-level regression model explaining first their win-stay behavior in the mixed valence block. As before, we used a forward model selection approach using AIC as a criteria.

In linear model for bias, we found a significant main effect of state and trait anxiety inventory scores (β=-0.04, p= 0.031) in the best model(Table 4.7) such that that higher the STAI score (corresponding to higher levels of anxiety) lower is the observed bias in win stay strategy for unpleasant vs. pleasant outcome condition.

In terms of average adaptive response in win-stay behaviour using the best model, there were no significant main effects(Table 4.8).

For lose switch behaviour only the pain catastrophizing scale (PCS) significantly accounted for the bias in strategy (β=0.11, p=0.04) such that higher the PCS score greater is the observed bias in the lose-switch performance in the unpleasant vs. pleasant outcome condition(Table 4.9).

PCS also showed a positive and marginally significant effect on adaptation in lose switch behaviour (β=0.12, p=0.054). Overall, this analysis suggests that trait scores like PCS and STAI could explain some of the variability in the negative bias observed in our sample(Table 4.10).

Finally, we tested if the trait-behaviour relationship we observed in the 70% correct feedback sessions is modulated by uncertainty. We found a significant main effect of STAI scores on bias (β=-0.03, p=0.02) and average adaptive response learning (β= 0.03, p=0.035)(Table 4.11 and Table 4.12)in the win-stay performance, regardless of the level of uncertainty. However, the relationship between STAI and these measures were almost equal and opposite such that while higher STAI scores were associated with lower choice bias in win stay strategy for unpleasant vs. pleasant outcome condition (could be related to generalization of unpleasant outcome effect to pleasant outcome trials leading to lesser difference between the two trial conditions), they were associated with higher average adaptive learning which together is consistent with high anxiety being correlated with high vigilance -related adaptive changes in behavior in the literature.

In case of lose-switch behavior, we found a significant negative interaction between uncertainty and STAI scores (β= -0.16, p=0.027) suggesting that the positive relation between STAI and lose switch bias decreased with increase

in uncertainty(Table 4.13). We also found a significant positive main effect of PCS (beta=0.11, p=0.017) and DSR (β=0.23, p=0.002) scores on average adaptive response for lose-switch performance in unpleasant vs. pleasant outcome condition, regardless of the level of uncertainty. In addition, a significant negative interaction was observed of DSR scores with uncertainty (β=-0.21, p=0.022)(Table 4.14).

Best GLM model for Lose-Switch behaviour(70% correct feedback blocks)			
Predictors	Estimate(β)	SE	p
(Intercept)	-0.35584	0.15244	**0.02**
condition [G]	0.12441	0.19309	0.519
bin	0.01468	0.13430	0.913
type[NEG]	-0.06870	0.21806	0.753
type[POS]	0.17856	0.20563	0.385
type[VAL]	0.35520	0.19761	0.072
condition [G] × type[NEG]	0.14368	0.29568	0.627
condition [G] × type[POS]	-0.41240	0.27700	0.137
condition [G] × type[VAL]	-0.30752	0.26473	0.245
bin × condition [G]	-0.12608	0.18098	0.486
bin × type[NEG]	0.25215	0.22823	0.269
bin × type[POS]	0.22372	0.18915	0.237
bin × type[VAL]	0.06323	0.19093	0.741
bin × condition[G] × type[NEG]	0.16685	0.31589	0.5974
bin × condition[G] × type[POS]	-0.28284	0.26359	0.2833
bin × condition[G] × type[VAL]	0.08498	0.26031	0.7441
Random Effects			
σ^2	3.29		
τ_{00subj_idx}	0.07		
ICC	0.02		
N_{subj_idx}	30		
Observations	1748		
Marginal R^2 /Conditional R^2	0.015 / 0.035		

Table 4.4: *Table showing results from the best model selected for lose-switch behaviour, across different type of* 70% *correct feedback blocks. Significant effects are shown in* ***BOLD***

Best for effect of uncertainty on Win-Stay behaviour			
Predictors	Estimate(β)	SE	p
(Intercept)	2.07803	0.21333	<**0.001**
condition [G]	-0.64341	0.12814	<**0.001**
type [VAL]	-0.36390	0.14072	**0.010**
prob incorr [30p]	-0.26775	0.29277	0.360
type [VAL] × prob incorr [30p]	0.43627	0.21057	**0.038**
condition [G] × prob incorr [30p]	-0.01622	0.19120	0.932
Random Effects			
σ^2	3.29		
τ_{00subj_idx}	0.65		
ICC	0.16		
N_{subj_idx}	57		
Observations	2667		
Marginal R^2 /Conditional R^2	0.030 / 0.190		

Table 4.5: *Table showing results from the best model selected for win-stay behaviour, across different type of* 70% *and* 80% *correct feedback blocks, including "VAL" and "SYM" feedback blocks. Significant effects are shown in* ***BOLD***

Best GLM model for effect of uncertainty on Lose-Switch behaviour			
Predictors	Estimate(β)	SE	p
(Intercept)	-0.07549	0.11652	0.517
condition [G]	-0.07378	0.09462	0.436
type [VAL]	0.19959	0.09992	**0.046**
prob incorr [30p]	-0.18895	0.11684	0.106
Random Effects			
σ^2	3.29		
τ_{00subj_idx}	0.04		
ICC	0.01		
N_{subj_idx}	57		
Observations	1837		
Marginal R^2 /Conditional R^2	0.006 / 0.019		

Table 4.6: *Table showing results from the best model selected for lose-switch behaviour, across different type of* 70% *and* 80% *correct feedback blocks, including "VAL" and "SYM" feedback blocks. Significant effects are shown in* ***BOLD***

Best linear model for Bias in win-stay behaviour in 70%			
Predictors	Estimate(β)	CI	p
(Intercept)	0.03	-0.01 – 0.07	0.088
STAI	-0.04	-0.09 – -0.00	**0.031**
Observations	31		
R^2 /Adjusted R^2	0.150 / 0.121		

Table 4.7: *Effects from linear model for Bias in win-stay behaviour in 70 percentage correct mixed feedback block. Significant effects are shown in* ***BOLD***

Best linear model for Adaptation in win-stay behaviour in 70%			
Predictors	Estimate(β)	CI	p
(Intercept)	0.14	0.10 – 0.18	**<0.001**
BDI	0.03	-0.01 – 0.07	0.158
Observations	31		
R^2 /Adjusted R^2	0.068 / 0.035		

Table 4.8: *Effects from linear model for adaptation in win-stay behaviour in 70 percentage correct mixed feedback block. Significant effects are shown in* ***BOLD***

Best linear model for Bias in lose-switch behaviour in 70%			
Predictors	Estimate(β)	CI	p
(Intercept)	0.15	0.05 – 0.26	**0.006**
STAI	-0.11	-0.23 – 0.01	0.082
PCS	0.11	0.01 – 0.22	**0.040**
PBQ	-0.08	-0.17 – 0.02	0.112
Observations	23		
R^2 /Adjusted R^2	0.376 / 0.278		

Table 4.9: *Effects from linear model for Bias in lose-switch behaviour in 70 percentage correct mixed feedback block. Significant effects are shown in* ***BOLD***

Best linear model for adaptation in lose-switch behaviour in 70%			
Predictors	Estimate(β)	CI	p
(Intercept)	0.37	0.25 – 0.49	**<0.001**
PCS	0.12	-0.00 – 0.24	0.054
Observations	23		
R^2 /Adjusted R^2	0.166 / 0.126		

Table 4.10: *Effects from linear model for adaptation in lose-switch behaviour in 70 percentage correct mixed feedback block. Significant effects are shown in* ***BOLD***

Best linear model for Bias in win-stay behaviour with uncertainty			
Predictors	Estimate(β)	CI	p
(Intercept)	0.03	-0.00 – 0.06	0.054
STAI	-0.03	-0.06 – -0.01	**0.020**
Observations	55		
R^2 /Adjusted R^2	0.098 / 0.081		

Table 4.11: *Effects from linear model for Bias in win-stay behaviour with uncertainty. Significant effects are shown in* ***BOLD***

Best linear model for Adaptation in win-stay behaviour with uncertainty			
Predictors	Estimate(β)	CI	p
(Intercept)	0.12	0.10 – 0.15	**<0.001**
STAI	0.03	0.00 – 0.06	**0.035**
Observations	55		
R^2 /Adjusted R^2	0.081 / 0.064		

Table 4.12: *Effects from linear model for Adaptation in win-stay behaviour with uncertainty. Significant effects are shown in* ***BOLD***

Best linear model for Bias in lose-switch behaviour with uncertainty			
Predictors	Estimate(β)	CI	p
(Intercept)	0.10	0.03 – 0.17	**0.006**
PSS	0.05	-0.02 – 0.13	0.173
STAI	0.04	-0.07 – 0.15	0.465
prob incorr [30p] × STAI	-0.16	-0.31 – -0.02	**0.027**
Observations	40		
R^2 /Adjusted R^2	0.190 / 0.123		

Table 4.13: *Effects from linear model for Bias in lose-switch behaviour with uncertainty. Significant effects are shown in* ***BOLD***

Best linear model for adaptation in lose-switch behaviour with uncertainty			
Predictors	Estimate(β)	CI	p
(Intercept)	0.40	0.31 – 0.48	<**0.001**
PCS	0.11	0.02 – 0.20	**0.017**
DSR	0.23	0.09 – 0.36	**0.002**
prob incorr [30p] × DSR	-0.21	-0.39 – -0.03	**0.022**
Observations	40		
R^2 /Adjusted R^2	0.348 / 0.293		

Table 4.14: *Effects from linear model for adaptation in lose-switch behaviour with uncertainty. Significant effects are shown in* ***BOLD***

Chapter 5

Discussion

Negativity bias has been operationalized in the literature as the phenomenon where a unit of negative valence induces a stronger response when compared to a unit of an equally strong positive valence stimulus (Norris *et al.*,2010). Behaviorally this bias has been documented as more rapid responses to aversive vs. appetitive stimuli, higher incidental learning of spatial location with unpleasant vs. pleasant stimuli and so on, while physiologically, this has been reported as higher ERP, fMRI and skin conductance responses to unpleasant vs. pleasant stimuli across visual and auditory modalities (Smith *et al.*,2003). In this study, we tested if keeping the value related motivational salience of feedback the same, valence modulated the choices and response strategy of participants in a probabilistic cue association learning task. In addition, given the role of uncertainty in emotion, particularly its aversiveness, we examined if there was an interaction between level of uncertainty and valence in eliciting a response bias in learning. To this end, we modified a standard probabilistic reinforcement learning task so as to dissociate valence from correct vs. incorrect feedback for learning about two cues and examined whether participants' performance showed a valence specific trend. We quantified the valence-based bias by deriving a new normalized measure of response strategy for unpleasant vs. pleasant outcome conditions as well as modeling the decision-making process using the RL-DDM(Pedersen *et al.*,2017) framework. Further, we looked at whether and how the bias in response strategy changed with the level of uncertainty and trait level questionnaire scores of participants on relevant dimensions like anxiety, depression, sensitivity to negative stimuli like pain, disgust etc.

