Extending the Drift Diffusion Model to the Cognitive Realm

Hriday Nayak *

February 28, 2021

Abstract

The drift diffusion model has had great success as a computational model to study the underlying processes of sensory and value-based decision making and the diffusion process may actually mimic how the brain integrates evidence and makes decisions. Here I review recent applications and extensions of the drift diffusion model to self-control, loss aversion, driving behaviour, racial biases and reinforcement learning with the aim of finding out whether the model is applicable to more cognitive tasks

1 Introduction

Decisions are ubiquitous. Every day, we make thousands of decisions, ranging from automatic low-level tasks like whether or not to look at a section of the screen to high-level tasks which require more deliberation like which movie to watch. In the last thirty years, major progress has been made in understanding how decisions are made in the brain.To explain any aspect of decision making, experiments are designed and then computational models are built on this experimental data. To make the models more cohesive and increase their explanatory power, brain data is included to see how the deliberation corresponds to activity in the brain. Out of a plethora of methods, electrophysiology[1] and

*Advised by: Julian Day-Cooney

[1]Extracellular Electrophysiology - A recording technique which involves the insertion of electrodes into the brain. It measures the change in the electrical activity in the neurons near the electrode and thus measures firing rate of neurons in spikes per second. It gives a great temporal resolution making it a very powerful technique. However, it is an invasive method and can only be used to measure the activity of a few neurons at a time.

functional magnetic resonance imaging (fMRI)[2] are specifically used to measure firing rates and to gauge where the decision-related brain activity is taking place, respectively.

Many real-life decisions require accumulation of evidence and information, either from the environment or from our memories, until it passes a threshold. This accumulation-to-threshold is explained by a group of models called sequential sampling models. Through extensive research, one of these models stands out as the most effective: The Drift Diffusion Model (DDM).

The DDM (Ratcliff,1978) [Rat78] postulates that decisions are made by the accumulation of noisy evidence over time which terminates once it reaches a threshold or a bound. The decision threshold is the amount of evidence needed to choose an alternative and make a decision. In the DDM, there is only one accumulation process whereas in other accumulator models the evidence for each response is accumulated independently. These models are like a race. The accumulation process that reaches the threshold first is what the subject decides. In the DDM, evidence accumulation is competitive. Figure 1 shows the drift diffusion process for a perceptual discrimination task as well as an ideal accumulator model (also called a race model).

The beauty of the DDM is that the brain makes decisions as it has been portrayed in Figure 1. The DDM is not just a model made to explain decision making but could be how the brain integrates evidence and makes decisions. The simple DDM is defined by four parameters, the starting point, the drift rate(μ) i.e. the rate of evidence accumulation, the value of the bounds (a and b), and the non-decision time, also called latency time which is the sum of the time before initiation of the accumulation and the time taken for action once a decision is reached. (Ratcliff,1978) [Rat78]

This review will begin with an in-depth description of the DDM, its advantages and the research on the neuronal populations representing the subparts of the DDM. Following this, I will discuss the applications of the DDM to different domains, with the aim of looking at its performance in more cognitive tasks. The review will conclude with possible future avenues.

1.1 Why DDM?

Certain aspects of the Drift Diffusion Model have made it very successful in all its applications so far. This section looks at the advantages of using the DDM as an analytic tool. The DDM explains response times (RTs) very well for both correct and error responses. It helps us visualize the effects of attributes like task difficulty and time pressure on the RTs, which serve as an important tool in analyzing behaviour. This explanatory power helps to differentiate the

[2]Functional Magnetic Resonance Imaging (fMRI) - An imaging technique that measures brain activity by detecting changes in blood flow. It is used to measure BOLD signals (blood-oxygen-level-dependent signals). It gives us great spatial resolution and can provide a clear image of how brain activity is localised. Another advantage is that it is non-invasive and is a relatively safe technique. However, it gives poor temporal resolution and does not show us moment to moment changes in activity.

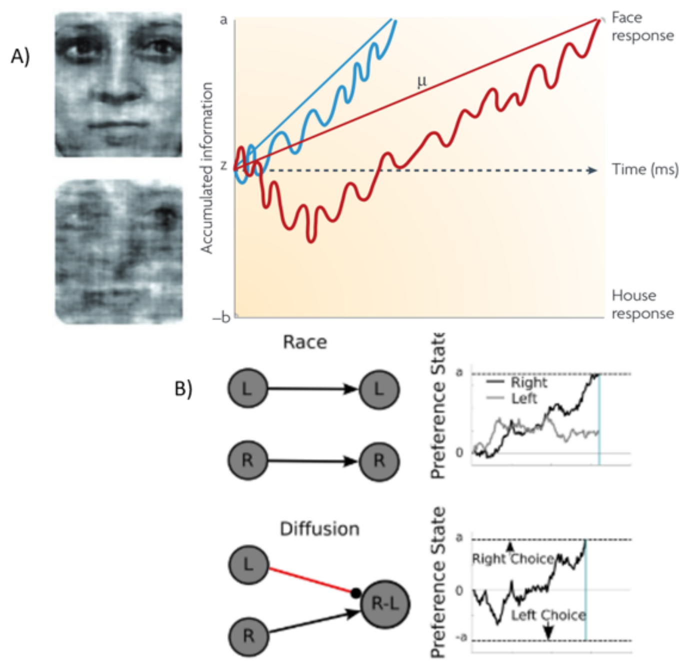

Figure 1: **A) z is the starting point for the process, a and -b are the thresholds. In the figure, there are two wobbly lines, which represent the decision process. The decision is encoded in a decision variable and this variable 'drifts' towards a threshold. A decision variable is a quantity which defines the possibility of one alternative over another and is driven by the integration of evidence over time. It can be thought of as a link between sensory evidence and the final choice. The straight lines connecting the starting point z and the thresholds are the drift rates. The image at the top is represented by the blue lines while the image at the bottom is represented by the red lines. In the figure, the blue lines have a much larger drift rate and as a result a short response time. Although intuitively we know that easier decisions will be made faster, this figure and the DDM, in general, gives us computational proof on the relation between decision difficulty and response time. Figure adapted from Heekeren et al., (2008) [HMU08]B) The graph adjacent to the race model shows two accumulation processes, one for each response for the race models, whereas, in the DDM, there is only one accumulation process which is a competition between the two alternatives. Figure adapted from Summerfield and Koechlin,(2008) [SK08]**

