Dissociable effects of cue validity on bias formation and reversal

Angelo Pirrone (apirrone1@sheffield.ac.uk)

Department of Psychology and Department of Computer Science, Sheffield University, Sheffield, S102TN, United Kingdom
Department of Psychology, Peking University, Beijing, 100871, China

Qi Zhang (zhangqipsy@qq.com) and Sheng Li (sli@pku.edu.cn)

Department of Psychology and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, 100871, China.
PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, 100871, China.
Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, 100871, China.

Abstract

In two experiments we manipulated the prior probability of occurrence for two alternatives. After a first learning session, in a second session the cue to bias the decision was reversed. Our investigation shows that subjects are able to learn the reverse bias only when the bias of the first session is in line with their expected outcome. When, during the first session, the actual outcome of the bias is not in line with the expected outcome, there is an inhibition for the reversal bias learning in the second session. We investigate this phenomenon with computational models of choice showing that the inhibition of reversal is due to an increase in the rate at which subjects accumulate evidence for repeated, unexpected stimuli. We discuss a possible theoretical explanation that links this phenomenon to similar results found in the literature on reversal learning and to the effect of novelty on learning.

Keywords: bias; reversal learning; drift diffusion model; random dot kinematogram

Introduction

The Drift Diffusion Model (DDM, Ratcliff & McKoon, 2008) is a celebrated model of decision making that has been shown to fit data in a broad variety of - but not limited to - perceptual decision making tasks (Ratcliff & McKoon, 2008; Bogacz et al., 2006). In the DDM, the decision-maker integrates difference in evidence supporting two alternatives until one of two decision boundaries is reached and a decision is made in favour of that alternative.

The DDM is determined by seven parameters: (1) the boundary separation, a, which describes the distance between the two decision boundaries; (2) the drift rate, v, which refers to the rate at which noisy information is accumulated over time within a trial and reflects the difficulty of the task; (3) the starting point of evidence accumulation, z, that describes whether the decision-maker starts to integrate evidence near one of the two boundaries (when z is equidistant from the two boundaries, the decision process is *unbiased*); (4) a parameter that incorporates the non-decision time component of a reaction time, ter, which is the time to encode the stimulus and execute the motor response; (5) inter-trial variability in drift rate, eta, (6) inter-trial variability in the starting point, sz and (7) inter-trial variability in the non decision time component, st.

In a typical task known as the random-dot kinematogram (RDK), the decision-maker is asked to identify the direction of a coherent subset of dots on a computer screen; typically, a percentage of dots move randomly while the remaining dots move in a coherent direction (Ratcliff & McKoon, 2008). Mulder et al. (2012) used an RDK task and biased their participants toward answering 'left' or 'right' by either manipulating the prior likelihood of occurrence for one of the two alternatives or the reward associated with one of the alternatives. The likelihood manipulation consisted of an arrow pointing left or right that was presented before the occurrence of each trial. The reward associated with one of the two alternatives was manipulated by assigning different payoffs to the occurrence of the alternatives. Mulder et al. (2012) found that both manipulation of bias, either as prior probability or as potential payoff, can be mapped onto a change in starting point of evidence accumulation, while the rate at which the evidence is accumulated is not affected. In our investigation, we are interested on whether and how, once a perceptual bias is 'learned', it can be modified and which parameters of the decision process are associated with the modification of bias.

Experiment 1

We focused on bias reversal: the situation in which, after a bias is learned, subjects are exposed to the opposite bias. We are interested in subjects' behavioural performance and in which parameters are affected both in the first bias learning phase and in the bias reversal learning phase. In our first study, subjects were performing an RDK task similar to the one used in Mulder et al. (2012). For simplicity, we define 'congruent' the session where there is, most of the time, congruency between the arrow and the direction of the dots (as compared to the 'incongruent' session) and we define the single trials as 'valid' or 'invalid' as a function of the correspondence between the arrow and the direction of the dots. A group of subjects was presented with a 'congruent' session while a second group was presented with an 'incongruent' session. After a considerable number of trials during which subjects learned the bias, the bias was re-

versed. In this way, the group that in the first phase of the experiment was presented with a congruent bias was afterwards switched to an incongruent bias and vice versa. Our results show that only the group that switched from a congruent cue to an incongruent cue (for simplicity defined as the 'c to i' group) could learn to modify the bias while the second group that switched from an incongruent cue to a congruent cue (the 'i to c' group) did not modify the bias when it was reversed.

Subjects

Six healthy subjects (4 female, mean age = 23.3, SD = 0.6) performed an RDK task. For this and the following studies, subjects were recruited among Peking University students and had normal or corrected-to-normal vision. Their participation was voluntary and rewarded monetarily. The procedure was approved by the ethical review board at Peking University and informed consent was obtained from each subject.