5.1 Response Times Analysis

We first looked at response times for choices made in the various blocks that differed in the nature of feedback they provided, namely: mixed, positive, negative vs. no valence. and found that compared to the no valence or symbolic block, reaction times in the positive valence block (having positively valenced image or a symbol as a feedback) were shorter while those in the negative valence block (having a negatively valenced image or a symbol as feedback) were longer. In Simpson et al(Simpson *et al.*,2000) a similar observation can be seen in terms of the negatively valenced stimuli. In this study adult participants were asked to respond whether the number of humans they are seeing in a valenced picture (either neutral or negatively valenced) was 0-1 or more than that. Two by third of the total images were neutral and were taken from IAPS(Lang *et al.*,1997a). And using an ANOVA on correct responses, they observed a significant main effect of valence: participants were responding slower to negative stimuli. And at the core of this difference we can only expect that the negative valence would be receiving more attention, even at an earlier stage of the decision making process, which is getting reflected in the later stages also. In Estes et al(Estes and Verges,2008) using a lexical decision making task, authors found that participants responded with slower response time for negative valence but, with a faster valence discrimination, in terms of accuracy. And such reliable occurrence of negative delay led researchers to speculate a possibility of negative stimuli eliciting a general suppression of motor activity(Fox *et al.*,2001). In context of the stroop effect, it was suggested that the behaviour could be a defense mechanism that responds to threat by freezing all activities(Algom *et al.*,2004), and this was called the motor suppression hypothesis. Motor suppression hypothesis was later questioned because not all threatening stimuli would lead to a freezing response, sometimes it would mount a quick and immediate response(Ferrari and Ferrari,1990;Searcy and Caine,2003).

When taking these suggestions into consideration, it is possible that the kind of response the negative stimuli gives rise to also depends on the task. Like, in studies that were challenging the motor suppression hypothesis, the stimulus valence would be more relevant, as they are more related to survival, when responding to an immediate threat(Ferrari and Ferrari,1990;Searcy and Caine,2003). In tasks where the valence is irrelevant, like the lexical decision making task, color naming or word naming task, where the participants have to disengage attention from the valence information to perform optimally, negative valence may result in longer response time. In that context, previous work is consistent with our findings regarding slower response times in negative vs. positive valence blocks since the valence information is not

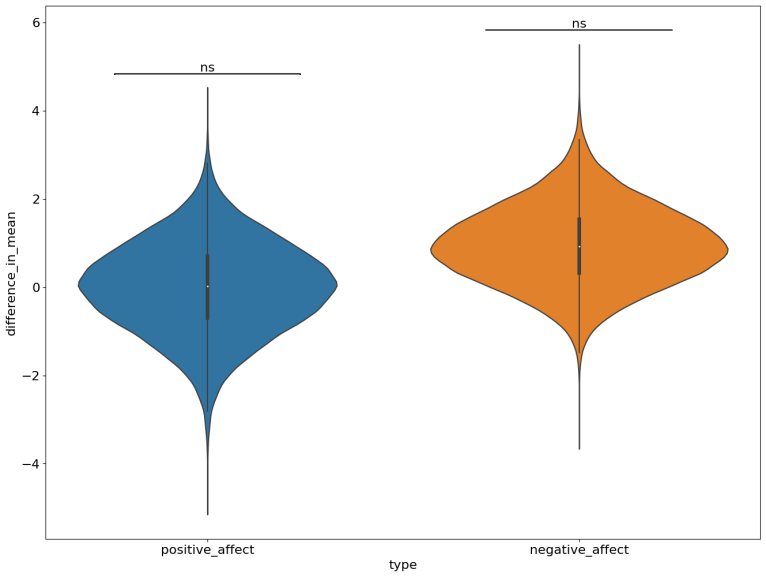


Figure 5.1: *Change in mean positive affect was computed by subtracting the mean positive affect score following the negative block from that following the positive block. This was implemented on 1000 random samples with replacement to construct the bootstrap distribution. Change in mean negative affect was computed by subtracting the mean negative affect score following the positive block from that following the negative block. This too was implemented on 1000 random samples with replacement to construct the bootstrap distribution*

relevant for performing optimally. And it is possible to have an opposite effect on tasks in which the negative information is required to be processed explicitly.

But, in our task, where the negative and positive stimuli were presented with certain probability, we were also suspecting the possibility of an induced affect, because of which we also collected PANAS(Watson *et al.*,1988) ratings after positive and negative blocks, but couldn't find any difference between positive and negative affect as given in figure 5.1. The non-significant difference in the induced affect suggests that the observed difference in RT may not be because of a difference in affect induced by the stimuli, but rather a consequence of the expected outcome in each block that differed in its valence. That is, in the case of a negative block, participants expected a negative valence feedback which led to slower response times, and the converse would be true for positive valence blocks. And in a mixed block, where the outcome could either be positively or negatively valenced, this apparent difference in RT vanished, as both the valence presented together likely canceled out each other's effect. However, consistent with previous research, RT did not capture the effect of condition, i.e., whether the trial outcome was

pleasant or unpleasant.

5.2 Win-stay and Lose-switch behaviour

In the case of win-stay behaviour, we found a significant main effect of condition, i.e., whether the Japanese alphabet was associated with an unpleasant or pleasant outcome , irrespective of the type of block, suggesting that the behaviour could be generalizable across blocks Similar observation was noted by Mellers et al(Mellers *et al.*,1999), where they investigated the effect of emotion on decision making and varying contexts. One experiment they conducted involved choosing between two options, where one was more risky with higher pay-off, while the other one was less risky, but with a reduced pay-off. Then they manipulated the valence context, in which the potential gains associated decisions were either negative, positive or neutral. For example, in a positively valenced context, in high risk, participant would be either getting a reward or nothing with 50% probability, where the reward amount would be high, and less amount, with a higher probability for less risky option. In the negative context, these potential gains were replaced with losses. Authors found that in a negatively valenced context, participants are more likely to stay with the initial choice after a win compared to positive or neutral contexts consistent with our results. Such observations have also been reported in the context of loss aversion. For example, Johnson et al(Johnson and Tversky,1983) investigated the effect of valence on an individual's risk taking behaviour. The valence context of the decision outcome was manipulated by framing it as gain or loss. The authors found that the participants were less likely to take risks in the negative valence condition in comparison to the positive valence condition confirming that negative valence may induce a more conservative strategy. Sokol et al(Sokol-Hessner *et al.*,2013) extended this line of work by implementing a gambling task after presenting participants with valenced images. It was found that participants, who were presented with negative valenced images were more likely to show loss aversive behaviour. In short, the possible explanations for the effect that we are observing could be that the negative emotion or valence increased an individual's sensitivity to potential losses(Lerner and Keltner,2001;Baumeister *et al.*,2001).

We used conditional probability to derive at two independent measures that quantify 1) bias in opting for a previously chosen option (or the alternative option) after a correct (or incorrect) trial of the same type (unpleasant vs. pleasant outcome valence) and 2) average adaptive response across each

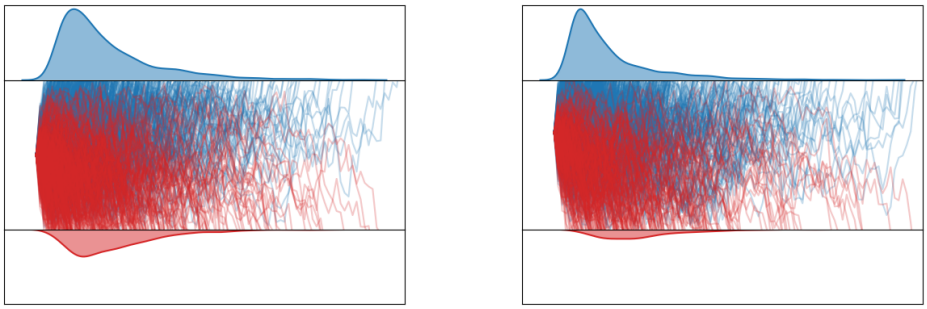
session. In the bootstrap comparison for difference between means in the experimental vs. no valence, control block, we observed a significant difference for mixed valence block ($p<0.05$ suggesting that the participants were more likely to stay after a win in unpleasant vs. pleasant outcome condition. This is consistent with the idea discussed above pertaining to negative valencemaking participants rely on more conservative strategies to reduce losses(Mellers *et al.*,1999;Johnson and Tversky,1983;Sokol-Hessner *et al.*,2013;Lerner and Keltner,2001;Baumeister *et al.*,2001). An absence of such a bias between unpleasant vs. pleasant conditions in either the positive only or negative only blocks suggests that the win-stay strategy is probably influenced by the relative value associated with both the outcomes. Even though an effect of valenced outcome was observed on RT and win-stay choice behavior in the separate valence blocks suggestive of the modulatory role of outcome valence on behavior, a bias in win-stay policy for negative vs. positive outcome valence using the new normalized measure could be seen only in the mixed vs. no valence block. We attribute this to the use of conditional probabilities in deriving the new bias measure which (unlike GLM) does not rely on any underlying assumptions (e.g. independence between events in sequential trials) and therefore is likely more sensitive to strategic biases in choices occurring over trials.