DDM from other sampling models like random walk models and accumulators which cannot model the RTs as accurately as the DDM (Ratcliff, 2004 [Rat04]). Although these conclusions are intuitive (harder tasks will have longer RTs), the DDM provides a computational framework for these conclusions. The response times are captured by RT distributions. These distributions can be represented by curves above their respective thresholds. The distribution encloses all the possible RT values for a particular experiment and its shape shows the variability in the response times and thus the variability in drift rates. The curves are shifted or skewed when task difficulty is changed or a time pressure is applied. Thus, the distributions help in giving an insight into the change in performances when task attributes are altered. Another advantage the DDM provides is that it explains the speed-accuracy trade-off well, (for an in-depth review see Bogacz et al., 2010 [BWFN10]). Higher decision thresholds will lead to more accurate answers since they require more evidence but will also lead to greater response times. On the other hand, lower thresholds will lead to fast responses, however, will result in a greater error rate. (see Figure 2A). Consider an investigation - more solid evidence will lead to catching the correct perpetrator but will take more time. However, quick justice could result in catching the wrong person. The DDM also provides a better understanding of choice biases. Biases can occur in two ways. First, there are starting point biases i.e. the starting point is closing to one bound, thus the decision-maker is inherently biased towards one decision. Second is the drift rate bias, in which the drift rate is higher for one response, biasing the decision to that alternative. Figures 2B and 2C show the effect of both these biases on the diffusion process.

2 Neural Correlates

While understanding the working of the brain is important, a major goal of neuroscience has been to map these processes to underlying circuits in particular regions of the brain. These regions are called neural correlates. This section will look at various studies pinpointing the correlates for the different sub-process of the diffusion process: evidence accumulation (drift rate), decision threshold, starting point bias, and comparison of alternatives. To find neural correlates in humans, the tool of choice would be fMRI, since it is non-invasive. However, causal studies[3] are also performed, giving more definitive proof that an area is responsible for a sub-process. Using fMRI, Rolls et al., (2010) [RGD10] found signatures in the dorsolateral prefrontal cortex (DLPFC) which could represent evidence accumulation. Philiastides et al., (2011) [PAHB11] showed the causal role of the DLPFC using trans-cranial magnetic stimulation.[4] They found that

[3] Causal Studies - Causal techniques are used to find direct causal relationships between brain regions and a specific function. Causal methods include inhibition of a particular area by stimulation, pharmacological inactivation and lesion studies. They have great explanatory power in finding neural correlates.

[4] Transcranial Magnetic Stimulation (TMS) - It is a non-invasive procedure in which neurons in the brain are stimulated by a magnetic field which induces electrical activity in those neurons.

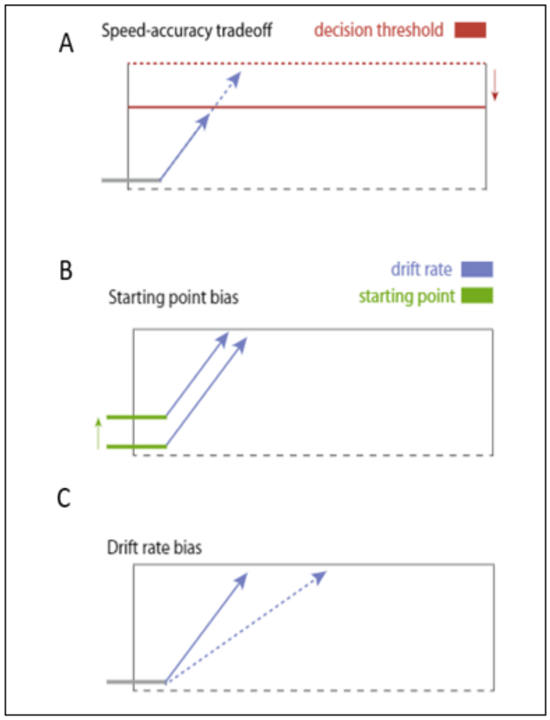

Figure 2: **a)The relationship between the speed-accuracy trade-off and the decision threshold. Lower thresholds can result in less accurate decisions. b) This figure shows the effects of the starting choice bias on the diffusion process. c)The effect of the drift rate bias on the diffusion process (Figures adapted from Mulder et al., 2014** [MvMF14]**)**

the drift rate was significantly reduced under the influence of the TMS while the non-decision time was almost unaffected, thus, showing the role of the DLPFC in evidence accumulation. Other studies have reported that areas like the frontal eye field (FEF) and intraparietal sulcus (IPS) could be responsible for evidence accumulation. (Basten et al., 2010 [BBHF10]; Ho et al., 2009 [HBS09]; Liu and Pleskac, 2011 [LP11])

The lateral intraparietal area (LIP), a subdivision of the IPS, has been also shown to represent sensory integration (Roitman and Shadlen, 2002 [RS02]). At this time, research points to a frontoparietal network (a network of areas in the frontal and parietal lobes of the brain) that is responsible for evidence accumulation. Studies regarding the decision threshold have pointed to a frontostriatal network which would include the anterior cingulate cortex (ACC), striatum, and the pre-supplementary motor area (pre-SMA) as candidate areas (Forstmann et al., 2008 [FDB+08]; Ivanoff et al., 2008 [IBM08]; Van Veen et al., 2008 [VVKC08]; Winkel et al., 2012 [WvMR+12]).

Kiani et al., (2014) [KCRN14] showed the response of neurons in the prearcuate gyrus during Changes of Mind. A change of mind would be a sudden change in the direction of the evidence accumulation. If the decision variable is drifting towards the upper threshold, a change of mind can be seen in a sudden reversal of direction towards the lower threshold. Mathematically, the sign of direction changes. The firing rates of these neurons peaked just before the saccade which could indicate the encoding of the decision threshold.

Various studies looking at value-based decision making have found encoding of subjective value and choice bias in the orbitofrontal cortex (OFC) (Forstmann et al., 2010 [FBD$^+$10]; Padoa-Schioppa and Assad,2006 [PSA06]; Summerfield and Koechlin, 2008 [SK08]). Other frontal areas such as the ACC, ventromedial prefrontal cortex (VMPFC), and DLPFC have been shown to encode starting point bias (Mulder et al., 2012 [MWR$^+$12]). The above-mentioned areas have also been shown to be responsible for the comparison of alternatives in choice tasks (Hare et al., 2011 [HSC$^+$11]; Hunt et al., 2012 [HKS$^+$12]).