Stimuli

The task consisted of an RDK task in which the prior probability of occurrence of one alternative was manipulated before the occurrence of each trial by presenting an arrow pointing left or right. Subjects were asked to fixate upon a cross on the centre of the screen, pay attention to the arrow that was presented before each RDK trial, and make their decision for each trial by pressing 'left' or 'right' on a keyboard. They were instructed to be as fast and accurate as possible in making a decision. The stimuli were similar to those used by Mulder et al. (2012): within a circle aperture of 5 deg, white dots with a size of 3x3 pixels were moving with a speed of 3 deg/s and a density of 16.7 $dots/deg^2/s$ on a black background. On each frame, noise dots followed a random, but constant direction. Signal dots had a limited lifetime of three frames after which they were redrawn in random locations. On each three consecutive frames (i.e. from 1 to 3, from 4 to 6 and so on) the dots constituting the signal were the same. The stimuli were generated on a personal computer using PsychoPy (Peirce, 2007) and presented on a 36 x 27 cm CRT screen with a refresh rate of 60 Hz at a viewing distance of 56 cm where the head of the subject was positioned on a chin rest. As done by Mulder et al. (2012), we matched the difficulty of the task across subjects. For this purpose, subjects performed a block of 200 trials of randomly presented stimuli with different coherence levels (respectively, 3, 10, 20, 30, or 40% coherence, 40 trials each) and by using the maximum-likelihood estimation procedure described in Palmer et al. (2005), for each subject, we computed the motion strength required for an expected accuracy level of 80%.

The cue to bias the decision was a white arrow presented on the centre of the screen, pointing to the left or to the right and indicating the correct response 80% of the times. One block consisted of 80 trials and after each block, subjects could take a self-paced break before continuing to the next trials. Subjects received 20 ¥ for their participation and were told that they could earn up to 50 extra ¥ if they would have reached 100% accuracy during the task. No penalty was introduced for a wrong answer. After each trial they were given a feedback ('correct' or 'wrong') and at the end of each block they were shown their average accuracy for the block. Subjects were also presented with neutral trials, in which instead of the arrow, they were presented only a rectangular shape equal to the rectangle constituting the body of the arrows of the biased trials.

Paradigm timing

Subjects performed 4 consecutive blocks of each biased session (i.e. the congruent or incongruent) and 4 consecutive blocks of neutral trials of an RDK task. At the beginning of each trial, subjects were presented a fixation cross for either 500 or 1500 ms after which the cue was presented for 1000 ms, followed by a fixation cross with a duration of 1000, 1500, 2000, or 2500 ms. Next, the RDK was presented for 1500 ms during which subjects had to make a decision by button press. After each trial, subjects were presented a feedback for 500 ms showing the number of points earned and whether the previous trial was correct or wrong. If subjects answered faster than 150 ms or slower than 1500 ms, they were presented with the words 'miss' (in Chinese) or 'too fast' (in Chinese) for 500 ms. Session-order was counterbalanced across subjects with a group performing a congruent-to-incongruent-cue manipulation and a second group performing an incongruent-to-congruent-cue experiment. The order of presentation of the neutral blocks was counterbalanced across subjects.

Behavioural analyses

A first inspection of the data showed that there was no cross-session learning and the performance of the subjects after the first block of each session was stable. For simplicity, in all the analyses we collapsed the data across blocks and we created five experimental conditions: (1) congruent session valid trials; (2) congruent session invalid trials; (3) incongruent session valid trials; (4) incongruent session invalid trials and (5) neutral trials.

To investigate the effects of our manipulation, we entered correct RTs, wrong RTs and accuracy levels, separately for each group, in three different 2 x 2 mixed-effect models with Session (congruent, incongruent), Validity (valid trials, invalid trials) and the interaction between Session and Validity as fixed factors and random effects for subject-specific constants and slopes, Figure 1. On
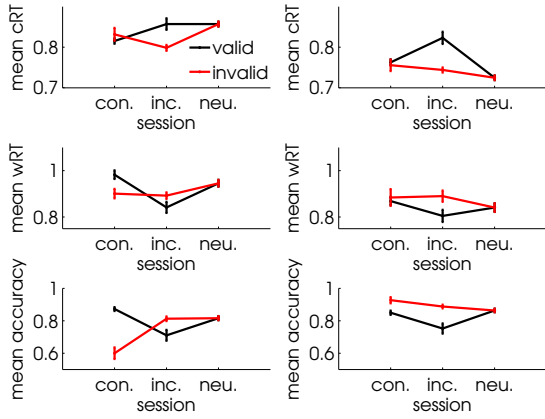
Figure 1: Mean correct RTs (first row), wrong RTs (second row) and accuracy levels (third row) for the 'c to i' group (left) and the 'i to c' group (right). Bars indicate standard errors of the group mean.

