Adaptation to Misinformation: A Drift Diffusion Model Analysis of Prior Likelihood

Pien Carlier – j.w.carlier@students.uu.nl
Master Thesis Applied Cognitive Psychology
27.5 ECTS
June 2024

Supervisor: Dr. Martijn Mulder – m.j.mulder@uu.nl
Assessor: Dr. Sjoerd Stuit – s.m.stuit@uu.nl

Abstract

Perceptual decision-making involves rapidly classifying sensory information to select appropriate responses. These decisions can be influenced by choice biases, which are systematic preferences for certain options and may rise from prior expectations of the likelihood of the occurrence of an event. While these biases increase decision efficiency, they can lead to suboptimal outcomes if expectations misalign with reality. In this study, we first investigated whether expected likelihoods of events induce proportional choice biases. Then, we examined the extent to which discrepancies between expected and actual likelihoods influence the development of these choice biases. We were particularly interested in any potential differences between adaptation to prior information inaccurately predicting a high likelihood of an event versus inaccurately predicting a low likelihood of an event. Participants (N=61) performed a visual discrimination task where prior information predicted the likelihood of occurrence for two options. We manipulated both the strength of the probability and prediction accuracy of the prior information. We then fitted a drift diffusion model (DDM) to the behavioral data of each participant. The starting point parameters of the DDM indicate that likelihood expectations do not result in proportional biases. Furthermore, participants showed no significant adaptation in response to inaccurate prior information, suggesting that bias development is mainly driven by prior information, regardless of the accuracy of this information.

INTRODUCTION

Perceptual decision-making is the process of classifying sensory information from the environment, allowing for the selection of an appropriate response (Summerfield & Blangero, 2017). These types of decisions are made rapidly and with little deliberation (Dutilh & Rieskamp, 2015), for example, when assessing a facial expression or the safety of crossing a street.

When making a perceptual decision, a choice bias may influence the decision process. A choice bias refers to a systematic preference for one option over other options, leading individuals to choose the preferred option more frequently and more quickly (Tobena et al., 1999; Kelly et al., 2019; Dunovan et al., 2014; White & Poldrack, 2014; Mulder et al., 2012; Summerfield & Blangero, 2017). Although a bias might lead to false observations or deductions, there are significant advantages to having a bias (Tobena et al., 1999). For example, by reducing the need to consider all options equally, choice biases allow for more efficient information processing, saving cognitive resources and resulting in quicker responses.

Such choice biases can be caused by prior knowledge, the information that an individual possesses before encountering sensory information, such as past experiences or information about the situation. This prior knowledge might shape expectations about the likelihood of different events and cause a bias for the option that is expected to be more likely. For example, when deciding whether it is safe to cross the road, prior knowledge of frequent accidents on this crossing might induce a bias for the option "dangerous to cross".

However, expectations of the likelihood of an event can have a mismatch with the actual likelihood of an event. A bias caused by inaccurate expectations can result in suboptimal, potentially dangerous choices. For instance, when someone is not aware of the frequent accidents that happen on the crossing, they might underestimate the likelihood of an accident. A bias caused by this underestimation might cause them to be more inclined to cross the street, exposing them to more risk than they are aware of.

A mismatch between the expectations and the actual likelihood of an event becomes apparent to an individual when the outcome of the decision is evaluated (Rangel et al., 2008). When such a mismatch exists, and the outcome of the decision does not coincide with the expectations, a prediction error occurs. The prediction error indicates that the initial representation of the decision was incorrect and calls for an update of the representation that is proportional to the prediction error (Rangel et al., 2008). As a result, future, similar decisions may yield a more appropriate response to the situation and result in the expected outcome.

However, after an update, the expectations of the likelihood of an event may approach, but still not be an exact representation of the true likelihoods and in turn still result in an undesirable choice bias. The prediction error after a mismatch might be underestimated or overestimated, for instance, when the mismatch between expected and actual occurrences is too small and therefore hardly observed. Consequently, it might require multiple experiences and updates to reach an accurate representation that results in a more beneficial choice bias.

To investigate this process of adaptation, this study aims to explore the extent to which mismatches between the expectations of the likelihood of events and the actual likelihood of events result in prediction errors, and therefore changes in a choice bias. To examine whether individuals strongly adhere to their prior knowledge or adjust well to the actual situation, we first seek to determine the extent to which accurate prior information of the likelihood of an event induces choice biases and examine whether these likelihood priors induce proportional biases. Then, we aim to determine the choice biases when the likelihood expectations turn out to be inaccurate. Specifically, we are interested in any potential differences between the adjustment in response to prior information inaccurately predicting a low versus high likelihood of an event.

One method for understanding perceptual decision-making and the influence of biases involves the use of computational models. These models aim to disentangle the underlying cognitive processes of a decision by simulating how their interaction results in observable behavior. In these models, a decision is the result of a stochastic process of evidence accumulation. This decision process is explained in terms of several parameters. From these parameters, specific conclusions about the underlying cognitive processes that influence decision-making can be drawn.

A powerful accumulation model is the drift diffusion model (Ratcliff, 1978) (Figure 1). When facing a binary choice, the DDM assumes the decision process begins at a starting point (z). From this starting point, evidence accumulation begins until a decision threshold (a, 0) for either one of the options is reached. The drift rate (v) is the rate of evidence accumulation, the average amount of evidence accumulated per time unit. The noise (η) is the standard deviation from the drift rate and refers to random disturbances, either in the sensory information or in the brain, which can cause the same information to result in different decision times or even different choices. The time it takes to get from the starting point to the threshold and reach a decision is the decision time. The reaction time is the decision time plus the stimulus-independent non-decision time (Ter), which includes the time it takes to encode the stimuli and execute the decision (for reviews, see Ratcliff & McKoon, 2008; Voss et al., 2004; Ratcliff et al., 2016).

For instance, when deciding whether it is safe to cross the road, the decision process starts at the starting point. From that point evidence accumulation begins for both options "safe to cross" and "dangerous to cross", for example by assessing the distance to the other side of the street, the speed of the cars, and the time remaining on the traffic light. Depending on how these situations appear, you may choose to cross the street, which means you have reached the decision threshold. For noise, this could be unpredictable factors such as sudden noises or unexpected changes in traffic patterns.

A starting point in the middle of the bounds indicates that for both options the same amount of evidence is required to reach the decision thresholds. When a starting point shifts closer to one boundary, it creates an asymmetric distance from the starting points to the boundaries. Less evidence for that option is needed to reach the decision, thus creating a choice bias for that option. Previous studies of choice bias have shown that in the drift diffusion model, prior knowledge of the likelihood of the options generally is reflected in the position of the starting points (Leite & Ratcliff, 2011, Mulder et al., 2012, White and Poldrack, 2014). Therefore, we expect adaptation to an inaccurate representation of the likelihoods to be reflected in a shift in the starting points.