Figure 3 summarizes the current research in finding neural correlates. Each dot in the figure represents a group of studies. The size of the dot represents the number of studies conducted. Thus, the figure shows every study conducted for the different parameters and sub-processes. The location of the dot shows the areas that are responsible for a sub-process. The colour of the dot shows the region of the brain the specific correlate is situated in and each region is represented by a unique colour as shown in the legend.

3 DDM In Perceptual and Lexical Tasks

This section will delve into the two most successful domains of applications of the DDM to behavioural data: perceptual tasks and lexical tasks. In the realm of perceptual decisions, researchers have applied the DDM to a simple dot motion-discrimination task (also called Newsome Dots or the Random Dot Kinematogram) and a categorization task. In the RDK task, the subject is shown a group of moving dots and needs to choose the average direction of the dots by a saccade. This task gives great control over task difficulty. It introduces motion perception and also requires the subject to compute the average motion of the dots which requires a large amount of integration. Figure 4 shows an RDK task.

Figure 1 shows an example categorization task, where the subject was required to place the given image in the house or the face category. This task also gives great control over changing the difficulty of the task but also requires evidence accumulation over time, and on the difficult trials, requires the comparison of the alternatives shown to the ideal response. For example, if the image has a low contrast the subject would need to compare the image to an ideal image of a face and that of a house and try to give the correct response. Both these tasks have been successfully modelled by the DDM. (see Gold and Shadlen, 2007 [GS07]; Heekeren et al., 2008 [HMU08] for an extensive review)

Ratcliff et al., (2004) [RMG04], showed the DDM applied to a lexical deci-

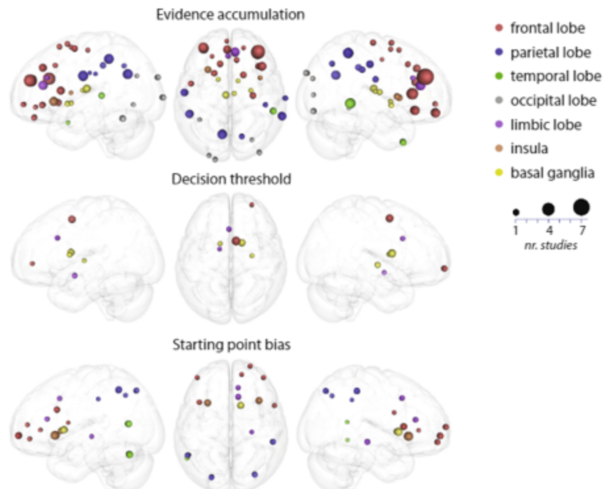

Figure 3: **Each dot is a separate group of studies. The size of the dot gives the number of studies conducted in a particular region. Regions have been highlighted as given in the legend. studies have been conducted to find correlates for evidence accumulation. This figure shows a frontoparietal network for the accumulation and a frontostriatal network for decision threshold. The starting point bias is almost only encoded by frontal networks. However, areas can have multiple functions and the distinctions are not always concrete. (Figure adapted from Mulder et al., 2014) [MvMF14]**

sion task. In this task, the human subject has to categorize the given stimuli into words and non-words. The stimuli were variable and were taken from a set of high-frequency words, low-frequency words, very low-frequency words, pseudowords, and non-words. The model explained the RTs for, both correct and error responses, and the probability of getting the decision correct very well for all types of stimuli. (see Table 3 in Ratcliff et al., 2004 [Rat04])

Recently, these applications have been extended to look at aging and IQ from a unique perspective. Studies have shown that older adults are slower than young adults due to longer non-decision times and a higher boundary, although age does not play a role in drift rate. This makes intuitive sense since older adults are usually more cautious and the decision thresholds prove this (Theisen et al., 2020 [TLvKV20]). Studies can also show how differences in IQ can affect decisions. Subjects with a higher IQ have higher drift rates but have almost equal non-decision times and boundary separations, as normal subjects. (Ratcliff et al., 2010 [RTM10]; Ratcliff and Mckoon, 2011 [RM11]).

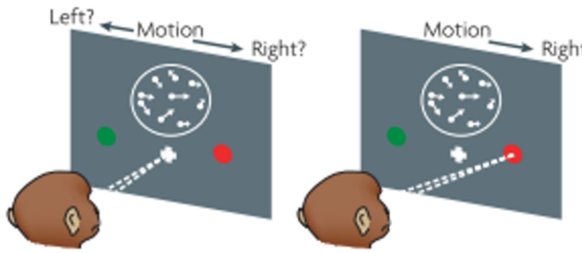

Figure 4: **An RDK task for a macaque monkey. The coherence of motion can be changed trial to trial. The monkey has to gauge the average motion and indicate its response by a saccade to one of the two targets on the screen. (Figure adapted from Heekeren et al., 2008) [HMU08]**

Studies looking at sleep deprivation, clinical populations, alcohol consumption, and reduced blood sugar have had success using a diffusion model analysis, thus proving that the DDM can be clinically useful. (See Forstmann et al., 2016 [FRW16] for an excellent review).

4 Extending the DDM to Economic Choices

This section will look at the modifications of the DDM for it to be applied to subjective tasks. All the tasks mentioned in the paper so far have had a defined correct response. This section will be an introduction into the domain of value-based choice. As I stated before, the model looked at so far is the simple DDM (sDDM) with 4 parameters. To extend the DDM to economic and subjective choices its computational framework behind the model needs to be modified. Milosavljevic et al., (2010) [MMH+10], compared the sDDM with 3 of its variants – the simple collapsing barrier DDM (scbDDM) [5], the full DDM (fDDM) [6], and the full collapsing barrier DDM (fcbDDM) [7], using the Bayes Information Criterion [8]. They found that the fDDM provided the best

[5] Simple Collapsing Barrier DDM (scbDDM) - It is a modification of the simple DDM in which the bound values (a and b) decrease as time progresses thus reducing the amount of evidence needed in the later stages of the trial. Just like sDDM it is defined by 4 parameters.