5.3 RL-DDM Capturing the Bias

We used RL-DDM for behavioural modeling since it is possible that the difference in behavior due to valence be captured by a parameter other than the learning rate which has not been able to consistently differentiate between outcome valence in the past(Palminteri *et al.*,2012). Recently, Kim et al(Kim and Anderson,2022) investigated how stimuli of same modality (liquid outcome – sweet juice or salty water) would influence attention and thereby, cue association learning. The authors found that the stimuli coding for primary reward or punishment biased attention leading to lower error rate and a shorter RT. In other words, the study failed to observe any selective attentional bias towards either valence, though stimuli that were previously associated with affective outcomes were found to draw more attention based on response ratios and RTs. However, given the learning rate codes for how much weight is given to the recent feedback, in a learning task with fixed probability, we do not expect it to capture much of the differences, as it is likely that after an initial exploratory phase within a session, the learning rate would decay in such studies and the difference over a session may not be able to capture finer differences. We overcame this limitation by using RL-

DDM(Pedersen *et al.*,2017) which enabled us to investigate other potential parameters that could explain the behavior better. It was found that across different types of feedback blocks, only in mixed valence feedback type, the posterior distribution for prior bias for unpleasant outcome condition was credibly higher than that of pleasant outcome condition suggesting that only when presented together in a session, the unpleasant outcome induced a prior bias such that the evidence integration process for decision making itself started from a point already closer to the response option associated with the unpleasant outcome(see Figure 5.2). From the DDM perspective, this suggests that while in the positive cue condition, the decision process is likely to hit both the incorrect decision boundaries with more probability. In the negative cue condition, the shift that the prior bias has towards the boundary for predicting that the outcome is going to be unpleasant, would make it easy to cross that boundary. As is evident from the modeling results, it is neither the quality of evidence nor the feedback error that makes the updating of values for the unpleasant outcome condition stronger, but merely the shifting of the the initial starting point that seems to captured bias towards the unpleasant vs. pleasant outcome condition . Though such a shift was found in the negative valence block as well, the value of z was not found to be credibly different between the conditions. No other block showed this shift, which suggests that it could be attributed to the presence of negative outcome stimuli. The shift in prior bias in negative valenced condition compared to positive valenced condition, could be a potential mechanism underlying such biases, such that the individuals would be shifting the prior bias, when the negative feedbacks are more, for not falsely detecting a negative feedback as positive, which would be a better strategy for survival. Additionally, we found the observed shift in prior bias to be reduced at a lower level of uncertainty.

Importantly, the learning rate parameter in our task did not change which is consistent with other studies. For example, Nassar et al(Nassar *et al.*,2012), hypothesized that changes in expected uncertainty would modulate learning rate, and the study implemented a probabilistic learning task. But, contrary to the hypothesis, they concluded that the learning rate may not change with the expected uncertainty. In a recent study, Duff et al(Duff *et al.*,2013), authors implemented a serial probabilistic learning task, and hypothesized that hippocampus plays an important role in modulating learning rate with changes in expected uncertainty. However, contrary to their hypothesis, they were unable to find modulation in learning rates with changes in expected uncertainty. This confirms with our expectation that some other parameter other than the learning rate would capture the valence based modulation on response strategy with changes in uncertainty. In our data, this turned out



(a) (b)

Figure 5.2: *A pictorial comparison between decision processes when there is a shift in prior bias(z) parameter as we have seen in the study. (a) A symbolic representation of diffusion decision process in positive valence condition of the mixed block. Where the upper boundary represents predicting or choosing positive outcome valence and lower boundary represents predicting or choosing negative outcome valence. (b) A symbolic representation of diffusion decision process in negative valence condition of the mixed block. Where the upper boundary represents predicting or choosing negative outcome valence and lower boundary represents predicting or choosing positive outcome valence. We can see that compared to positive valence condition(a), the negative valence condition(b) is now making lesser number of incorrect responses because of a shift in prior bias(z). This is because the shift in prior bias makes it easy to reach upper boundary compared to that of the lower one.*

to be the prior bias parameter.

A signal detection theory perspective of such behaviours could be found in Macmillan et al(Macmillan and Creelman,2004, chapters: 2,4,8,9), where it was suggested that both the sensitivity and response bias would influence the signal detection, where sensitivity helps to distinguish signal and noise and bias would make the response in a particular way. In DDMRatcliff and McKoon,2008) framework, drift rate is a parameter coding for strength of evidence(Ratcliff and McKoon,2008;Wagenmakers *et al.*,2008), which isn't changing with the condition, as there is no credible difference between the posterior distributions, which could suggest that it is not at the level of detection or evidence that the observed bias is a result of. Rozin et al(Rozin and Royzman,2001) suggested that negative events would have a larger impact on evaluation, behaviour, judgement, etc, when compared to positive events and posited three possible mechanisms for the bias.

- Evolutionary factors prioritizing detection of threat and danger
- Cognitive processes as drawing greater attention for negative stimuli
- Affective processes giving stronger response to a negative stimuli

It was suggested that the survival of ancestor depends on the ability to respond to threat in the environment, which makes it advantageous to prioritize processing of negative information to respond quickly to it. But, most of such evolutionary explanations are coming from post-hoc analysis, which makes it less powerful as an observation. And like we discussed, the threatening nature of the negative stimuli depends on how relevant that is for optimal performance, in a task where valence information is irrelevant, it is difficult to make the assumption that it can be explained using the same evolutionary argument. However, the bias is still observed, which leads us to the question about heterogeneity in the bias observed and possible relation to individual traits.

5.4 Traits and Bias

We did not find any significant relationship between the coefficient of pleasant vs. unpleasant outcome condition for prior bias in the RL-DDM(Pedersen *et al.*,2017) model and any of the trait variables after running a forward model selection. However, the normalized variables for bias and average adaptive response for win-stay and lose-switch behaviour showed some notable trends. Importantly, the bias defined for win-stay behaviour showed a negative relation with the state and trait anxiety inventory(Spielberger *et al.*,1971) scores such that individuals who scored high on anxiety in this scale were associated with a reduced bias for win-stay strategy for unpleasant vs. pleasant outcome condition. Bar-Haim et al(Bar-Haim *et al.*,2007) investigating the magnitude of threat related attentional bias in relation to anxiety via a meta analysis of 112 studies found more attentional bias towards threatening stimuli for anxious individuals belonging to both clinical and non-clinical populations, irrespective of the kind of threatening stimuli used. Cisler et al(Cisler and Koster,2010), further suggest three possible mechanisms for the existence of such biases, which are the following.

- Attention being selective: individuals with anxiety disorder may pay special attention to threatening stimuli, which would result in an increased magnitude of negative bias
- Difficulty in disengaging attention from negative stimuli for people having anxiety disorder.
- Biased interpretation of something ambiguous being negative

Further Duncan et al(Duncan *et al.*,2017) using a neuroimaging study, investigated the possible neural correlates for such bias in relation to anxiety. And found that amygdala, which are expected to be processing emotional stimuli are hyper active in individuals with anxiety disorder, when compared to healthy controls. On the other hand, the anterior cingulate cortex, involved in attentional control and emotion regulation was less active in individuals with anxiety disorder when compared to healthy controls. Insula, known to be associated with processing introspective information has also been shown to have an increased activity compared to healthy controls. Overall these studies suggest that individuals with anxiety disorder may have difficulty in attention regulation and emotion control. Older studies such as Mathews et al(Mathews and MacLeod,1986) and Mogg et al(Mogg *et al.*,2004) points towards the same direction, where the anxiety measures are having a positive relation to the negative bias being observed.

However it is important to notice that the negative valence in most of these previous studies are of threats, which is task relevant, in terms of the performance in the task. Whereas, in a study like ours, the negative valence is not related to task performance and is more passive. It is possible that the bias-anxiety relation is modulated by task context. In addition, it is possible that lower bias in response strategy observed in individuals with high anxiety scores is due to generalization of unpleasant outcome valence to pleasant outcome trials in these individuals leading to lesser difference between the two trial conditions. This is consistent with a significant positive association between high STAI scores and high average adaptive response measure since individuals with high anxiety have been previously associated with indulging in high vigilance related adaptive changes in behavior (Brown *et al.*,2007).

In the case of lose switch behaviour, we found a significant main effect of pain catastrophizing scale (PCS-McNeil and Rainwater,1998;Osman *et al.*,1997;Sullivan *et al.*,2001), a measure that aims to quantify an individual's tendency to catastrophize in response to pain) on the bias. By catastrophization, what we mean is the tendency to magnify the significance of pain symptoms, ruminate on pain related thoughts, and to feel overwhelmed and helpless. In Quartana et al(Quartana *et al.*,2009) it was found that individuals scored high in PCS paid more attention to pain related information, which are supported by subsequent studies(Campbell *et al.*,2010;Blustein *et al.*,2010). However, it should be noted that these studies explicitly used pain related stimuli, and not generic negatively valenced ones. This implies the possibility that the relation of catastrophizing and response strategy bias may not be stimuli modality specific.

5.5 Role of Uncertainty and Individual traits in valence based bias in response strategy

In terms of the response times, there was no significant difference observed between the 70% vs. 80% correct feedback sessions. Higher uncertainty was associated with greater bias in win-stay behavior for unpleasant vs. pleasant outcome trials. Similarly, Lose-switch behaviour was found to have a positively affected in the mixed vs. no valence feedback blocks. This is in support of the argument put forward by Andersen et al(Anderson *et al.*,2019), which posit that uncertainty itself could be aversive.

Bias defined using the new coordinate system for win-stay behavior was absent in the 80% vs. 70% correct feedback sessions suggesting the causal contribution of expected uncertainty in eliciting a response strategy bias for unpleasant vs. pleasant outcome condition. The same holds true for the difference in prior bias parameter in the RL-DDM(Pedersen *et al.*,2017) model for the two tested levels of uncertainty. A recent meta analysis, Leu et al(Leung *et al.*,2022) suggest that in the case of anxiety and depression, there exists a positive relationship between uncertainty and cognitive biases including negative bias. The amplification of negative bias in depression and anxiety was particularly in relation to attention, interpretation and memory bias. And higher levels of anxiety and depression could further modulate the bias, regardless of the associated level of uncertainty. This made the authors to propose that the interaction between uncertainty and negative affect, could lead to a heightened sensitivity to negative affect. Interestingly, in the combined data across two levels of uncertainty, while there was no significant effect of STAI in the best model there was a significant negative interaction between uncertainty and STAI suggesting that with higher uncertainty (70% correct feedback blocks) the relationship between STAI and bias would become more negative. On the other hand, DSR(Olatunji *et al.*,2007;van Overveld *et al.*,2011) scores, quantifying the sensitivity to disgust was more positively correlated with the average adaptive response measure of lose switch behaviour in the 80% vs. 70% correct feedback sessions suggesting which could be because the negative stimulus being used are also disgusting in nature.

5.6 Conclusions

We report a bias in win-stay strategy in unpleasant vs pleasant outcome trials in the mixed valence vs. no valence feedback blocks of our experiment using

not only the newly derived normalized measure of bias but also in the prior bias parameter of the RL-DDM model. This is similar to a signal detection theory scenario where the detection threshold is shifted so as to not miss a potentially negative outcome, even when the valence information may not be relevant for optimal performance. The fact that such a bias was observed only in the mixed valence and not positive or negative valence alone blocks signifies the role of relative evaluation of outcome valence in this bias even though distinct effects were found in the reaction time analysis of individual valence blocks (and which was potentially canceled out in the mixed valence block). We ruled out the role of induced emotions in the observed reaction time differences by showing an absence of significant differences in the PANAS questionnaire scores implemented after negative and positive feedback blocks. Detection of response strategy bias only in the mixed valence block suggests a possible difference in strategy that an individual would apply in relation to an unpleasant vs. pleasant correct outcome as seen in the new dimension for bias (which unlike GLM does not make any assumptions about the independence of events across sequential trials) defined on win-stay behaviour.