We hypothesize that prior knowledge that predicts a high likelihood of the occurrence of an event causes a larger choice bias than a low-likelihood prior because high-likelihood prior knowledge reduces the uncertainty of a choice to a greater extent compared to low-likelihood prior knowledge. We assume the choice bias caused by prior knowledge of the likelihood to be at its minimum when uncertainty is at its maximum, which corresponds to a 50% expected likelihood of the occurrence of an event. Conversely, we assume the choice bias to be at its maximum when a 100% likelihood is expected. Between these extremes, we anticipate that all choice biases caused by likelihood expectations of intermediate percentages follow a linear relationship, meaning each change in percentage points would result in an equivalent absolute change in starting points (Figure 2A).

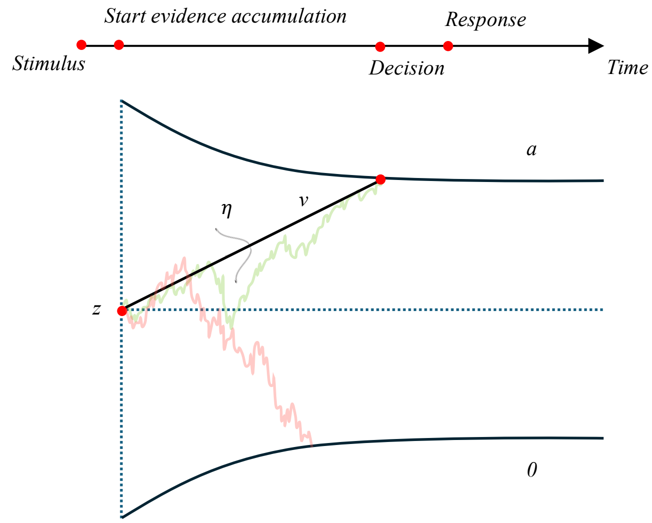

Figure 1. Schematic representation of the drift-diffusion model, featuring the starting point (z), drift rate (v), bounds (a, 0), and noise (η). The green and red path serve as an example of a decision where evidence is sometimes accumulated for the option corresponding to the lower bound (path moves towards the lower bound) and sometimes for the option corresponding to the upper bound (path moves towards the upper bound). The green pad results in a decision for the option at the upper bound and the red path in a decision at the lower bound. The decision time is the time it takes from the beginning of evidence accumulation until reaching the bound and making the decision.

However, we predict that if prior expectations of the likelihood of an event turn out to be inaccurate, mismatches in higher likelihood expectations result in more pronounced choice bias adaptations than mismatches in lower likelihood expectations (figure 2B). The reason for this is that expectations of a higher likelihood of the occurrence of an event restrict the area for uncertainty more than lower likelihood expectations, which could make the mismatch more noticeable. Consequently, the mismatch would sooner result in a prediction error and thus in changes in the choice bias, reflected by a shift in the starting point.

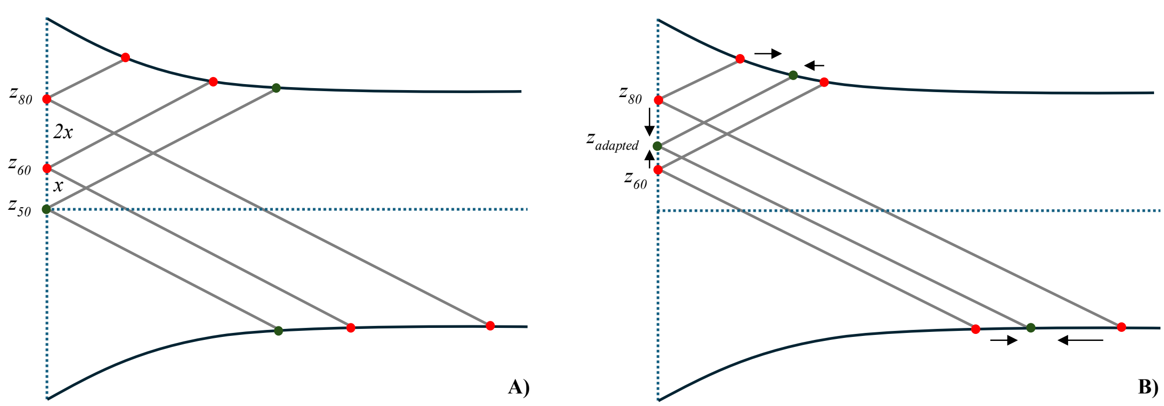

Figure 2. **A)** Expectations of starting points as a result of prior expectations of the likelihood of the occurrence of an event. We expect the starting point to be closest to the middle when the uncertainty is at its maximum, which is when a 50% likelihood of occurrence is expected (z_{50}). We expect that as the likelihood of an event increases, the starting points increases proportionately. Therefore, the difference between the starting points when a 50% (z_{50}) versus 60% (z_{60}) likelihood is expected to be half as small as the difference between the starting points at an expected likelihood of 60% (z_{60}) versus 80% (z_{80}). **B)** Expectations of changes in the starting points when the prior expectations of the likelihood of an event turn out to be inaccurate. We expect mismatches in both high and low likelihood expectations to result in shifts in the starting point in the direction of the actual perceived likelihood of the event ($z_{adapted}$). However, we expect this shift to be more pronounced when a higher likelihood of an event was expected. Therefore, we expect a mismatch when an 80% likelihood of an event is expected while the actual likelihood is 60% to result in bigger starting point shifts than when prior expectations of a 60% likelihood mismatches with an actual 80% likelihood.

METHOD

Participants. From 61 participants (49 female, 10 male, 2 other, mean age = 21.3, SD = 2.5), 51 participants reported to be right-handed, 10 to be left-handed. Participants were recruited via Utrecht University's Sona Systems (https://www.sona-systems.com/) (55) and rewarded 0.75 PPU or were recruited via online media (6). All participants provided informed consent before participating in the study. The study was built in Gorilla Experiment Builder (https://app.gorilla.sc, Anwyl-Irvine et al., 2019) and approved by the Faculty Ethics Review Committee.

Stimuli. All stimuli were generated in Python. Stimuli consist of a square with 150×150 pixels. The square consists entirely of 2500 black and white squares (Figure 3). Since our environment is usually ambiguous, there are two levels of difficulty, providing a balance between situations where a discrepancy between expectations and reality might be more noticeable, in the easy stimuli, and less noticeable, in the hard stimuli. In the easy level, the black-white ratio will be 53/47 (1325 squares of one color and 1175 squares of the other color). In the hard level, the color ratio will be 52/48 (1300 squares of one color and 1200 squares of the other color).