[6] Full DDM (fDDM) - Along with the 4 parameters of the sDDM it has an additional 4 parameters : a standard deviation parameter characterizing the noise in the accumulation process, a starting point bias parameter(zm), a range of latency times giving the distribution from which the latency time (non-decision time) is sampled every trial and a range of bias giving the distribution of the bias parameter.

[7] Full Collapsing Barrier DDM (fcbDDM) - It is defined by the same 8 parameters as the fDDM but now has collapsing barriers like the scbDDM.

[8] Bayes Information Criterion - BIC is a criterion used to compare different models for a given set of data samples. It strikes a balance between model complexity and model fit. Lower

quantitative description of their data and could be the model used by the brain. They conducted this analysis on a subjective value task in which the subjects had to choose between two food options.

However, the fDDM, as described in the paper, was suitable only for binary choice and did not take attention or fixations into account. When we choose between many alternatives, we often foveate (position our fovea centralis, the part of the eye with the sharpest vision) on the preferred option and this can bias our choice. This was not incorporated in the fDDM. To change this, Krajbich and Rangel, (2011) [KR11], proposed a novel drift-diffusion process for subjective multi-alternative decisions, the Attentional Drift Diffusion Model (aDDM). Their model included a fixation bias and explained the data for both binary and trinary value-based choices. The aDDM can also be extended to simple purchasing tasks, in which subjects need to decide whether or not to buy a product for the given price (Krajbich et al., 2012 [KLCR12]). The model explained RTs for different sets of stimuli but also showed the adverse impact of visual fixations. When subjects looked at the product more, they were more likely to buy it. On the other hand, if they looked at the price for a longer period, they were more likely to reject it. The effects of visual fixations and affection thus seem to bridge the perceptual and economic domains together. Thus, we can see that the DDM has been successfully modified to purchasing decisions and value-based choices, both for binary and trinary choices. Attempts have been made to extend the aDDM to quaternary choice (von Boguslawski and Mildén, 2015 [vBM15]), with mixed results. Although they have modeled choice well, the sample size may be small and the results may not be very significant. Still, it is another step towards modeling more complex tasks.

5 Cognitive Tasks

This section will look at novel applications of the DDM to more cognitive tasks. The purchasing and value-based experiments talked about in previous experiments were simplified accounts of real-life decisions. This section will address experiments looking at more complex behaviors and decisions.

5.1 Self-Control

Berkman et al. (2017) [BHL+17], put forth an alternative model for self-control. Rather than a competition between the impulsive and deliberative processes, they defined self-control as a value-based choice between two alternatives. Rather than modeling self-control with dual-process models, they used the drift diffusion model. By using the example of a dieter choosing between a salad and a burger, they looked as self-control as a comparison between the subjective values of two alternatives, thereby eliminating the need for a 'control' system. The decision would be governed by the values of the decision thresholds

the Δ BIC score, better the model. It penalizes the model for having too many observations and parameters and rewards the model for fitting the data well.

and the two alternatives. Their model captures internal events like effort expenditure by incorporating it into the value-integration process. Effort can be an opportunity cost that is compared with the benefits that an alternative pose., thus the task that needs more effort can be avoided, favoring the impulsive option over the deliberative option. This view of self-control could lead to understanding why damage to prefrontal cortices, areas thought to participate in evidence accumulation and comparison of alternatives, results in more impulsive decisions. It can also lead to further research into the realms of goal-attainment and motivation.

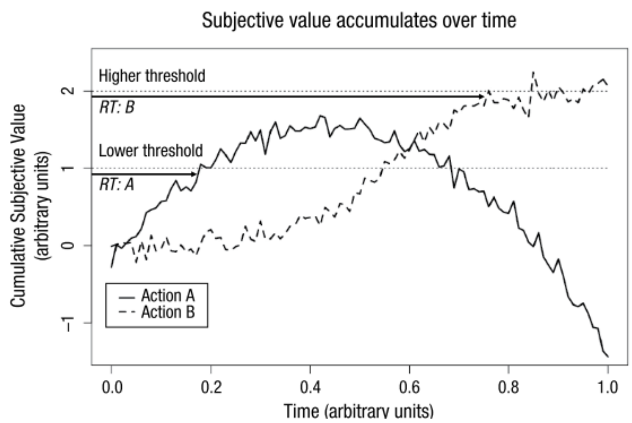

Figure 5: **Subjective value accumulates over time just as sensory information does. The value of Action A accumulates rapidly but falls over after some time whereas the value for Action B rises slowly but ultimately reaches a higher point. A person with a lower decision threshold would pick Action A and could have poor self-control. Under time pressure, A would be the action chosen. However, for a person with a higher decision threshold or with no time pressure, B would be chosen. This explanation can be heightened by taking the example of Action A being eating pizza and Action B as eating salad. (Figure adapted from Berkman et al., 2017 [BHL+17])**

5.2 Loss Aversion

Loss Aversion is one of the central tenets of Prospect Theory (Kahneman and Tversky, 1979 [KT79]), which proposes that when faced with risk or uncertainty, decision-makers are loss averse i.e. they place a greater weight on losses than they do on gains. An experiment that highlights this is when decision-makers are offered a gamble with a 50% probability to gain 11$ and a 50 % probability to lose 10$, they often reject the gamble. Despite the gamble having a positive expected value, it seems unattractive and is rejected. Zhao et al.,

(2020) [ZWB20], applied the DDM to this psychological phenomenon. They modified the full DDM to incorporate the unequal weightage of losses against gains but also incorporated a pre-valuation bias. This bias behaved similarly to the starting point bias, and represented a predisposition towards rejection, by being closer to the rejection bound (Figure 6). Since the starting point for the diffusion process is closer to the rejection bound, the decision-maker is biased towards rejecting the gamble. It takes less evidence for the decision variable to cross the bound, thus the RTs for rejection will be shorter and the probability of rejection will be greater. This bias corresponds to prior experience and introduces the concept of learning into the experiment. As they show in their paper, during trial blocks with higher payoffs –i.e. trials in which the possible gains were much greater than the possible losses, this pre-valuation bias was closer to the rejection bound, meaning it took a larger gain to loss ratio to convince the subject to accept the gamble, in this case, 1.83, whereas in trial blocks with lower payoffs i.e. trials in which the possible gains were almost equal to the possible losses, the pre-valuation bias was farther from the rejection bound, meaning that it took a smaller gain to loss ratio to convince the subject to accept the gamble ,in this case, a 1.25 gain-to-loss ratio. Thus, they show the influence of prior gambles and the prior rewards on the current gamble. Using the Deviance Information Criterion, a model criterion similar to BIC which penalizes the model with greater variance, and thus uncertainty in the data, they showed how the DDM outperforms older models explaining loss aversion. Through the incorporation of a starting point bias, in the form of the pre-valuation bias, their model captures the choice probabilities and the RTs for both rejected and accepted gambles. Thus, the DDM has been successfully modelled for another task, more cognitive than those of the economic and perceptual domains.