Moreover, the bias defined both in terms of a new dimension to explain win-stay behaviour as well as observed as a shift in prior bias parameter of the RL-DDM model was found to be amplified with uncertainty, in line with our hypothesis as well as previous work allowing our results to be relevant to clinical populations having anxiety and depression disorders. Our results also makes an important suggestion regarding the use of ambiguious stimuli to test negativity bias since ambiguity is intrinsically associated with uncertainty which itself is typically perceived as aversive. Also, the fact that bias could be partially explained by trait measures and the trait-bias relationship itself is getting modulated by the uncertainty, makes it important to define the bias in the context of the population under consideration.

5.7 Limitations

The inferences that we have drawn in our study are likely restricted to our narrow subject pool of undergraduates in a university instead of being broadly generalizable to the population at large. The images that we used weren't matched for lower level visual features like luminosity, spatial frequency, colour or complexity, largely due to the difficulty in doing so across multiple variables and images matched for arousal were assumed to

be more relevant in demonstrating the role of valence in how affective feedback impacts performance strategy. Consistent with earlier work (Warriner *et al.*,2013;Gable and Harmon-Jones,2008), valence and motivation ratings in our data set were highly correlated, which could beg the question if the bias we report is due to the valence or motivational properties of the images used. However, it is also a separate theoretical question about whether motivation could be considered as an axis separate from valence and arousal. While older studies such as Lang et al(Lang *et al.*,1997b), suggest that we consider affect, activation and action to understand emotions, a more recent account by Lisa Barrett(Barrett,2006) argues that motivation is closely related but distinct from the core affect proposed by Russel et al(Russell,2003). Core affect is a more basic, non-conceptual experience of valence and arousal, which is applicable across emotions. But, motivation is referring to the process that is driving behaviour, which can either be pursuit of rewards or avoidance of punishment. Barret argues that the emotion could be arising from the interaction between core affect and motivation and hence be too tightly intervened to be studied in isolation.

5.8 Future Direction

We would like to investigate the mechanism driving the affect-uncertainty interaction further by studying the neural signals associated with cognitive control such as frontomedial theta(Cavanagh and Cohen,2022) in our task. Interestingly Wu et al(Wu *et al.*,2020) suggest mechanisms through which the cognitive control would get implemented under uncertainty, which also posit the possibility of cognitive control network(CCN) playing a more general role than previously expected. We also plan to quantify uncertainty independently by pupillometry(Brunyé and Gardony,2017) for which the task will be designed to control for luminosity associated with the valenced images. Further, even though responses to the intolerance of uncertainty questionnaire(Morriss *et al.*,2022a) have been found to correlate with anxiety (Carleton *et al.*,2007), we would like to explicitly test if it is able to explain our results better in future iterations of the experiment. Lastly, given that while pain-killers like acetaminophen have been shown to attenuate responses to valenced stimuli(DeWall *et al.*,2010;Durso *et al.*,2015) and uncertainty seems to amplify it, it would be interesting to use pharmacological interventions in our task to shed light on the mechanistic explanation of uncertainty's role in negativity bias.

References

Akaike H (1974). A new look at the statistical model identification. IEEE transactions on automatic control 19(6), 716–723.

Algom D, Chajut E, Lev S (2004). A rational look at the emotional stroop phenomenon: a generic slowdown, not a stroop effect. Journal of experimental psychology: General 133(3), 323.

Allman JM, Hakeem A, Erwin JM, Nimchinsky E, Hof P (2001). The anterior cingulate cortex: the evolution of an interface between emotion and cognition. Annals of the New York Academy of Sciences 935(1), 107–117.

Anderson EC, Carleton RN, Diefenbach M, Han PK (2019). The relationship between uncertainty and affect. Frontiers in psychology 10, 2504.

Angela JY, Dayan P (2005). Uncertainty, neuromodulation, and attention. Neuron 46(4), 681–692.

Bandyopadhyay D, Pammi VC, Srinivasan N (2019). Incidental positive emotion modulates neural response to outcome valence in a monetarily rewarded gambling task. Progress in Brain Research 247, 219–251.

Bar-Haim Y, Lamy D, Pergamin L, Bakermans-Kranenburg MJ, Van Ijzendoorn MH (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. Psychological bulletin 133(1), 1.

Barrett LF (2006). Solving the emotion paradox: Categorization and the experience of emotion. Personality and social psychology review 10(1), 20–46.

Bates D, Mächler M, Bolker B, Walker S (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67(1), 1–48.

Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD (2001). Bad is stronger than good. Review of general psychology 5(4), 323–370.

Beck AT, Steer RA, Carbin MG (1988). Psychometric properties of the beck depression inventory: Twenty-five years of evaluation. Clinical psychology review 8(1), 77–100.

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961). An inventory for measuring depression. Archives of general psychiatry 4(6), 561–571.

Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007). Learning the value of information in an uncertain world. Nature neuroscience 10(9), 1214–1221.

Berke JD (2018). What does dopamine mean? Nature neuroscience 21(6), 787–793.

Blustein DL, Murphy KA, Kenny ME, Jernigan M, Pérez-Gualdrón L, Castaneda T, Koepke M, Land M, Urbano A, Davis O (2010). Exploring urban students' constructions about school, work, race, and ethnicity. Journal of Counseling Psychology 57(2), 248.

Broad CD (1940). Sir arthur eddington's the philosophy of physical science1. Philosophy 15(59), 301–312.

Brown JS (1948). Gradients of approach and avoidance responses and their relation to level of motivation. Journal of comparative and physiological psychology 41(6), 450.

Brown LH, Silvia PJ, Myin-Germeys I, Kwapil TR (2007). When the need to belong goes wrong: The expression of social anhedonia and social anxiety in daily life. Psychological Science 18(9), 778–782.

Brunyé TT, Gardony AL (2017). Eye tracking measures of uncertainty during perceptual decision making. International Journal of Psychophysiology 120, 60–68.

Cacioppo JT, Berntson GG (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. Psychological bulletin 115(3), 401.

Cacioppo JT, Gardner WL, Berntson GG (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. Personality and Social Psychology Review 1(1), 3–25.

Campbell CM, Kronfli T, Buenaver LF, Smith MT, Berna C, Haythornthwaite JA, Edwards RR (2010). Situational versus dispositional measurement of catastrophizing: associations with pain responses in multiple samples. The Journal of Pain 11(5), 443–453.

Carleton RN (2016). Fear of the unknown: One fear to rule them all? Journal of anxiety disorders 41, 5–21.

Carleton RN, Norton MPJ, Asmundson GJ (2007). Fearing the unknown: A short version of the intolerance of uncertainty scale. Journal of anxiety disorders 21(1), 105–117.

Carmichael ST, Price J (1996). Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. Journal of Comparative Neurology 371(2), 179–207.

Cavanagh JF, Cohen MX (2022). Frontal midline theta as a model specimen of cortical theta. The Oxford Handbook of EEG Frequency 178.

Cisler JM, Koster EH (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. Clinical psychology review 30(2), 203–216.

David JS, Nicola G, Bradley PC, Angelika vdL (2002). Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64(4), 583–639.

DeBruine L, Jones B (2018). Determining the number of raters for reliable mean ratings. Open Science Framework .

DeWall CN, MacDonald G, Webster GD, Masten CL, Baumeister RF, Powell C, Combs D, Schurtz DR, Stillman TF, Tice DM, *et al.* (2010). Acetaminophen reduces social pain: Behavioral and neural evidence. Psychological science 21(7), 931–937.

Doya K (2008). Modulators of decision making. Nature neuroscience 11(4), 410–416.

Duff MC, Kurczek J, Rubin R, Cohen NJ, Tranel D (2013). Hippocampal amnesia disrupts creative thinking. Hippocampus 23(12), 1143–1149.

Duncan L, Yilmaz Z, Gaspar H, Walters R, Goldstein J, Anttila V, Bulik-Sullivan B, Ripke S, of the Psychiatric Genomics Consortium EDWG,

Thornton L, *et al.* (2017). Significant locus and metabolic genetic correlations revealed in genome-wide association study of anorexia nervosa. American journal of psychiatry 174(9), 850–858.

Durso GR, Luttrell A, Way BM (2015). Over-the-counter relief from pains and pleasures alike: Acetaminophen blunts evaluation sensitivity to both negative and positive stimuli. Psychological science 26(6), 750–758.

Edwards LC, Pearce SA, Turner-Stokes L, Jones A (1992). The pain beliefs questionnaire: An investigation of beliefs in the causes and consequences of pain. Pain 51(3), 267–272.

Eppinger B, Herbert M, Kray J (2010). We remember the good things: Age differences in learning and memory. Neurobiology of Learning and Memory 93(4), 515–521.

Eppinger B, Kray J (2011). To choose or to avoid: age differences in learning from positive and negative feedback. Journal of Cognitive Neuroscience 23(1), 41–52.

Eppinger B, Schuck NW, Nystrom LE, Cohen JD (2013). Reduced striatal responses to reward prediction errors in older compared with younger adults. Journal of Neuroscience 33(24), 9905–9912.

Ereira S, Pujol M, Guitart-Masip M, Dolan RJ, Kurth-Nelson Z (2021). Overcoming pavlovian bias in semantic space. Scientific reports 11(1), 3416.

Estes Z, Verges M (2008). Freeze or flee? negative stimuli elicit selective responding. Cognition 108(2), 557–565.

Ferrari SF, Ferrari MAL (1990). Predator avoidance behaviour in the buffy-headed marmoset, callithrix flaviceps. Primates 31, 323–338.

Forstmann BU, Ratcliff R, Wagenmakers EJ (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. Annual review of psychology 67, 641–666.

Fox E, Russo R, Bowles R, Dutton K (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? Journal of experimental psychology: General 130(4), 681.

Frank MJ, Seeberger LC, O'reilly RC (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306(5703), 1940–1943.

Gable PA, Harmon-Jones E (2008). Approach-motivated positive affect reduces breadth of attention. Psychological Science 19(5), 476–482.

Gelman A, Carlin JB, Stern HS, Rubin DB (1995). Bayesian data analysis. Chapman and Hall/CRC.

Gelman A, Hill J (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge university press.

Goodyear MD, Krleza-Jeric K, Lemmens T (2007). The declaration of helsinki.

Gray V, Douglas KM, Porter RJ (2021). Emotion processing in depression and anxiety disorders in older adults: systematic review. BJPsych Open 7(1), e7.