To determine these difficulties, we conducted a pilot study without any prior information. Too difficult stimuli might cause participants to entirely base their decisions on the prior information and fail to notice inaccurate prior information they will encounter. Conversely, when the difficulty is too low, participants might disregard the prior information, since it does not help in reaching a decision. The chosen ratios lead to an accuracy of 85% for easy stimuli and 75% for hard stimuli, which should provide a balance wherein participants use, but do not entirely depend on their prior knowledge.

Experimental design. The experiment consists of five blocks, one for each condition. For each participant, the order of the conditions is randomized. Each block consists of 100 trials, all with a break after 50 trials. Within each block, half of the stimuli is of level easy and half of level hard, randomly distributed. The task is to indicate whether there is more black or more white present in a stimulus by pressing the p-key or the q-key (for a schematic overview of a block, see Figure 4).

Figure 3. Stimuli of level easy with more black squares (left) and stimuli of level hard with more white squares (right).

For each participant, the keys remained consistent throughout the experiment. The key-color correspondence is counterbalanced to account for the effect of handedness.

At the beginning of each block, prior information indicates the dominance of a color in the upcoming trials, and it is stressed that this information will be helpful in their judgment. The five conditions are as follows (for an overview, see Table 1): In the **50-Match** condition, participants are accurately informed that in the upcoming trials, both black and white colors will be dominant in 50% of the stimuli. In the **60-Match** and **80-Match** conditions, participants are accurately informed about the respective 60% and 80% dominance of a color in the upcoming trials. In the **60-Mismatch** condition, a 60% color dominance is predicted, while the actual dominance of that color is 80%. In the **80-Mismatch** condition, an 80% dominance of a color is predicted, while the actual dominance of that color is only 60%.

In the 60-Match and 80-Mismatch conditions, the same color is predicted to be more dominant, while in the 80-Match and 60-Mismatch conditions, the other color is predicted to be more dominant. The assignment of which color is dominant in which pair of conditions is counterbalanced between all participants.

	50% predicted dominance	**60% predicted dominance**	**80% predicted dominance**
50% actual dominance	50-Match	-	-
60% actual dominance	-	60-Match	80-Mismatch
80% actual dominance	-	60-Mismatch	80-Match

Table 1. Five conditions formed by different combinations of actual and predicted dominance of a color in the upcoming trials. Within each condition, half of the stimuli is of level easy and half is of level hard.

Procedure. At the beginning of the experiment, participants were asked for consent and instructed to minimize distractions, maximize screen brightness, and set their screens to fullscreen mode. Subsequently, participants were given the task instructions, directing them to indicate whether they perceived more black or white in an image by pressing the p- and q-keys. Additionally, participants were informed about the duration of the experiment and were briefed that before each block, they would be given information about the probability of a color being the right answer. Participants underwent 15 practice trials to familiarize themselves with the task's design and pace. No prior information was given during these practice trials. After each trial, they received feedback on whether they chose the correct answer.

Before each block, participants were presented with the prior information about the dominance of black and white colors in the upcoming trials, both with text and with a visualization (figure 4). This prior information was repeated during the break between trials. To keep a focus on this prior information, a simplified graph is shown during the trials, showing the ratio of both colors as indicated by the prior information. In each trial, a fixation cross was presented for a duration determined by a random number drawn from a normal distribution between 600 ms and 1200 ms. Responses faster than 250 ms were categorized as premature, and a notification 'too fast' was displayed for 1500 ms. Failure to respond within 2000 ms was considered a 'too late' response (see Figure 4**)**.

After completion of the experiment, participants were debriefed. To indicate whether participants noticed the inaccuracies in the prior information, they were asked whether they suspected the information provided before each trial might not always have been accurate. If they answered yes, they were asked to indicate how many times they thought the prior information was not accurate. Hereafter, everyone was informed that in 2 out of 5 times, the prior information was not accurate.

Behavioral analysis. During the experiment, response times and choices were measured. Descriptive results were obtained using Jasp (JASP Team (2024) JASP (Version 0.18.3)). Response times lower than 250 ms are considered too fast and removed. Response times higher than three times the standard deviation from the average response times are considered outliers and also removed (1.90%). To indicate the effect of accurate prior information on bias development, we first compared the accuracy within the biased trials between conditions 50-Match, 60-Match, and 80-Match using a 3×2 (Prior $\times$ Difficulty) repeated measures ANOVA. Then, we compared the reaction time for biased choices with the reaction times for unbiased choices, also by performing a 3×2 (Prior $\times$ Difficulty) repeated measures ANOVA. To indicate the effect of inaccurate prior information, we compared the proportion of choices in favor of the option predicted to be more dominant by the prior information between the different conditions using a $2 \times 2 \times 2 \times 2$ (Prior $\times$ Difficulty $\times$ Match $\times$ BiasedTrial) repeated measures ANOVA.

DDM analysis. To investigate the underlying processes behind the observable behavior, we used the drift diffusion model. We designed four different models using the PyDDM Python package (Shinn et al., 2020). Each model reflects a different hypothesis on bias development in response to (inaccurate) prior information. We compared these models to determine what model best describes the behavioral data and thereby offers the most accurate view of such bias development. Since previous studies showed that choice biases caused by probabilistic cueing are mostly reflected in the position of starting points (Leite & Ratcliff, 2011, Mulder

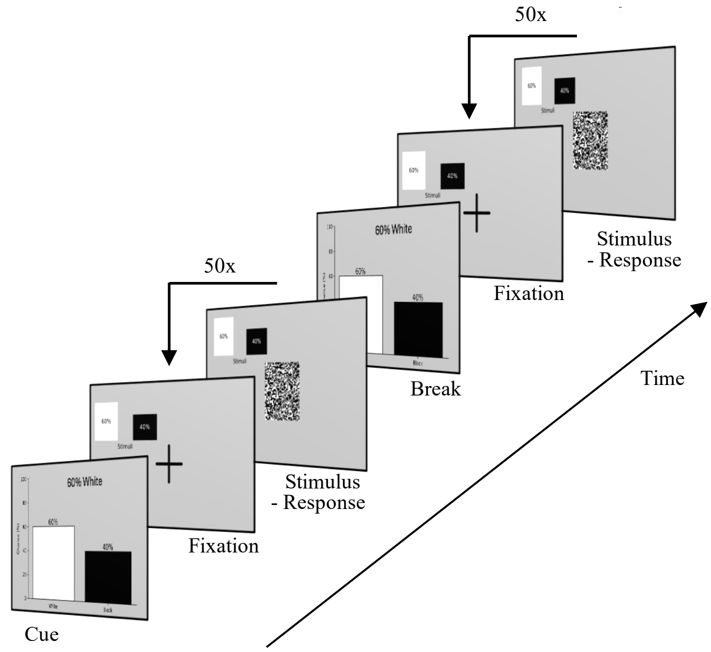

Figure 4. Schematic overview of one block. Each block starts with a cue, predicting the proportion of correct black and white choices both in text and with a bar graph. Then, the first 50 trials follow. Each trial consists of a fixation screen, shown between 600-1200 ms, and the stimulus, shown for 2000 ms. During the fixation and the stimuli, a simplified graph of the color prediction is shown. During the break, the cue predicting the proportion of correct black and white choices is repeated. When participants indicate they are ready, they continue with the last 50 trials.

et al., 2012, White and Poldrack, 2014), the focus in these four models lies on the starting point parameters.