5.3 Driving Tasks

Recently, the DDM has been modified for different types of driving tasks. Cooper and Strayer, (2008) [CS08] conducted an experiment to determine the effects of cell-phone usage on driving. They used a 3D driving simulation during which subjects were engaging in a conversation they found interesting using a hands-free phone. Ratcliff and Strayer, (2014) [RS14] conducted an analysis on this study using a one-boundary drift diffusion model (Figure 7A) and found that distracted drivers have longer non-decision times and lower drift rates resulting in longer response times and slower uptake of information. Thus, this study provided computational proof as to why distracted drivers have higher chances of being in a car crash.

Building on this, Daneshi et al., (2020) [DAT20] used a one-boundary DDM to model time-to-collision to an obstacle. In this task, the subjects had to stay on their trajectory for as long as possible but prevent collision with the lead vehicle. (Figure 7B). They conducted this task with and without time pressure and found that both the drift rate and the decision threshold were higher for the trials with time pressure. This could mean that under time pressure drivers have

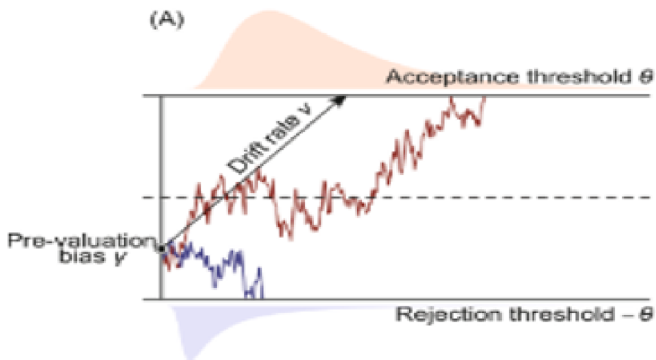

Figure 6: **A drift diffusion process for loss aversion. The pre-valuation bias γ appears as a starting point bias towards the rejection threshold. Since the distance from the thresholds is now unequal, rejection is more likely and will have a shorter response time since it takes less evidence to reach the threshold. (Figure adapted from Zhao et al., 2020 [ZWB20])**

greater evidence accumulation but also can be uncertain about their decisions thus increasing their decision thresholds and their margins of safety.

Both the previous studies have looked at simple braking and driving around tasks. The DDM has also been applied to more complex tasks such as accepting or rejecting a turn at an intersection. Zgonnikov and Abbink,(2020) [ZA20] used a modified full collapsing barrier DDM (fcbDDM) with variable drift rates to model a driving task which had subjects accept or reject a left turn with an oncoming car which could block them (Figure 7C). Evidence accumulation involved gauging the distance from the oncoming car and its speed and using this information to compute a time-to-arrival. Greater the time-to-arrival, greater the probability to turn. They found a positive relationship between response time and time-to-arrival (Figure 7D). Their model also accurately predicted their data well. This modification of a variable drift rate could be of huge importance to similar experiments that look at dynamic real-life scenarios. More research into driving behaviours could have applications in computer-driven cars and in making traffic interactions safer.

5.4 Racial Bias

In light of the increasing police brutality, researchers have tried to study racial biases in police officers and trainees and regular people. A first-person shooter task (FPST) has been developed for these studies. In the task, the participants are instructed to shoot armed targets and to not shoot unarmed targets. The

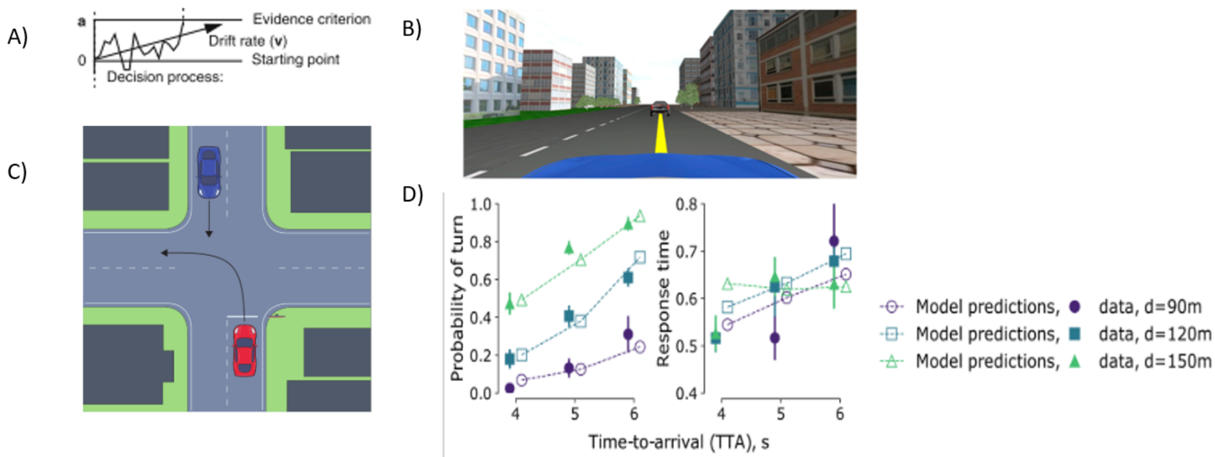

Figure 7: **A) A one-boundary diffusion model for driving tasks. The parameters remain the same as the sDDM. Figure adapted from Ratcliff and Strayer, (2014) [RS14] B) Participants have to stick to the yellow line for as long as possible and need to drive around once the obstacle gets too close. Figure adapted from Daneshi et al., (2020) [DAT20] C) The participants are in the red car. The speed of the blue oncoming car is variable. Participants have to decide whether or not to turn left. Figure adapted from Zgonnikov and Abbink, (2020) [ZA20] D) The model predictions are given by the dotted lines. They fit the data well. The probability to turn increases with an increase in time to arrival and an increase in distance from the oncoming car. The reaction time also shows a positive relationship with time-to-arrival. Figure adapted from Zgonnikov and Abbink, (2020) [ZA20]**

targets can be either Black or White men.