Green RE, Goldfried MR (1965). On the bipolarity of semantic space. Psychological Monographs: General and Applied 79(6), 1.

Haberkamp A, Glombiewski JA, Schmidt F, Barke A (2017). The disgust-related-images (dirti) database: Validation of a novel standardized set of disgust pictures. Behaviour research and therapy 89, 86–94.

Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD (2016). Mesolimbic dopamine signals the value of work. Nature neuroscience 19(1), 117–126.

Harding EJ, Paul ES, Mendl M (2004). Cognitive bias and affective state. Nature 427(6972), 312–312.

Hershberger WA (1986). An approach through the looking-glass. Animal Learning & Behavior 14(4), 443–451.

Hirsh JB, Mar RA, Peterson JB (2012). Psychological entropy: a framework for understanding uncertainty-related anxiety. Psychological review 119(2), 304.

Hochberg Y (1988). A sharper bonferroni procedure for multiple tests of significance. Biometrika 75(4), 800–802.

Hojat M, Shapurian R, Mehryar AH (1986). Psychometric properties of a persian version of the short form of the beck depression inventory for iranian college students. Psychological reports 59(1), 331–338.

Johnson EJ, Tversky A (1983). Affect, generalization, and the perception of risk. Journal of personality and social psychology 45(1), 20.

Kahneman D, Tversky A (1984). Choices, values, and frames. American psychologist 39(4), 341.

Kim H, Anderson BA (2022). Primary rewards and aversive outcomes have comparable effects on attentional bias. Behavioral Neuroscience .

Kim H, Somerville LH, Johnstone T, Alexander AL, Whalen PJ (2003). Inverse amygdala and medial prefrontal cortex responses to surprised faces. Neuroreport 14(18), 2317–2322.

Koban L, Corradi-Dell'Acqua C, Vuilleumier P (2013). Integration of error agency and representation of others' pain in the anterior insula. Journal of Cognitive Neuroscience 25(2), 258–272.

Konishi S, Kitagawa G (2008). Information criteria and statistical modeling .

Kroenke K, Spitzer RL (2002). The phq-9: a new depression diagnostic and severity measure.

Kurdi B, Lozano S, Banaji MR (2017). Introducing the open affective standardized image set (oasis). Behavior research methods 49, 457–470.

Kurtz JL, Wilson TD, Gilbert DT (2007). Quantity versus uncertainty: When winning one prize is better than winning two. Journal of Experimental Social Psychology 43(6), 979–985.

Kurz HD (2016). Adam smith on markets, competition and violations of natural liberty. Cambridge Journal of Economics 40(2), 615–638.

Lagisz M, Zidar J, Nakagawa S, Neville V, Sorato E, Paul ES, Bateson M, Mendl M, Løvlie H (2020). Optimism, pessimism and judgement bias in animals: a systematic review and meta-analysis. Neuroscience & Biobehavioral Reviews 118, 3–17.

Lang PJ, Bradley MM, Cuthbert BN, *et al.* (1997a). International affective picture system (iaps): Technical manual and affective ratings. NIMH Center for the Study of Emotion and Attention 1(39-58), 3.

Lang PJ, Bradley MM, Cuthbert BN, *et al.* (1997b). Motivated attention: Affect, activation, and action. Attention and orienting: Sensory and motivational processes 97, 135.

Lerner JS, Keltner D (2001). Fear, anger, and risk. Journal of personality and social psychology 81(1), 146.

Leung CJ, Yiend J, Trotta A, Lee TM (2022). The combined cognitive bias hypothesis in anxiety: A systematic review and meta-analysis. Journal of Anxiety Disorders 89, 102575.

Lewin K (1936). A dynamic theory of personality: Selected papers. The Journal of Nervous and Mental Disease 84(5), 612–613.

Lindquist KA, Satpute AB, Wager TD, Weber J, Barrett LF (2016). The brain basis of positive and negative affect: evidence from a meta-analysis of the human neuroimaging literature. Cerebral cortex 26(5), 1910–1922.

Loewenstein GF, Weber EU, Hsee CK, Welch N (2001). Risk as feelings. Psychological bulletin 127(2), 267.

Macmillan NA, Creelman CD (2004). Detection theory: A user's guide. Psychology press.

Mathews A, MacLeod C (1986). Discrimination of threat cues without awareness in anxiety states. Journal of abnormal psychology 95(2), 131.

McNeil DW, Rainwater AJ (1998). Development of the fear of pain questionnaire-iii. Journal of behavioral medicine 21, 389–410.

Mellers B, Schwartz A, Ritov I (1999). Emotion-based choice. Journal of Experimental Psychology: General 128(3), 332.

Mendl M, Burman OH, Parker RM, Paul ES (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. Applied Animal Behaviour Science 118(3-4), 161–181.

Miller NE (1959). Liberalization of basic sr concepts: Extensions to conflict behavior, motivation and social learning. Psychology: A study of a science, Study 1, 196–292.

Minsky M (1961). Steps toward artificial intelligence. Proceedings of the IRE 49(1), 8–30.

Mogg K, Philippot P, Bradley BP (2004). Selective attention to angry faces in clinical social phobia. Journal of abnormal psychology 113(1), 160.

Mohebi A, Pettibone JR, Hamid AA, Wong JMT, Vinson LT, Patriarchi T, Tian L, Kennedy RT, Berke JD (2019). Dissociable dopamine dynamics for learning and motivation. Nature 570(7759), 65–70.

Morriss J, Bradford DE, Wake S, Biagi N, Tanovic E, Kaye JT, Joormann J (2022a). Intolerance of uncertainty and physiological responses during instructed uncertain threat: a multi-lab investigation. Biological psychology 167, 108223.

Morriss J, Tupitsa E, Dodd HF, Hirsch CR (2022b). Uncertainty makes me emotional: Uncertainty as an elicitor and modulator of emotional states. Frontiers in Psychology 13.

Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. Nature neuroscience 15(7), 1040–1046.

Nassar MR, Wilson RC, Heasly B, Gold JI (2010). An approximately bayesian delta-rule model explains the dynamics of belief updating in a changing environment. Journal of Neuroscience 30(37), 12366–12378.

Norris CJ (2021). The negativity bias, revisited: Evidence from neuroscience measures and an individual differences approach. Social neuroscience 16(1), 68–82.

Norris CJ, Gollan J, Berntson GG, Cacioppo JT (2010). The current status of research on the structure of evaluative space. Biological psychology 84(3), 422–436.

Olatunji BO, Sawchuk CN, De Jong PJ, Lohr JM (2007). Disgust sensitivity and anxiety disorder symptoms: Psychometric properties of the disgust emotion scale. Journal of Psychopathology and Behavioral Assessment 29, 115–124.

Osman A, Barrios FX, Kopper BA, Hauptmann W, Jones J, O'Neill E (1997). Factor structure, reliability, and validity of the pain catastrophizing scale. Journal of behavioral medicine 20, 589–605.

Palminteri S, Justo D, Jauffret C, Pavlicek B, Dauta A, Delmaire C, Czernecki V, Karachi C, Capelle L, Durr A, *et al.* (2012). Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. Neuron 76(5), 998–1009.

Parisi N, Katz I (1986). Attitudes toward posthumous organ donation and commitment to donate. Health Psychology 5(6), 565.

Paul K, Pourtois G (2017). Mood congruent tuning of reward expectation in positive mood: evidence from frn and theta modulations. Social Cognitive and Affective Neuroscience 12(5), 765–774.

Paulus MP, Angela JY (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. Trends in cognitive sciences 16(9), 476–483.

Pedersen ML, Frank MJ, Biele G (2017). The drift diffusion model as the choice rule in reinforcement learning. Psychonomic bulletin & review 24, 1234–1251.

Peters A, McEwen BS, Friston K (2017). Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. Progress in neurobiology 156, 164–188.

Quartana PJ, Campbell CM, Edwards RR (2009). Pain catastrophizing: a critical review. Expert review of neurotherapeutics 9(5), 745–758.

Rakos RF, Laurene KR, Skala S, Slane S (2008). Belief in free will: Measurement and conceptualization innovations. Behavior and Social Issues 17, 20–40.

Ramsey FP (2016). Truth and probability. Readings in Formal Epistemology: Sourcebook 21–45.

Rangel A, Camerer C, Montague PR (2008). A framework for studying the neurobiology of value-based decision making. Nature reviews neuroscience 9(7), 545–556.

Ratcliff R, McKoon G (2008). The diffusion decision model: theory and data for two-choice decision tasks. Neural computation 20(4), 873–922.

Reinen JM, Whitton AE, Pizzagalli DA, Slifstein M, Abi-Dargham A, McGrath PJ, Iosifescu DV, Schneier FR (2021). Differential reinforcement learning responses to positive and negative information in unmedicated individuals with depression. European Neuropsychopharmacology 53, 89–100.

Rescorla RA (1988). Behavioral studies of pavlovian conditioning. Annual review of neuroscience 11(1), 329–352.

Rozin P, Royzman EB (2001). Negativity bias, negativity dominance, and contagion. Personality and social psychology review 5(4), 296–320.

Ruscheweyh R, Marziniak M, Stumpenhorst F, Reinholz J, Knecht S (2009). Pain sensitivity can be assessed by self-rating: development and validation of the pain sensitivity questionnaire. Pain 146(1-2), 65–74.

Russell JA (2003). Core affect and the psychological construction of emotion. Psychological review 110(1), 145.

Scherer KR (1999). Appraisal theory. .

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

Searcy YM, Caine NG (2003). Hawk calls elicit alarm and defensive reactions in captive geoffroy's marmosets (callithrix geoffroyi). Folia Primatologica 74(3), 115–125.

Sidman M, Center DB (1991). Coercion and its fallout. Behavioral Disorders 16(4), 315–317.

Simon HA (1966). Theories of decision-making in economics and behavioural science. Springer.

Simpson JR, Öngür D, Akbudak E, Conturo TE, Ollinger JM, Snyder AZ, Gusnard DA, Raichle ME (2000). The emotional modulation of cognitive processing: an fmri study. Journal of Cognitive Neuroscience 12(Supplement 2), 157–170.

Slovic P, Finucane ML, Peters E, MacGregor DG (2007). The affect heuristic. European journal of operational research 177(3), 1333–1352.

Smith NK, Cacioppo JT, Larsen JT, Chartrand TL (2003). May i have your attention, please: Electrocortical responses to positive and negative stimuli. Neuropsychologia 41(2), 171–183.

Sokol-Hessner P, Camerer CF, Phelps EA (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. Social cognitive and affective neuroscience 8(3), 341–350.

Spielberger C, Gorsuch R, Lushene P, Vagg P, Jacobs G (1983). Manual for the state-trait anxiety inventory (form y). redwood city. CA: Mind Garden, Inc .