In all four models, the decision thresholds, the bounds, correspond to black choices (upper bound) and white choices (lower bound). All models estimate a baseline starting point parameter across all trials for each participant, which reveals a potential bias for black or white, regardless of any prior information. The effect of the prior information is captured by an additional z-value (Δz). Together, the baseline and the additional z-value determine the starting point in each condition. Since in conditions 60-Match, 60-Mismatch, 80-Match, and 80-Mismatch the color predicted to be dominant was in fact dominant, even in the mismatch conditions, we assume the effect of the additional z-value information was always in favor of the color predicted to be dominant.

Therefore, when black is predicted to be dominant, the starting point will be determined by a baseline plus an additional z-value. When white is predicted to be dominant, the starting point will be determined by the baseline, minus the additional z-value. A higher z-value indicates a larger effect of the prior information on the starting point compared to the baseline, and thus a larger bias.

Previous research has shown differences in difficulties are usually reflected in the drift rate, since the efficiency of the uptake of information differs (Ratcliff & McKoon, 2008; Leite & Ratcliff, 2011). Therefore, in all models, there will be two drift rate parameters, one for easy and one for hard stimuli. Furthermore, all models have collapsing bounds, since all trials have a degree of urgency. Therefore, over time, the bounds shift closer to the middle, and less evidence is needed to reach a decision. For all models, we assume exponentially collapsing bounds, expressed by parameter B, the bound at t = 0, and parameter tau, the rate of collapse. Conventionally, the noise is set a 1 for all four models.

1) Null model. The Null model is designed after a theory wherein participants disregard both the prior information they receive beforehand and their experiences in previous trials. If this view is correct, all conditions result in the same starting point, solely determined by any natural biases for black or white. Therefore, the null model only has a baseline starting point parameter, z_{base}.

2) Prior model. This model reflects a view where participants behave primarily per the prior information, regardless of the accuracy of this information. If this theory is correct, there would be no difference in the starting points between an accurate and inaccurate prediction of color domination. Therefore, in this model, there are three starting point parameters: As in the Null model, a baseline z_{base} over all trials is calculated. Furthermore, there are parameters $z\text{-}60_{prior}$ and $z\text{-}80_{prior}$, which are the additional values (Δz) to the baseline when the prior information predicts a 60% (conditions 60-Match and 60-Mismatch) and an 80% (conditions 80-Match and 80-Mismatch) dominance of color, respectively. We assumed that when the prior predicts a white dominance, the starting points shift towards the lower bound ($z - \Delta z$) and when the prior indicates a black dominance, the starting point shifts towards the upper bound ($z + \Delta z$). Following the hypothesis, the conditions predicting an 80% dominance would result in an additional value (Δz) twice as large as the conditions predicting a 60% dominance.

3) Trials model. The Trials model is based on a hypothesis where biases are caused by the image of the ratio of black-and-white answers that participants develop throughout the block and not by the prior information. If this hypothesis is correct, all conditions with the same ratio of black and white being the correct answer result in the same starting point. Therefore, in this model, there are three starting point parameters. As in the Prior model and the Null model, there is a baseline parameter z_{base}. Furthermore, there are parameters $z\text{-}60_{trials}$ and $z\text{-}80_{trials}$, which represent the additional value to z_{base} when the actual color ratio is 60/40 (conditions 60-Match and 80-Mismatch) and 80/20 (conditions 80-Match and 60-Mismatch), respectively**.**

4) Full model. This model considers both a possible influence of the prior information and the perception of the color ratio during the trials. Therefore, in this model, there are five different starting point parameters. A baseline parameter z_{base} and additional value parameters for all conditions 60-Mismatch, 60-Match, 80-Mismatch, and 80-Match that, when added to the baseline, each result in a separate starting point.

To determine which of these four models best describes the empirical data, we compared both the Bayesian Information Criterion (BIC) (Schwartz, 1978) and the Akaike Information Criterion (AIC) (Akaike, 1974). The BIC is a criterion for model selection based on a likelihood function. The likelihood function indicates how well the model explains the data. The model fit can be increased by adding parameters. However, this might result in overfitting, which leads to poor generalization of new situations and reduced predictive performance. To avoid this, a penalty is added for the number of parameters. Therefore, The BIC reflects a balance between model fit and model complexity. We also calculated the AIC, which is more tolerant for complex models. Lower BIC and AIC values indicate a better fit.

For the BIC and the AIC values of all four models, see Table 2. The BIC and AIC values do not conclusively determine the winning model. The BIC identifies the Prior model as the winning model, while the AIC selects the Full model as the winning model. We decided to continue our analysis with the Full model, since this model allows us to test for the effect of inaccurate prior information.

To assess the performance of the full model, we conducted a goodness-of-fit analysis (Figures 5 and 6). By plotting the empirical data and the data predicted by the Full model, we can assess how well the Full model predicts the data. Datapoints closer to the diagonal suggest a good fit, since the the estimated data would be more similar to the empirical data.

A 2 × 2 (Prior × Match) repeated measures ANOVA on the starting points of the Full model in Jasp allows for a comparison between the biases developed in each condition and thus for a comparison between the adaptation to inaccurate expectations of high likelihood and the adaptation to inaccurate expectations of low likelihood, answering our main questions.

.

	Null model	Prior model	Trials model	Full model
Mean BIC (%)	163.37 (11.48%)	**132.89 (55.74%)**	140.11 (22.95%)	138.63 (9.84%)
Mean AIC (%)	138.12 (0%)	99.22 (27.86%)	106.44 (16.39%)	**96.54 (55.74%)**

Table 2. Mean BIC and AIC values for all four models, and the percentage of participants for which the model had the lowest value.