Using a hierarchical DDM [9], Johnson et al., (2017) [JHCP17] analysed an FPST in which participants were rewarded for correct shooting decisions. They found that participants had a starting point bias towards the to-shoot decision, however, this was independent of race and can be explained by the rewarding outcome for to-shoot decisions. Evidence accumulation was stronger to shoot armed Black men than to shoot armed White targets, thus participants have a higher drift rate when it comes to shooting armed Black men and this results in shorter response time and a greater likelihood to shoot armed Black men. Following this study, Johnson et al., (2020) [JSCF20] looked at the effects of sleep deprivation and caffeine on racial biases. They found that subjects were more likely to shoot unarmed Black men than unarmed White men and this bias was not affected by either sleep deprivation or caffeine. Caffeine did not mitigate the errors caused by sleep deprivation. It only reduced response times. They also found that subjects set a wider threshold for White men than for Black men, showing that they needed lesser evidence when it came to making a decision when they were shown a Black man as the target. Surprisingly, they found that overall, participants who were given a placebo had a higher starting point to shoot White targets.

This study reaffirmed a conclusion that Johnson et al., (2017) [JHCP17] had come to, proving that for unarmed targets, subjects had a lower drift rate for Black men than for White men and for armed targets, had a higher drift rate for Black men than for White men.

5.5 Reinforcement Learning

The DDMs looked at in the review so far have not incorporated an element of learning into the process, however, the drift diffusion process modeled for loss aversion hinted at the influence of past outcomes. Recently, the DDM has been applied to learning tasks as well. These groups of models are called reinforcement learning drift diffusion models (RLDDM). They unify the DDM and the theory of reinforcement learning (See Seo and Lee, (2012) for an excellent review [SL12]). These models have been applied to probabilistic selection tasks (PST), which present the participants with two options with the goal being to pick the option with a greater probability to be rewarded. However, the participants need to learn the probabilities and rewards of these options as the trials go on. The simplest RLDDM has 4 parameters: a learning rate [10], threshold

[9]Hierarchical DDM – The HDDM analyzes the data at the population level rather than at an individual level. This means that fewer trials can be conducted per participant, but the parameters can easily be recovered and can still capture all the aspects of the data as the simple DDM does.

[10]Learning Rate - The learning rate determines how sensitive the decision maker is to previous outcomes. A learning rate that is too low is not optimal since the learning will be very slow, however a learning rate that is too high will induce forgetting the outcomes that happened a few trials back.

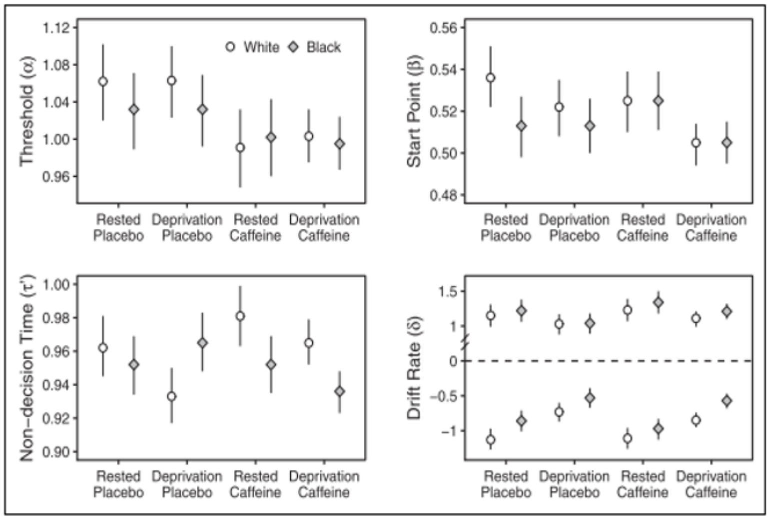

Figure 8: **The x-axis shows the 4 groups of patients: Patients with a whole night's sleep on a placebo, patients with a whole night's sleep on caffeine, patients who had not slept for 24 hours on placebo, and patients who had not slept for 24 hours on caffeine. The top left panel shows that subjects had wider thresholds for white men than for black men. The top right panel shows that subjects, surprisingly, had a starting point bias to shoot white men. The bottom left panel shows the effect of sleep deprivation and caffeine. The bottom right panel shows the drift rates for the different trials. All the negative drift rates i.e. below the dashed lines are for unarmed targets while those above the dashed lines are for armed targets. Thus, Black armed men produced a higher drift rate in the participants while subjects had a lower drift rate for unarmed black men than those for white men. Figure adapted from Johnson et al., (2020) [JSCF20]**

values, a scaling parameter vmod [11], which ensures that the difference in values and probabilities for the two choices are transformed into an appropriate scale in the DDM framework, and the non-decision time.

Fontanesi et al., (2019) [FGSR19] showed that the RLDDM can explain both reaction times and choice probabilities very well. However, it also shows the learning throughout the task and thus successfully combined the RL models and the DDM (Figure 9).

The accuracy of the responses increased steadily and the RTs decreased throughout the task, showing that the participants learnt the probability of

[11]Scaling Parameter - It is analogous to the drift rate in the DDM. It helps to convert the effect of previous outcomes into an appropriate scale that can be incorporated into the DDM framework.