Spielberger CD, Gonzalez-Reigosa F, Martinez-Urrutia A, Natalicio LF, Natalicio DS (1971). The state-trait anxiety inventory. Revista Interamericana de Psicologia/Interamerican journal of psychology 5(3 & 4).

Starling MJ, Branson N, Cody D, McGreevy PD (2013). Conceptualising the impact of arousal and affective state on training outcomes of operant conditioning. Animals 3(2), 300–317.

Steer RA, Rissmiller DJ, Beck AT (2000). Use of the beck depression inventory-ii with depressed geriatric inpatients. Behaviour research and therapy 38(3), 311–318.

Stone M (1960). Models for choice-reaction time. Psychometrika 25(3), 251–260.

Sullivan MJ, Thorn B, Haythornthwaite JA, Keefe F, Martin M, Bradley LA, Lefebvre JC (2001). Theoretical perspectives on the relation between catastrophizing and pain. The Clinical journal of pain 17(1), 52–64.

Sutton RS, Barto AG, *et al.* (1998). Introduction to reinforcement learning, volume 135. MIT press Cambridge.

van de Vijver I, Ridderinkhof KR, de Wit S (2015). Age-related changes in deterministic learning from positive versus negative performance feedback. Aging, Neuropsychology, and Cognition 22(5), 595–619.

van Overveld M, de Jong PJ, Peters ML, Schouten E (2011). The disgust scale-r: A valid and reliable index to investigate separate disgust domains? Personality and Individual Differences 51(3), 325–330.

Verburg M, Snellings P, Zeguers MHT, Huizenga HM (2019). Positive-blank versus negative-blank feedback learning in children and adults. Quarterly Journal of Experimental Psychology 72(4), 753–763.

Wagenmakers EJ, Farrell S (2004). Aic model selection using akaike weights. Psychonomic bulletin & review 11, 192–196.

Wagenmakers EJ, Ratcliff R, Gomez P, McKoon G (2008). A diffusion model account of criterion shifts in the lexical decision task. Journal of memory and language 58(1), 140–159.

Warriner AB, Kuperman V, Brysbaert M (2013). Norms of valence, arousal, and dominance for 13,915 english lemmas. Behavior research methods 45, 1191–1207.

Watson D, Clark LA, Tellegen A (1988). Development and validation of brief measures of positive and negative affect: the panas scales. Journal of personality and social psychology 54(6), 1063.

Wiecki TV, Sofer I, Frank MJ (2013). Hddm: Hierarchical bayesian estimation of the drift-diffusion model in python. Frontiers in neuroinformatics 14.

Williams N (2014). The gad-7 questionnaire. Occupational medicine 64(3), 224–224.

Worthy DA, Hawthorne MJ, Otto AR (2013). Heterogeneity of strategy use in the iowa gambling task: A comparison of win-stay/lose-shift and reinforcement learning models. Psychonomic bulletin & review 20, 364–371.

Wu T, Spagna A, Chen C, Schulz KP, Hof PR, Fan J (2020). Supramodal mechanisms of the cognitive control network in uncertainty processing. Cerebral Cortex 30(12), 6336–6349.

Xu S, Sun Y, Huang M, Huang Y, Han J, Tang X, Ren W (2021). Emotional state and feedback-related negativity induced by positive, negative, and combined reinforcement. Frontiers in Psychology 12, 647263.

Chapter 6

Additional data and Codes

6.1 Codes

Please visit: `https://drive.google.com/drive/folders/1GHMU7-H8OSbgxsT9FPb9tKuS4BGnRlya?usp=sharing`

6.2 Forward Model Selection For RL-DDM models

Forward Selection Model For RL-DDM in "VAL" block			
S.No	Parameters estimated separately for conditions	DIC	Δ_{DIC}
1	Baseline	3117.394	26.324
2	$[z]$	3139.814	48.744
3	$[\alpha]$	3097.650	6.58
4	$[m]$	3091.070	0
5	$[a]$	3153.781	62.711
6	$[m, \alpha]$	3094.190	3.12
7	$[z, \alpha]$	3105.231	14.161
8	$[m, z]$	3096.817	5.747
9	$[v, m, \alpha]$	3098.870	7.8

Table 6.1: *Forward model selection table for RL-DDM in* 70% *correct feedback blocks*

Forward Selection Model For RL-DDM in "VAL" and "SYM" block				
S.No	depends on condition	depends on type	DIC	Δ_{DIC}
0	Baseline	Baseline	6445.525	442.345
1	$[z]$	Baseline	6446.598	443.418
2	$[z]$	$[z]$	6257.006	253.826
3	$[m]$	Baseline	6359.103	255.923
4	Baseline	$[m]$	6249.539	246.395
5	$[m]$	$[m]$	6152.601	149.421
6	$[\alpha]$	Baseline	6369.231	366.051
7	Baseline	$[\alpha]$	6243.977	240.797
8	$[\alpha]$	$[\alpha]$	6157.103	153.923
9	Baseline	$[z]$	6240.940	237.76
10	$[z]$	$[a]$	6151.906	148.726
11	Baseline	$[a, z]$	6134.962	131.782
12	$[z]$	$[a, z]$	6133.003	129.823
13	$[m]$	$[a]$	6064.686	61.506
14	Baseline	$[a, m]$	6115.413	112.233
15	$[m]$	$[a, m]$	6017.908	14.728
16	$[\alpha]$	$[a]$	6066.323	63.143
17	Baseline	$[a, \alpha]$	6111.020	107.84
18	$[\alpha]$	$[a, \alpha]$	6027.356	24.176
19	Baseline	$[a]$	6148.259	145.079
20	$[a]$	Baseline	6460.116	456.936
21	$[a]$	$[a]$	6175.853	172.673
22	$[m, z]$	$[a, m]$	6016.073	12.893
23	$[m]$	$[a, m, z]$	6007.198	4.018
24	$[m, z]$	$[a, m, z]$	6008.323	5.143
25	$[m, \alpha]$	$[a, m]$	6020.355	17.175
26	$[m]$	$[a, m, \alpha]$	6013.322	10.142
27	$[m, \alpha]$	$[a, m, \alpha]$	6024.885	21.705
28	$[m, z, \alpha]$	$[a, m, z]$	6011.946	8.766
29	$[m, z]$	$[a, m, z, \alpha]$	6003.180	0
30	$[m, z, \alpha]$	$[a, m, z, \alpha]$	6011.548	8.368

Table 6.2: *Forward model selection table for RL-DDM in* 70% *and* 80% *correct feedback blocks*

We were trying to see what all parameters should be fitted separately for positive and negative conditions. When done the forward modeling on mixed feedback blocks("VAL"), it was observed that model fitting scaling fac-

tor separately for each condition(Table 6.1), it was having the lowest DIC. However, when including z and α, the DIC difference(Δ_{DIC}) was less than 10, meaning that there is a substantial evidence in support of the more complex model. Also, a model fitting α and m separately for each condition has $\Delta_{DIC}<5$, suggesting even stronger evidence for the more complex model.
However, when fitted a model for combined data from mixed and symbolic(Table 6.2) blocks, we found that the best model varies m and z with condition. In principle, fitting the model on the combined data, or separately fitting them won't make any difference as they belong to same set of participants. But, when α is included now for the model on combined data set, the DIC difference is again less than 10. Thus we decided to take the more complex model, which has α, m, and z fitted independently for each condition and all variables fitted independently for block types.

6.3 Posterior Predictive Checks For RL-DDM

In Bayesian statistics, the posterior predictive probability is given by:

$$p(\hat{y_p}|y) = \int_\theta p(\hat{y_p}, \theta|y)d\theta \tag{6.1}$$

Let's consider the conditional probability inside the integral

$$p(\hat{y_p}, \theta|y) = \frac{p(\hat{y_p} \cap \theta \cap y)}{p(y)} = \frac{p(\hat{y_p} \cap \theta \cap y)}{p(\theta \cap y)} . \frac{p(\theta \cap y)}{p(y)} = p(\hat{y_p}|\theta, y) \cdot p(\theta|y) \tag{6.2}$$

Using this, we can rewrite 6.1 as

$$p(\hat{y_p}|y) = \int_\theta p(\hat{y_p}|\theta, y) \cdot p(\theta|y)d\theta \tag{6.3}$$

The second probability term in the integral is posterior probability, given by

$$p(\theta|y) = \frac{p(y|\theta}{p(y)} p(\theta) \tag{6.4}$$

Which is the posterior probability distribution or the distribution, estimated using Markov Chain Monte Carlo method, and all chains converged(Gelman Rubin statistics >1.01).
Let's define the probability of a posterior predictive value being less than or equal to a particular value

$$F(y_i|y) = p(y_i > \hat{y_p}|y) = \sum_{\hat{y_p}<y_i} \int_\theta p(\hat{y_p}, \theta|y)d\theta \tag{6.5}$$

Then if we are setting the $\alpha = 0.05$ the upper and lower limits of the high density interval will give

$$F(y_L|y) = \frac{\alpha}{2} \tag{6.6}$$

$$F(y_U|y) = 1 - \frac{\alpha}{2} \tag{6.7}$$

And we get

$$\sum_{\hat{y_p}=y_L}^{y_U} p(\hat{y_p}|y) = 1 - \alpha \tag{6.8}$$

For each type and condition, we divided the data into 10 bins and computed the y_L and y_U seperately for each. For any given bin, if the observed value lies within the range $[y_L, y_U]$(denoted by error bars), we can say that the predictive distribution of the model is well capturing the data, which is desirable.