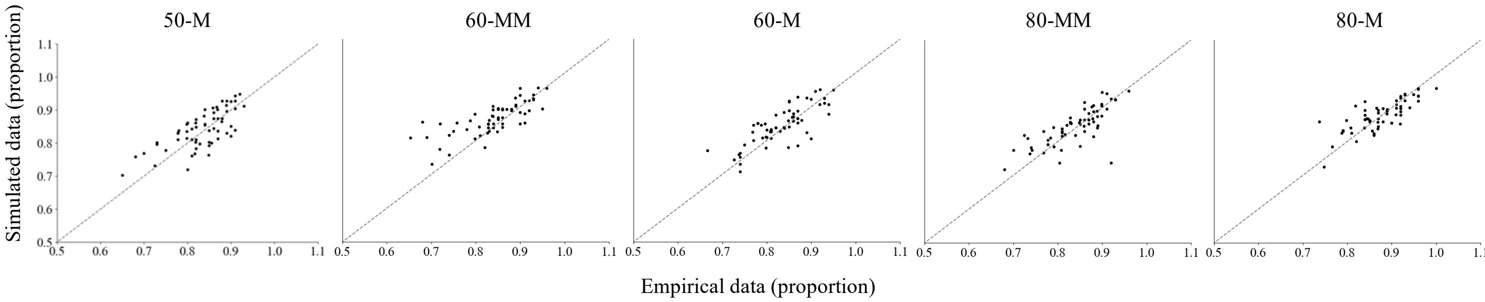

Figure 5. Goodness-of-fit for the proportion of correct answers for the Full model. For conditions (from left to right) 50-Match, 60-Mismatch, 60-Match, 80-Mismatch, and 80-Match, the graph shows for each participant the actual overall proportion of correct choices (x-axis) and the overall proportion of correct choices predicted by the full model for that same participant (y-axis).

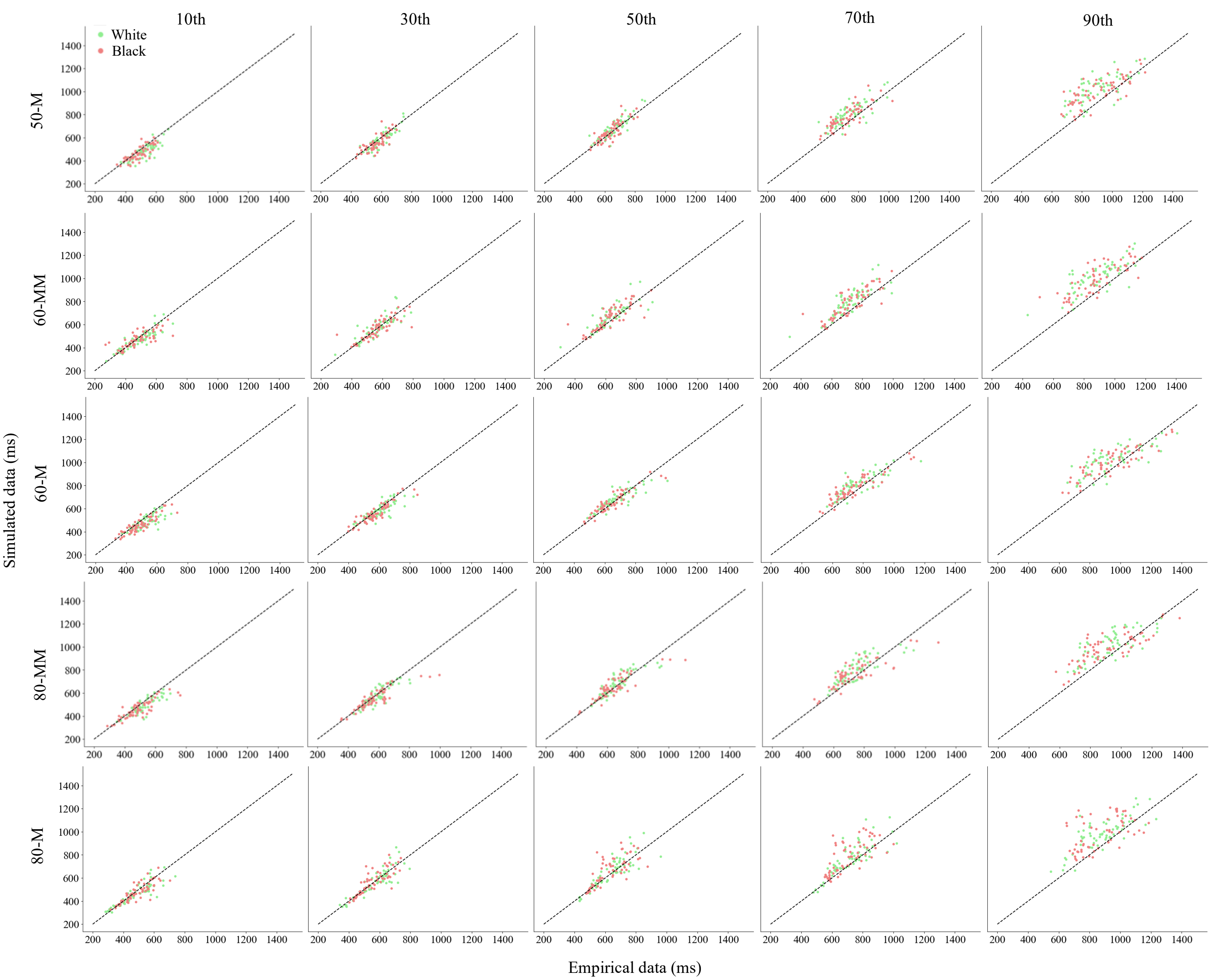

Figure 6. For each participant the empirical (x-axis) and the simulated data (y-axis) for conditions (from top to bottom) 50-Match, 60-Mismatch, 60-Match, 80-Mismatch, and 80-Match. From left to right the 10th, 30th, 50th 70th, and 90th quantile of the reaction time distribution, separated by choice for black or for white. The model seems to fit the data moderately well, except for a tendency to overestimate the reaction times for the 70th and 90th quantile.

Results

Descriptive results

Accuracy. For the descriptive statistics of accuracy within the biased trials, see Table 3. We conducted a 3 × 2 (Prior × Difficulty) repeated measures ANOVA. Prior has levels 50, 60, and 80, corresponding to the accurately predicted dominance of one color. Difficulty has levels easy and hard, the difficulties of the stimuli. There are significant main effects for Prior ($F(2,120) = 49.8$, $p < 0.001$) and for Difficulty ($F(1,60) = 169.9$, $p < 0.001$). Furthermore, no interaction effect between Prior and Difficulty was found ($F(2,120) = 5.7$, $p = 0.004$).

Reaction time. For reaction times, we compared the mean reaction times in the 50-Match condition with the mean reaction times for biased choices in 60-Match and 80-Match, which are the choices in favor of the option that was predicted to be more dominant by the prior information. The reaction times can be found in Table 4. A 3 × 2 (Prior × Difficulty) repeated measures ANOVA shows a main effect for Prior ($F(2,120) = 41.1$, $p < 0.001$) and a main effect for Difficulty ($F(1,60) = 9.2$, $p = 0.003$). There is no significant interaction effect between Prior and Difficulty ($F(2,120) = 0.6$, $p = 0.551$).