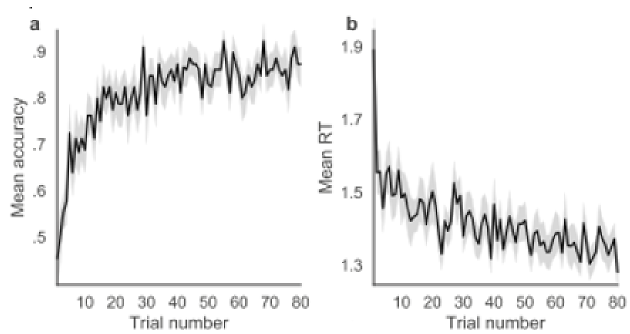

Figure 9: **The top panels show the power of the RL models whereas the bottom two represent the contribution of the DDM to the RLDDM. Thus, two prominent theories can be unified to give a better account of decision making, a) The accuracy increases as the trial number increases thus showing the effect of learning on the task. b) The RT decreases, once again showing the effect of learning. Figure adapted from Fontanesi et al., (2019) [FGSR19]**

the rewards and improved their performance. They also conducted an analysis to find out which RLDDM explained the data the best. The RLDDM can be modified by having different learning rates for negative and positive outcomes. The threshold can either be fixed or variable, and the scaling parameter can either be linear or sigmoid. Thus, there can be 8 types of the RLDDM. Using the Watanabe-Akaike Information Criterion [12], they found that the full RLDDM i.e. with dual learning rates, one for positive and one for negative outcomes, with sigmoid scaling parameters and with variable bounds explain the date the best (Figure 10).

Pedersen et al., (2017) [PFB17] used the RLDDM to gain a different perspective on ADHD patients and the effects of medication. They found that medication increased boundary separation, lowered learning rates, increased non-decision time, and increased the drift rate scaling, showing the shift towards focusing on accuracy rather than speed. Thus, the RLDDM has the potential to be used in many clinical experiments.

6 Discussion

Recently, researchers have tried applying the Drift Diffusion Model, a popular computational model for sensory decision-making, to more cognitive and complex tasks. These studies have shown that the DDM can explain a variety of psy-

[12]Watanabe Akaike Information Criterion - Another model criterion like BIC and DIC. It penalizes the model in a way, similar to that of DIC but it takes the summation of the variance of each posterior draw. It is more computationally taxing than both BIC and DIC but gives a better approximation of how good the model is.

Model	η	v	a	p_{WAIC}	$-$lppd	WAIC
RLDDM 1	One	Linear	Fixed	111	5129	10481
RLDDM 2	Two	Linear	Fixed	134	5051	10369
RLDDM 3	One	Linear	Modulated	145	4942	10174
RLDDM 4	Two	Linear	Modulated	159	4866	10048
RLDDM 5	One	Sigmoid	Fixed	137	4930	10135
RLDDM 6	Two	Sigmoid	Fixed	159	4861	10039
RLDDM 7	One	Sigmoid	Modulated	164	4672	9672
RLDDM 8	Two	Sigmoid	Modulated	190	4613	9607

Figure 10: **Lower the WAIC score, better the model. The dual learning rate means that positive and negative outcomes will have a different effect on the decision-maker. The threshold is variable, thus when the decision is the easiest the threshold is the lowest. The scaling parameter can be fixed or s-shaped.**

chological phenomena and real-life decisions. The model captures the response times and the behaviour of the participants in these tasks and helps in giving a greater insight into the underlying processes of deliberation and decision-making. Further research into the DDM could have powerful applications in consumer behaviour, traffic behaviour and laws, computer-driven vehicles and could have important clinical and social applications. Though the tasks looked at in this review are more complex than sensory and simple value-based choice, they are still a step away from explaining important and life-changing decisions. The model comparison criteria used in this review (BIC, DIC, WAIC) may represent a caveat in the literature. These criteria penalize the models in different manners and lack of uniformity in the literature could result in the selection of incorrect models (Churchland and Kiani, 2016 [CK16]) . Future research should aim to refine the DDM framework and attempt to resolve the debates about the dynamics of the drift-diffusion process. Future work should also aim to conduct more extensive research in finding definitive neural correlates and circuits for the parameters. New techniques such as calcium imaging and optogenetics, if adapted to work in primates, hold interesting possibilities. Another interesting innovation in the literature is the quantum drift-diffusion model (Rosendahl et al., 2020 [RBC20]), which looks at evidence as a quantum particle of information and the threshold as a square attractor. This may open new and fascinating avenues for improving computational models as a whole. In the last 10 years, the extensions of the Drift Diffusion Model have led to tremendous progress in understanding how cognitive decisions are made. The DDM has the potential to model more complex decisions and holds a lot of promise for the future.

References

[BBHF10] U. Basten, G. Biele, H.R. Heekeren, and C.J. Fiebach. How the brain integrates costs and benefits during decision making. *Proc.*

Natl. Acad. Sci. U. S. A., 2010.

[BHL+17] E.T. Berkman, C.A. Hutcherson, J.L. Livingston, L.E. Kahn, and M. Inzlicht. Self-Control as Value-Based Choice. *Curr. Dir. Psyhchol. Sci.*, 2017.

[BWFN10] R. Bogacz, E.J. Wagenmakers, B.U Forstmann, and S. Nieuwenhuis. The neural basis of the speed-accuracy tradeoff. *Trends Neurosci.*, 2010.

[CK16] A.K. Churcland and R. Kiani. Three challenges for connecting model to mechanism in decision-making. *Curr. Opin. Behav. Sci.*, 2016.

[CS08] J.M. Cooper and D.L. Strayer. Effects of simulator practice and real-world experience on cell-phone-related driver distraction. *Hum. Factors*, 2008.

[DAT20] A. Daneshi, H. Azarnoush, and F. Towhidkhah. A one-boundary drift-diffusion model for time to collision estimation in a simple driving task. *J. Cogn. Psychol.*, 2020.

[FBD+10] B.U. Forstmann, S. Brown, G. Dutilh, J. Neumann, and E.J. Wagenmakers. The neural substrate of prior information in perceptual decision making: A model-based analysis. *Front. Hum. Neurosci.*, 2010.

[FDB+08] B.U. Forstmann, G. Dutilh, S. Brown, J. Neumann, D.Y. von Cramon, K.R. Ridderinkhof, and E.J. Wagenmakers. Striatum and pre-sma facilitate decision-making under time pressure. *Proc. Natl. Acad. Sci. U. S. A.*, 2008.

[FGSR19] L. Fontanesi, S. Gluth, M.S Spektor, and J. Rieskamp. A reinforcement learning diffusion decision model for value-based decisions. *Psychon. Bull. Rev.*, 2019.

[FRW16] B.U. Forstmann, R. Ratcliff, and E.J. Wagenmakers. Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annu. Rev. Psychol.*, 2016.