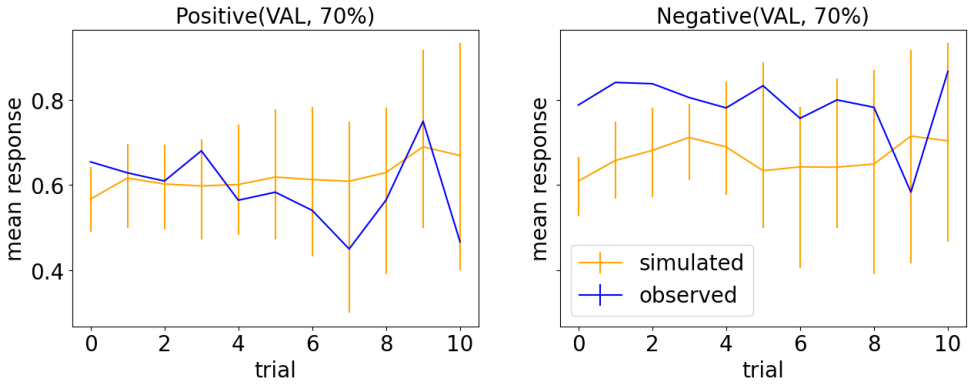


Figure 6.1: *Posterior predictive values compared against observed data, for mixed feedback block with 70% correct feedback. Error bar in yellow gives the highest density intervals for the predictions made at each bin. We can observe that there is a discrepancy in the early bins. However, otherwise the predictions and the data are closely matching. Mean response is plotted against trial bin.*

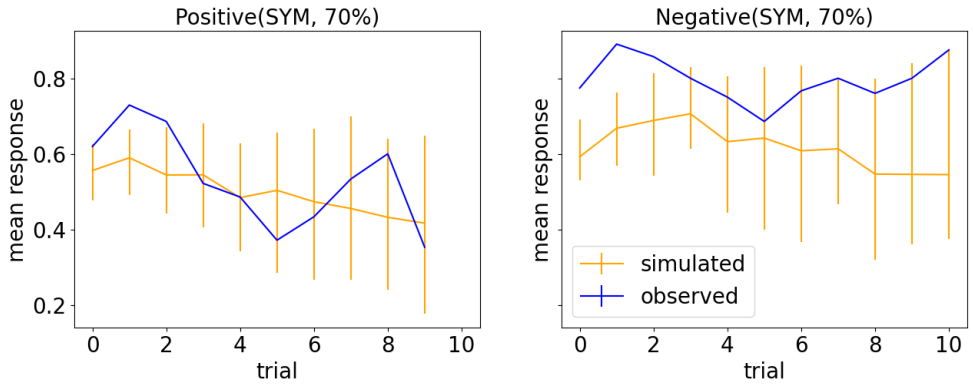


Figure 6.2: *Posterior predictive values compared against observed data, for symbolic feedback block with 70% correct feedback. Error bar in yellow gives the highest density intervals for the predictions made at each bin. We can observe that there is a discrepancy in the early bins. However, otherwise the predictions and the data are closely matching. Negative and positive feedback condition is just denoting different shapes used, and has not valence. Mean response is plotted against trial bin.*

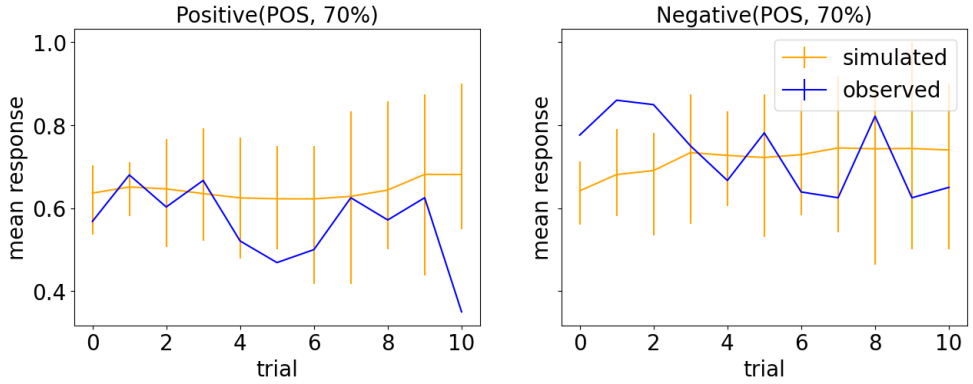


Figure 6.3: *Posterior predictive values compared against observed data, for positive feedback block with 70% correct feedback. Error bar in yellow gives the highest density intervals for the predictions made at each bin. We can observe that there is a discrepancy in the early bins. However, otherwise the predictions and the data are closely matching. In the block there was either a positive feedback(image) or a shape, depending on the association it has with the Japanese alphabet. Mean response is plotted against trial bin.*

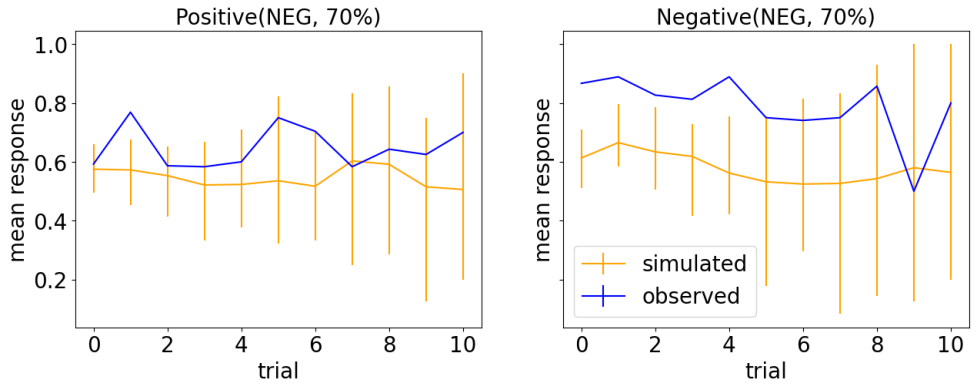


Figure 6.4: *Posterior predictive values compared against observed data, for negative feedback block with 70% correct feedback. Error bar in yellow gives the highest density intervals for the predictions made at each bin. We can observe that there is a discrepancy in the early bins. However, otherwise the predictions and the data are closely matching. In the block there was either a negative feedback(image) or a shape, depending on the association it has with the Japanese alphabet. Mean response is plotted against trial bin.*

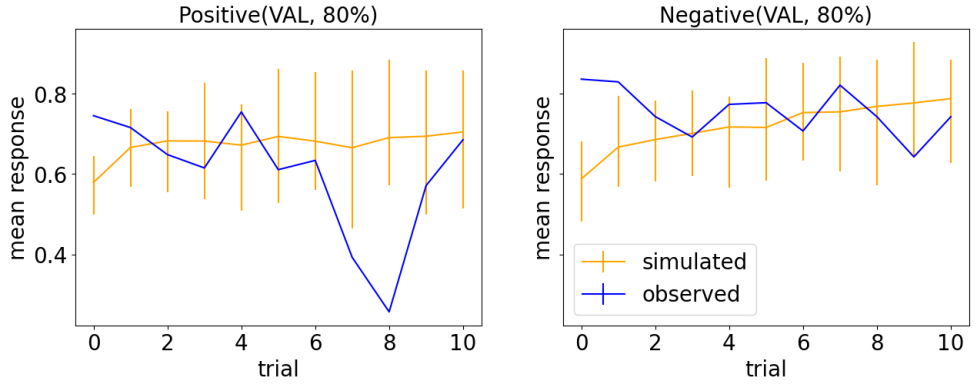


Figure 6.5: *Posterior predictive values compared against observed data, for mixed feedback block with 80% correct feedback. Error bar in yellow gives the highest density intervals for the predictions made at each bin. We can observe that there is a discrepancy in the early bins, and towards later bins in positive condition, which seems like a switching. However, otherwise the predictions and the data are closely matching. Mean response is plotted against trial bin.*

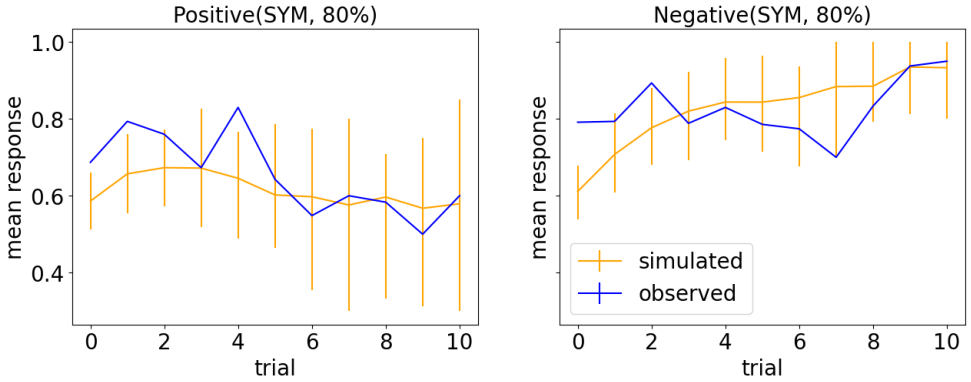


Figure 6.6: *Posterior predictive values compared against observed data, for symbolic feedback block with 80% correct feedback. Error bar in yellow gives the highest density intervals for the predictions made at each bin. We can observe that there is a discrepancy in the early bins. However, otherwise the predictions and the data are closely matching. Negative and positive feedback condition is just denoting different shapes used, and has not valence. Mean response is plotted against trial bin.*

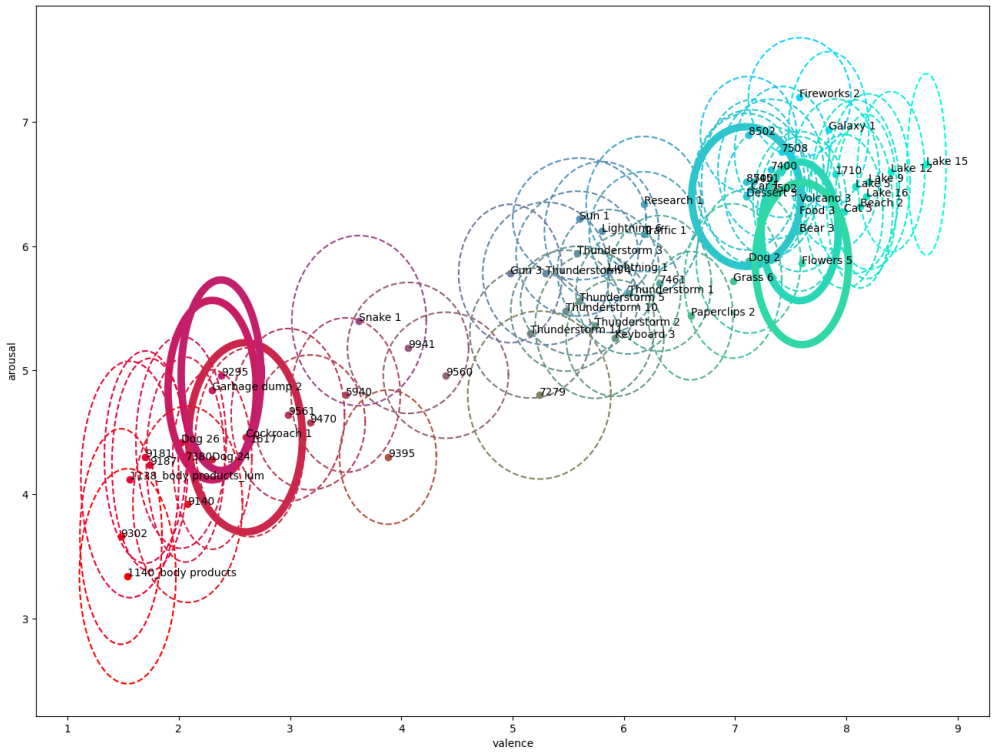


Figure 6.7: *valence-arousal plane. X axis is valence and Y axis is arousal. Ellipse axes, which are standard error in valence(along x axis) and standard error in arousal(along y axis) and the center of the ellipse denotes mean value of valence and arousal for that particular picture. The thicker ellipses represents the images picked for the experiment.*