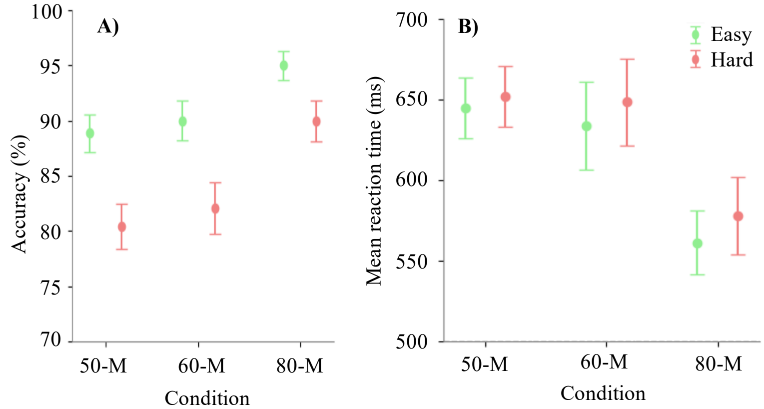

Figure 7. A) Accuracy (%) within the choices in favor of the option that was predicted to be more dominant, separated by prior predictions of 50, 60 or 80 dominance and by difficulties easy and hard. For prior strength 50, where an equal dominance is predicted, all trials are considered. **B)** Mean of the individual median reaction time (ms) for all choices for the option predicted to be more dominant separated by prior strengths 50, 60, and 80 and by difficulties easy and hard. For prior strength 50 all choices are considered.

	50-M	60-M	80-M
Accuracy (%) easy (SD)	88.9 (6.8)	90 (7.2)	95 (5.3)
Accuracy (%) hard (SD)	80.4 (8.2)	82.1 (9.3)	90 (7.4)

Table 3. Accuracy (%) for both easy and hard stimuli in conditions 50-Match, 60-Match, and 80-Match. For conditions 50-Match, it is the accuracy for all trials. For 60-Match and 80-Match, it is the accuracy for all biased trials, which are the trials of the color that was predicted to be more dominant. For a visualization, see Figure 7A.

	50-M	60-M	80-M
Mean RT easy (SD)	645 (74.2)	634 (108.9)	561 (78.1)
Mean RT hard (SD)	652 (75)	648 (107.3)	578 (95.5)

Table 4. Mean of all individual's median reaction times (ms) for easy and hard stimuli for conditions 50-Match, 60-Match, and 80-Match. For a visualization, see Figure 7B.

Biased choices. To indicate the effect of the prior information on the bias development in the match versus mismatch conditions, we compared the proportion of choices made in favor of the color predicted to be more dominant by the prior information. This comparison was made within the trials where the biased color, the dominant color, was the correct choice (for descriptives, see Table 5). Additionally, we made the comparison within the trials where the unbiased color, the non-dominant color, was the correct choice (for descriptives, see Table 6).

We performed a 2 × 2 × 2 × 2 (Prior × Match × Difficulty × Biased-trials) repeated measures ANOVA. Prior has levels 60 and 80, corresponding to the level of dominance predicted by the prior information. Match has levels yes and no, referring to the accuracy of the prior information. Difficulty has levels easy and hard, for the difficulty of the stimuli. Biased-trials has level 'biased', for the trials where the color predicted to be dominant was the correct choice, and level 'unbiased', for the trials where the non-dominant color was the correct choice

The ANOVA shows main effects for Prior ($F(1,60) = 40.492$, $p < 0.01$) and for Biased-trials ($F(1,60) = 1748.2$, $p < 0.001$). Furthermore, there is an interaction effect between Prior and Match ($F(1,60) = 22.7$, $p < 0.001$). Post hoc t-tests show a significant difference between condition 80-M and conditions 80-MM, 60-MM, and 60-M (80-MM: $t(60) = 4.8$, 60-MM: $t(60) = 5.7$, 60-M: $t(60) = -7.6$, all $ps < 0.001$). No significant differences were found between conditions 60-M and 60-MM ($t(60) = -2.9$, $p = 0.031$). The p-values were adjusted for multiple comparisons using the Bonferroni correction.

Biased Trials	60-M	60-MM	80-MM	80-M
Easy-mean (SD)	90.5 (7.2)	91.3 (6.9)	94.0 (5.8)	95.0 (5.3)
Hard-mean (SD)	80.4 (9.3)	81.8 (10.2)	83.4 (8.9)	89.7 (7.4)

Table 5. Per condition, the proportion of choices (%) made for the color predicted to be more dominant by the prior information, within the trials where the dominant color was the correct choice. Separated by the difficulty of the stimulus. For a visualization, see Figure 8.

Unbiased Trials	60-M	60-MM	80-M	80-M
Easy-mean (SD)	12.7 (10.4)	19.7 (18.7)	24.3 (15.8)	24.3 (15.8)
Hard-mean (SD)	24.8 (12.4)	30.0 (16.8)	37.7 (19.4)	37.7 (19.4)

Table 6. Per condition, the proportion of choices (%) made for the color predicted to be more dominant by the prior information, within the trials where the non-dominant color was the correct choice. Separated by the difficulty of the stimulus. For a visualization, see Figure 9.

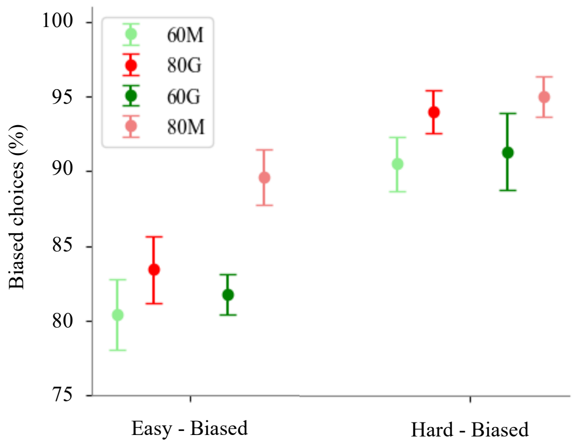

Figure 8. Per condition the proportion of biased choices (%) within the biased trials, separated by difficulty.

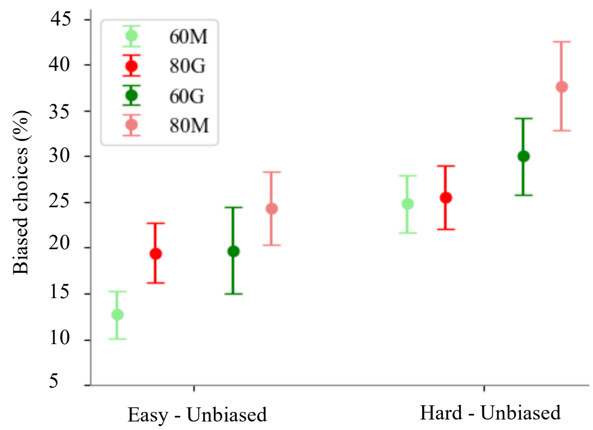

Figure 9. Per condition the proportion of biased choices (%) within the unbiased trials, separated by difficulty.