[GS07] J.I. Gold and M.N. Shadlen. The Neural Basis of Decision Making. *Annu. Rev. Psychol.*, 2007.

[HBS09] T.C. Ho, S. Brown, and J.T. Serences. Domain general mechanisms of perceptual decision making in human cortex. *J. Neurosci.*, 2009.

[HKS+12] L.T. Hunt, N. Kolling, A. Soltani, M.W. Woolrich, M.F.S Rushworth, and T.E.J. Behrens. Mechanisms underlying cortical activity during value-guided choice. *Nat. Neurosci.*, 2012.

[HMU08] H.R. Heekeren, S Marrett, and L.G. Ungerleider. The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.*, 2008.

[HSC+11] T.A. Hare, W. Schultz, C.F. Camerer, J.P. O'Doherty, and A. Rangel. Transformation of stimulus value signals into motor commands during simple choice. *Proc. Natl. Acad. Sci. U. S. A.*, 2011.

[IBM08] J. Ivanoff, P. Branning, and R. Marios. fmri evidence for a dual process account of the speed-accuracy tradeoff in decision-making. *PLoS One 3*, 2008.

[JHCP17] D.J. Johnson, C.J. Hopwood, . Cesario, and T.J. Pleskac. Advancing research on cognitive processes in social and personality psychology: A hierarchical drift diffusion model prime. *Soc. Psychol. Personal. Sci*, 2017.

[JSCF20] D.J Johnson, M.E. Stepan, J. Cesario, and K.M. Fenn. Sleep deprivation and racial bias in the decision to shoot: A diffusion model analysis. *Soc. Psychol. Personal. Sci*, 2020.

[KCRN14] R. Kiani, C.J. Cueva, J.B. Reppas, and W.T. Newsome. Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. *Curr. Biol.*, 2014.

[KLCR12] I. Krajbich, D. Lu, C. Camerer, and A. Rangel. The attentional drift-diffusion model extends to simple purchasing decisions. *Front. Psychol.*, 2012.

[KR11] I. Krajbich and A. Rangel. Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl. Acad. Sci. U. S. A.*, 2011.

[KT79] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 1979.

[LP11] T. Liu and T.J. Pleskac. Neural correlates of evidence accumulation in a perceptual decision task. *J. Neurophysiol.*, 2011.

[MMH+10] M. Milosavljevic, J. Malmaud, A. Huth, C. Koch, and A. Rangel. The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgm. Decis. Mak*, 2010.

[MvMF14] Mulder.M.J, L. van Maanen, and B.U. Forstmann. Perceptual decision neurosciences - a model-based review. *Neuroscience*, 2014.

[MWR+12] M.J. Mulder, E.J. Wagenmakers, R. Ratcliff, W. Boekel, and B.U. Forstmann. Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *J. Neurosci.*, 2012.

[PAHB11] M.G. Philiastides, R. Auksztulewicz, H.R. Heekeren, and F. Blankenburg. Causal role of dorsolateral prefrontal cortex in human perceptual decision making. *Curr. Biol.*, 2011.

[PFB17] M.L. Pedersen, M.J. Frank, and G. Biele. The drift diffusion model as the choice rule in reinforcement learning. *Psychon. Bull. Rev.*, 2017.

[PSA06] C. Padoa-Schioppa and K=J.A. Assad. Neurons in the orbitofrontal cortex encode economic value. *Nature*, 2006.

[Rat78] R. Ratcliff. A Theory of Memory Retrieval. *Psychol. Rev.*, 1978.

[Rat04] R. Ratcliff. A Comparison of Sequential Sampling Models for Two-Choice Reaction Time. *Psychol. Rev.*, 2004.

[RBC20] M. Rosendahl, A. Bizyaeva, and J. Cohen. A Novel Quantum Approach to the Dynamics of Decision Making. *Cognitive Science Society*, 2020.

[RGD10] E.T. Rolls, F. Grabenhorst, and G. Deco. Decision-making, errors, and confidence in the brain. *J. Neurophysiol.*, 2010.

[RM11] R. Ratcliff and G. Mckoon. Effects of aging and iq on item and associative memory. *Pyschol. Rev.*, 2011.

[RMG04] R. Ratcliff, G. McKoon, and P. Gomez. A diffusion model account of the lexical decision task. *Psychol. Rev.*, 2004.

[RS02] J.D. Roitman and M.N. Shadlen. Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task. *J. Neurosci.*, 2002.

[RS14] R. Ratcliff and D. Strayer. Modeling simple driving tasks with a one-boundary diffusion model. *Psychon. Bull. Rev.w*, 2014.

[RTM10] R. Ratcliff, A. Thapar, and G. Mckoon. Individual differences , aging , and iq in two-choice tasks. *Cogn. Psychol.*, 2010.

[SK08] C. Summerfield and E. Koechlin. A neural representation of prior information during perceptual inference. *Neuron*, 2008.

[SL12] H. Seo and D. Lee. Neural basis of learning and preference during social decision-making. *Curr. Opin. Neurobiol,*, 2012.

[TLvKV20] M. Theisen, V. Lerche, M. von Krause, and A. Voss. Age differences in diffusion model parameters: a meta-analysis. *Pyschol. Rev.*, 2020.

[vBM15] M. von Boguslawski and P. Mildén. The attentional drift-diffusion model for simple choice in the quaternary case, measuring the effect of permutation of item location on choice. *Behaviour*, 2015.

[VVKC08] V. Van Veen, M.K. Krug, and C.S. Carter. The Neural and Computational Basis of Controlled Speed-Accuracy Tradeoff during Task Performance. *J. Cog. Neurosci.*, 2008.

[WvMR+12] K. Winkel, L. van Maanen, R. Ratcliff, M.E. Van der Schaaf, M.R. Van Schouwenburg, R. Cools, and B.U. Forstmann. Bromocriptine does not alter speed-accuracy tradeoff. *Front. Neurosci.*, 2012.

[ZA20] A. Zgonnikov and D. Abbink. Should i Stay or Should I go? Evidence Accumulation Drives Decision Making in Drivers. *PsyArXiv*, 2020.

[ZWB20] W.J. Zhao, L. Walasek, and S. Bhatia. Psychological mechanisms of loss aversion: A drift-diffusion decomposition. *Cogn. Psychol.*, 2020.