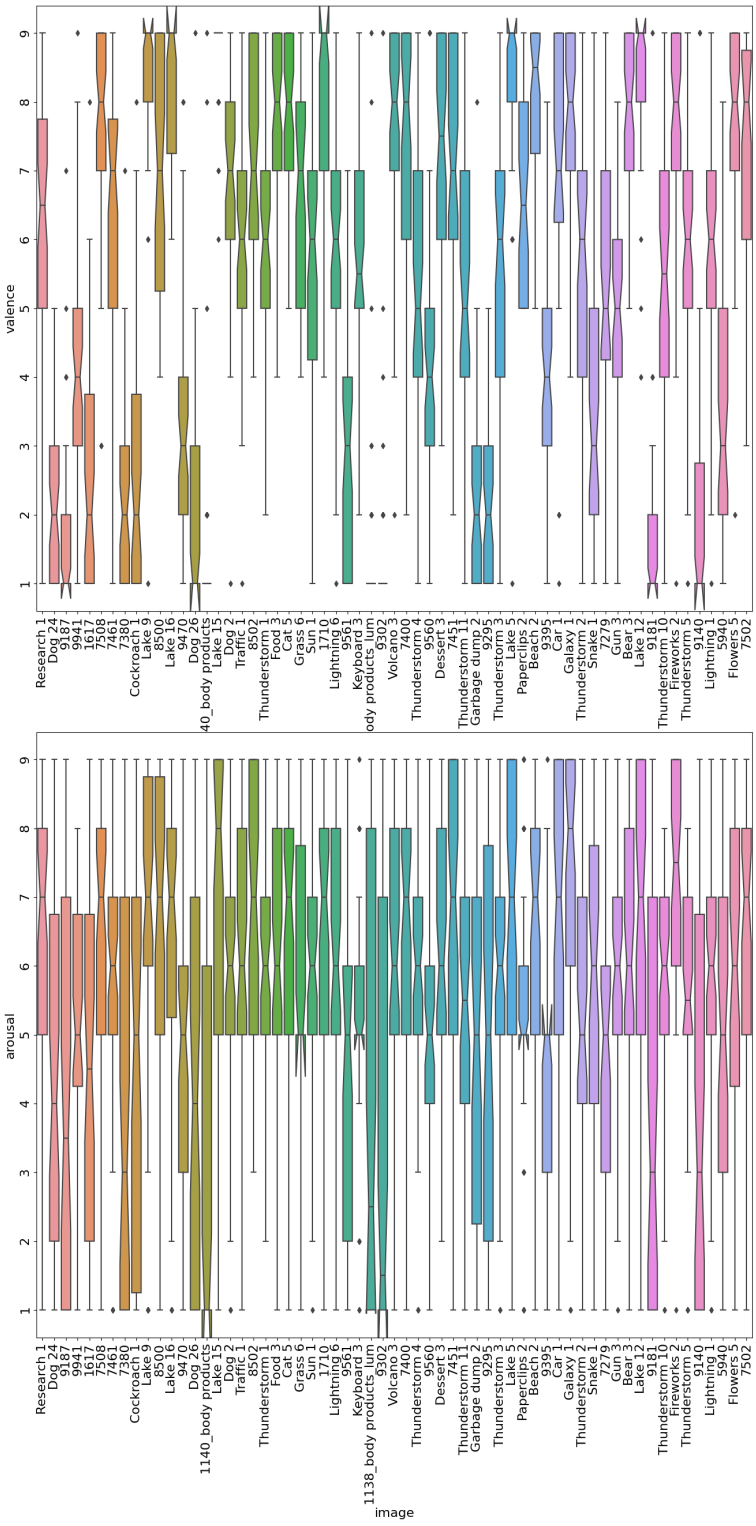


Figure 6.8: *Valence and arousal rating for all the images*

GLMs for RT- 70 percentage correct feedback forward model selection			
S.No	Equation	AIC	Δ_{AIC}
1	$rt \sim 1 + (1\|subj_i dx)$	8284.711	209.1
2	$rt \sim 1 + win_{t-1} + (1\|subj_i dx)$	8289.753	214.14
3	$rt \sim 1 + stay_t + (1\|subj_i dx)$	8290.12	214.51
4	$rt \sim 1 + condition + (1\|subj_i dx)$	8293.729	218.12
5	$rt \sim 1 + bin + (1\|subj_i dx)$	8140.66	65.05
6	$rt \sim 1 + type + (1\|subj_i dx)$	8199.62	124.01
7	$rt \sim 1 + bin + win_{t-1} + (1\|subj_i dx)$	8146.28	70.67
8	$rt \sim 1 + bin + stay_t + (1\|subj_i dx)$	8146.68	71.07
9	$rt \sim 1 + bin + condition + (1\|subj_i dx)$	8149.70	74.09
10	$rt \sim 1 + bin + type + (1\|subj_i dx)$	**8075.609**	**0**
11	$rt \sim 1 + bin + type + win_{t-1} + (1\|subj_i dx)$	8080.04	4.43
12	$rt \sim 1 + bin + type + stay_t + (1\|subj_i dx)$	8080.38	4.77
13	$rt \sim 1 + bin + type + condition + (1\|subj_i dx)$	8084.65	5.04
14	$rt \sim 1 + bin + type + block : type + (1\|subj_i dx)$	8094.82	19.21

Table 6.3: *Forward model selection table for response time in* 70% *correct feedback blocks*

GLMs for RT- 70% and 80% correct feedback forward model selection			
S.No	Equation	AIC	Δ_{AIC}
1	$rt \sim 1 + (1\|subj_i dx)$	5622.218	52.065
2	$rt \sim win_{t-1} + (1\|subj_i dx)$	5624.855	54.702
3	$rt \sim condition + (1\|subj_i dx)$	5630.811	60.658
4	$rt \sim bin + (1\|subj_i dx)$	**5570.153**	0
5	$rt \sim type + (1\|subj_i dx)$	5627.508	57.355
6	$rt \sim prob - incorr + (1\|subj_i dx)$	5626.107	55.954
7	$rt \sim bin + win_{t-1} + (1\|subj_i dx)$	5573.343	3.19
8	$rt \sim bin + condition + (1\|subj_i dx)$	5578.749	8.596
9	$rt \sim bin + type + (1\|subj_i dx)$	5573.776	3.623
10	$rt \sim bin + prob - incorr + (1\|subj_i dx)$	5574.264	4.111

Table 6.4: *Forward model selection table for response time in* 70% *and* 80% *correct feedback blocks*

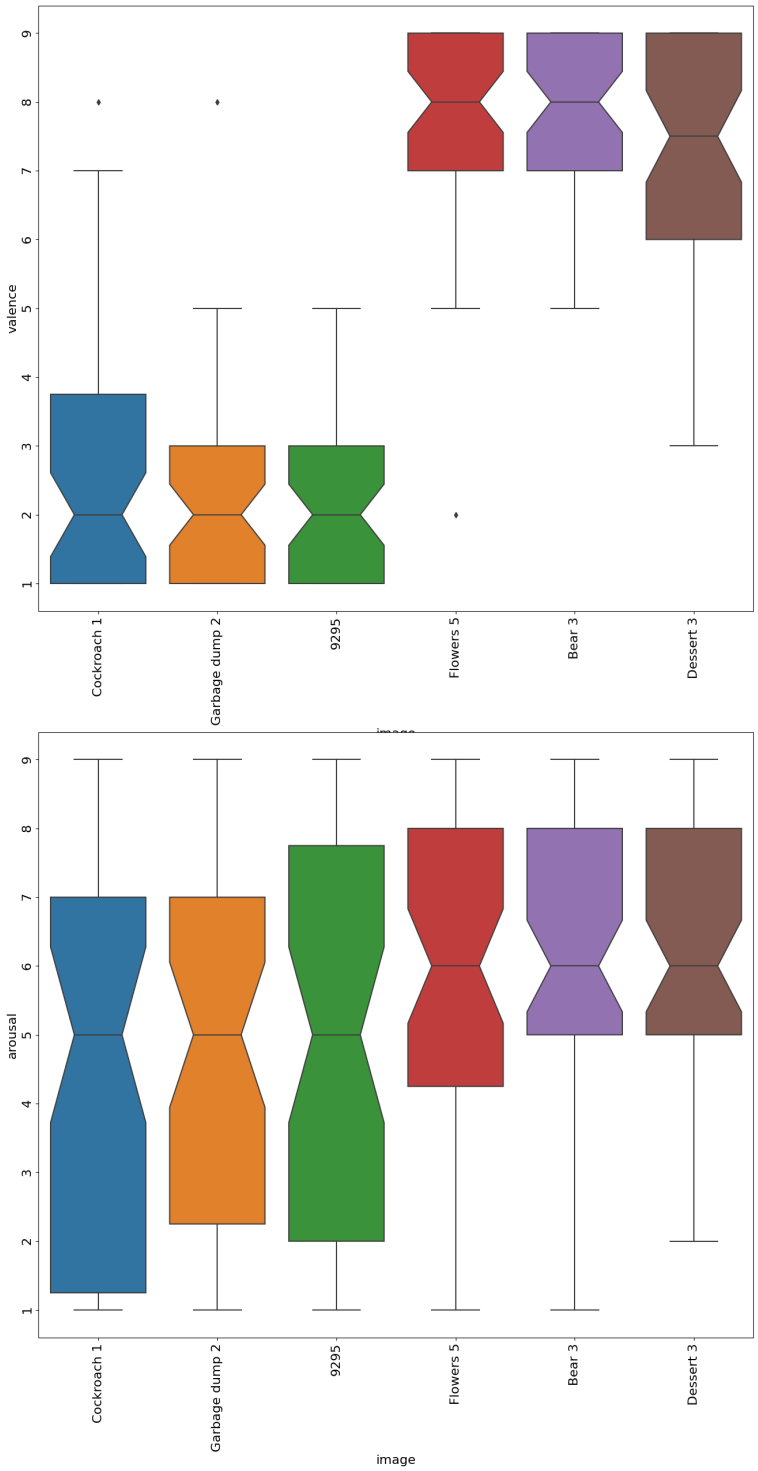


Figure 6.9: *Valence and Arousal Ratings of the Stimuli used for the task*