DDM analysis

The full model seemed to best fit the data (see method section for the BIC and AIC values and the goodness-of-fit). In this section, we will analyze the parameters of the full model to determine the choice biases caused by prior information and perception of the dominance of one color in conditions 60-M, 60-MM- 80-M, and 80-MM.

All parameters of the full model can be found in Table 6. For z_{base}, a value above 0 implies a natural choice bias for black and below 0 for white. A one-sample Wilcoxon signed-rank test shows that z_{base} is significantly higher than 0 (V = 1454.0, $p < 0.001$). *z-60-MM, z-60-M, z-80-MM*, and *z-80-M* are the additional value to z_{base}. These values represent the effect of the prior information on the bias development. A larger z-value indicates a larger bias.

Since we are specifically interested in the bias development caused by the prior information, we performed a 2 × 2 repeated measures ANOVA (Prior × Match) on the additional values in conditions 60-MM, 60-M, 80-MM, and 80-M. Conventionally, the values are scaled between 0 and 1. The values can be found in Table 7 and a visual representation in Figure 10.

We found a significant main effect for factor Prior ($F(1,60) = 46.5$, $p < 0.001$). However, no significant main effect for Match ($F(1,60) = 2.3$, $p = 0.137$) and interaction effect between Prior and Match ($F(1,60) = 3.4$, $p = 0.070$) were found.

	60-M	**60-MM**	**80-MM**	**80-M**
z-value	0.038	0.053	0.077	0.110
Std. Dev	0.06	0.07	0.06	0.08

Table 7. Additional z-values of conditions 60-MM, 60-M, 80-MM, and 80-M, relative to z_{base}. Larger z-values indicate a larger bias for the color that was predicted to be dominant. Negative z-values would indicate a choice bias for the color predicted to be non-dominant.

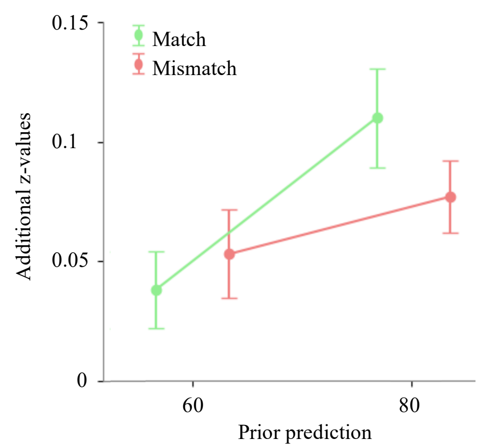

Figure 10. Additional z-values for conditions 60-Match, 60-Mismatch, 80-Match, 80-Mismatch.

	v_{easy}	v_{hard}	z_{base}	*z-60-MM*	*z-60-M*	*z-80-MM*	*z-80-M*	*B*	*tau*	*Ter*
Mean	1.529	1.061	0.050	0.106	0.076	0.154	0.220	1.013	0.864	0.301
Std. Deviation	0.405	0.290	0.096	0.147	0.127	0.121	0.164	0.269	0.427	0.053

Table 6. Mean and standard deviation for all parameters for the full model. v_{easy} and v_{hard} are the drift parameters, corresponding to easy and hard stimuli. z_{base} is a baseline estimate for the starting point, separate from any prior information. *z-60-MM, z-60-M, z-80-MM*, and *z-80-M* are estimates of the additional value to the baseline, representing the effect of the prior information in conditions 60-MM, 60-M, 80-MM, and 80-M, respectively. *B* is the decision threshold, *tau* represents the collapsing rate and *Ter* is the non-decision time.

DISCUSSION

In this study, we aimed to explore the adaptation to misinformation. We first investigated whether expectations of different degrees of likelihood of an event induce proportional biases, i.e. whether the difference in the extent of a choice bias between a 60% versus 80% expected likelihood of an event happening is twice the size of the difference between a 50% versus 60% expected likelihood. In addition, we were interested in the extent to which participants adapt their expectations in response to inaccurate information. Therefore, we compared the biases caused by prior information accurately predicting the probabilities of an event with the biases caused by prior information inaccurately predicting the probabilities. We were particularly interested in any potential differences between adaptation to prior information inaccurately predicting a high likelihood of an event versus inaccurately predicting a low likelihood of an event.

We conducted a two-alternative forced choice task where participants indicated whether an image contained more black or white. They received prior information about the likelihood of one color being correct. We manipulated both the probability strength and prediction accuracy, and recorded choices and reaction times. We then fitted a drift diffusion model on each participant's behavioral data to identify their choice biases.

We hypothesized that accurate expectations of the probability of an event will induce proportional biases. Furthermore, we expected participants to be well able to detect inaccuracies in their expectations and update their beliefs, reflected in their biases, accordingly. Finally, we expected a better adjustment to prior information inaccurately predicting a high likelihood of an event than a low likelihood, since an inaccurate expectation of a high likelihood of an event might be more noticeable.

For each participant, their bias was determined by the starting point parameters, calculated by the DDM. The biases in condition 80-Match were significantly higher than in the 60-Match condition. This aligns with previous research, suggesting that biases arising from prior information predicting the probability of an event are primarily reflected in the starting point parameters (Leite & Ratcliff, 2011, Mulder et al., 2012, White and Poldrack, 2014). Although we expected the difference between the 60-Match and 80-Match conditions to be twice the size of the difference between the base starting point and the 60-Match condition, the difference in the starting points is approximately the same size between the base starting point and the starting point in the 60-Match condition (0.038 increase) as the difference between the 60-Match and the 80-Match condition (0.034 increase). Also considering the fact no interaction effect was found between factors Prior and Match in the starting point parameters, indicating that the inclusion of the mismatch conditions did not influence the extent of the biases developed in the match conditions, this refutes the hypothesis of biases developing proportionally to prior information predicting a likelihood of an event.

Apparently, a decrease in uncertainty is not necessarily paired with a constant increase in the bias and differs for varying degrees of uncertainty, while this degree of uncertainty should not matter rationally. For instance, there is no rational reason why the difference in bias between a 74% and 75% expected likelihood differs from the bias difference between a 75% and 76% expected likelihood.

The starting points in the base, 60-Match, and 80-Match conditions suggest that biases might develop according to a pattern of saturation. Initially, when the uncertainty is at its maximum and no color dominance is predicted, the bias is at its minimum (base). Then, when the uncertainty decreases, bias increases rapidly (60-M), but this change diminishes over time (80-M), suggesting that bias does not increase proportionality with the prior likelihood. Alternatively, biases might develop according to a sigmoid logistic relationship, where as the expected likelihood of an event increases, they initially exhibit exponential growth, then gradually slow down before reaching an asymptote.

An explanation could be that people may categorize probabilities into specific ranges. For instance, we may consider probabilities of 50%-55% as ‘around even chance,’ 56%-75% as ‘likely,’ 76%-90% as ‘very likely’, and above 90% as ‘almost certain’. When the extent of a bias is determined by these mental categories and heuristics, the result may be a bias smaller or bigger than ‘rational’. For example, an expected likelihood of 60% might fall just at the lower end of such a range, causing a slightly higher than rational bias, while an expected likelihood of 80% might fall at the upper end, resulting in a slightly lower than rational bias. However, to address this adequately, further research testing a wider range of probabilities is necessary.

Regarding the adaptation to inaccurate prior expectations, the lack of significant differences in starting points between both the 60-Match and 60-Mismatch and between the 80-Match and 80-Mismatch conditions proves the ability to timely detect inaccurate information or update expectations of the likelihood accordingly was below anticipated levels.

Although the starting points in the Mismatch conditions did not differ significantly from the starting points in the Match conditions, participants chose the option predicted to be dominant significantly less often in the 80-Mismatch condition than in the 80-Match condition, while there was no significant difference in the proportion of biased choices between the 60-Match and 60-Mismatch conditions.

Combined with the fact that in the questionnaire after the experiment a substantial majority (88.5%) reported suspecting the prior information may not always have been accurate, this raises the question of whether there would have been significant differences in bias development in match versus mismatch conditions if participants were given a wider opportunity to update their beliefs. This could be achieved, for instance, by increasing the contrast between the predicted and the actual likelihoods or by providing participants with more time to detect the inaccurate predictions and update their beliefs.

However, it is worth noting that even if significant differences were observed between the match and mismatch conditions, any potential disparities in adapting to prior information inaccurately predicting either a low or high likelihood of an event may not solely be attributed to these varying strengths of predicted likelihood. That is, in the 60-Mismatch condition, the prior information underestimates the likelihood of an event, whereas the 80-Mismatch condition overestimates the likelihood. When either an over- or an underestimation of a predicted likelihood would be more noticeable, this might result in a more accurate adaptation to the inaccurate information in these conditions. Therefore, in future research, it can be considered to remove this potential influence by ensuring that both conditions involve either an overestimation or an underestimation.

Furthermore, the lack of significant differences in starting points estimated by the Full model between the 60-Match and 60-Mismatch conditions, as well as between the 80-Match and 80-Mismatch conditions, explains why the Prior model also resulted in low BIC and AIC values. Indeed, it appears that biases are primarily driven by prior information, regardless of its accuracy.

The question remains whether we have designed the drift diffusion models to best capture the participants' behavior. In all four models, we did not take into account the possibility of the biases being reflected in the drift rate. While the changes in bias caused by prior probability cues are primarily reflected in changes in the starting points (Leite & Ratcliff, 2011, Mulder et al., 2012, White and Poldrack, 2014), there are instances of predictive probability cues resulting in changes in the drift rate as well as in the starting points. For instance, in Dunovan et al. (2014), where cues predicted a 50%, 70%, or 90% chance of a stimulus being a face or a house and these predictive cues influenced both the starting point and the drift parameters. Thus, future studies could consider exploring potential changes in the drift rate as well.

Additionally, it is important to note that the classic DDM we used does not dynamically account for learning effects. A combined DDM with a learning model, such as RLDDM (Reinforcement Learning Drift Diffusion Model), could provide a more comprehensive understanding of the process of belief updating in response to inaccurate prior information.

In conclusion, this study sheds light on the complexity of adaptation to inaccurate expectations of the likelihood of an event happening. We found evidence against the idea that accurate likelihood expectations induce proportional biases. Furthermore, while participants generally showed awareness of potential inaccuracies in the prior information, they still strongly adhered to their prior expectations, as evidenced by the lack of significant differences in starting points between match and mismatch conditions. The data does indicate a trend towards better adjustment when the prior information inaccurately predicts a high likelihood compared to a low likelihood, as seen in the proportion of biased choices. However, additional research is necessary to conclusively address this.

REFERENCES

[1] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/tac.1974.1100705

[2] Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N. Z., & Evershed, J. K. (2019). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

[3] Dunovan, K., Tremel, J. J., & Wheeler, M. E. (2014). Prior probability and feature predictability interactively bias perceptual decisions. *Neuropsychologia*, *61*, 210 221. https://doi.org/10.1016/j.neuropsychologia.2014.06.024

[4] Dutilh, G., & Rieskamp, J. (2015). Comparing perceptual and preferential decision making. *Psychonomic Bulletin & Review*, *23*(3), 723–737. https://doi.org/10.3758/s13423-015-0941-1

[5] Kelly, S., Corbett, E. A., & O'Connell, R. G. (2019). Multifaceted adaptation of the neural decision process with prior knowledge of time constraints and stimulus probability. *bioRxiv (Cold Spring Harbor Laboratory)*. https://doi.org/10.1101/715318

[6] Leite, F. P., & Ratcliff, R. (2011). What cognitive processes drive response biases? A diffusion model analysis. *Judgment and Decision Making*, *6*(7), 651–687. https://doi.org/10.1017/s1930297500002680

[7] Mulder, M. J., Wagenmakers, E., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the Brain: A diffusion model analysis of prior probability and potential payoff. *˜the œJournal of Neuroscience/˜the œJournal of Neuroscience*, *32*(7), 2335–2343. https://doi.org/10.1523/jneurosci.4156-11.2012

[8] Rangel, A., Jr., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. In *Nature Reviews Neuroscience* (Vol. 9, Issue 7, p. 545). https://doi.org/10.1038/nrn2357

[9] Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295x.85.2.59

[10] Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

[11] Ratcliff, R., Smith, P. L., Brown, S., & McKoon, G. (2016). Diffusion Decision Model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

[12] Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2). https://doi.org/10.1214/aos/1176344136

[13] Shinn, M., Lam, N. H., & Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *eLife*, *9*. https://doi.org/10.7554/elife.56938

[14] Summerfield, C., Blangero, A., & University of Oxford. (2017). Perceptual Decision-Making: What do we know, and what do we not know? In *Decision Neuroscience* (p. 149). Elsevier Inc. http://dx.doi.org/10.1016/B978-0-12-805308-9.00012-9

[15] Tobeña, A., Marks, I., & Dar, R. (1999). Advantages of bias and prejudice: an exploration of their neurocognitive templates. *Neuroscience & Biobehavioral Reviews/Neuroscience and Biobehavioral Reviews*, *23*(7), 1047–1058. https://doi.org/10.1016/s0149-7634(99)00036-6

[16] Voß, A., Rothermund, K., & Voß, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*(7), 1206–1220. https://doi.org/10.3758/bf03196893

[17] White, C. N., & Poldrack, R. A. (2014). Decomposing bias in different types of simple decisions. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(2), 385–398. https://doi.org/10.1037/a0034851