correct RTs, for the 'c to i' group, the interaction between Session and Validity resulted almost significant [$F_{(1, 15.10)} = 3.45$, $p = .08$]. On correct RTs, for the 'i to c' group, no significant effect was found ($p > .43$). On wrong RTs, for the 'c to i' group, the main effect of Session resulted significant [$F_{(1, 6.55)} = 8.63$, $p = .02$] and the interaction effect between Session and Validity was also significant [$F_{(1, 26.13)} = 7.56$, $p = .01$]. Post-hoc tests with Bonferroni correction showed that, in the congruent session, subjects had faster errors for the invalid trials ($M = .90$, S.E.$= .03$) than for the valid trials ($M = .98$ S.E. $= .02$; $p = .03$). On wrong RTs, for the 'i to c' group, no significant effect was found ($p > .46$). On accuracy levels, for the 'c to i' group, the interaction between Session and Validity resulted significant [$F_{(1,11)} = 15.49$, $p = .002$]. In the congruent session, subjects were less accurate for the invalid trials ($M = .61$, S.E. $= .10$), than for the valid trials ($M = .87$, S.E. $= .05$; $p = .04$). On accuracy levels, for the 'i to c' group, no significant effect was found ($p > .30$).

Model fitting

We fitted the diffusion model to RT distributions and proportions of correct and wrong responses using the Diffusion Model Analysis Toolbox (Vandekerckhove & Tuerlinckx, 2007) for MATLAB in which parameters are estimated by minimising a chi-square function. We decided to represent the reaction time distributions of correct and error responses in terms of five deciles, the 1st, 3rd, 5th, 7th and 9th decile of the RT distribution.

We fitted each subject separately, and for each subject we allowed the drift rate to vary freely across conditions and the starting point of evidence accumulation to vary

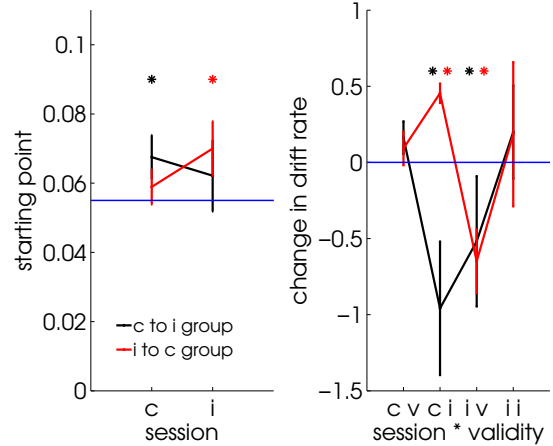as a function of the session (congruent vs. incongruent).

Figure 2: Effects of bias in DDM parameters. Average effects across subjects for starting point (left) and for change in drift rate with regards to the drift rate of the neutral trials (right). The blue horizontal line represents the unbiased level *a/2* (left) and the level at which the drift is zero (right). Asterisks show whether Wilcoxon Signed Rank tests resulted significant. X axis refers to (c v) congruent session valid trials (c i) congruent session invalid trials (i v) incongruent session valid trials and (i i) incongruent session invalid.

Figure 2 shows how the different parameters of the DDM were affected during the task; in particular we show how the starting point varied across sessions and what is the proportional change in drift rate as compared to the drift rate of the neutral trials.

Wilcoxon Signed Rank tests were performed to assess whether the starting point of evidence accumulation differed from the unbiased level (H0: median value of starting point = a/2, $p < 0.05$) and whether the proportional change in drift rate (with regards to the neutral trials) differed from zero (H0: median value of $\Delta v / v_{unbiased} = 0$, $p < 0.05$). Results are reported in Figure 2.

group	*a*	*ter*	*eta*	*sz*	*st*
c to i	0.11	0.63	0.15	0.04	0.30
i to c	0.11	0.63	0.32	0.06	0.32

Table 1: Parameter's values for the 'c to i' and the 'i to c' group.

In Table 1 are reported the mean values of the boundary separation, *ter*, *eta*, *sz* and *st* for the two groups; these parameters, being practically identical between the two groups, did not differ between groups ($p > .11$) as shown by mixed models with Group as fixed factor and random effects for subject-specific constants.

Experiment 2

In a second study, we tested the hypothesis that subjects in the 'i to c' group could not reverse the bias since they had to firstly learn a rule not in line with their expected outcome (i.e. that the arrow would point in the future direction of the dots) and this prevents them from overwriting the new rule when the bias is reversed.

We tested this hypothesis by manipulating the expectations of the subjects: all subjects performed an incongruent-to-congruent experiment as the 'i to c' group of the first study. However, before starting the experiment, one experimental group performed a training block in which there was always congruency between the arrow and the direction of the dots (congruent training group, for simplicity 'c training' group), and a second group performed a training block in which there was always incongruency between the arrow and the direction of the dots (incongruent training group,'i training' group). Our prediction was that only the group that had an incongruent training and for which the following incongruent session of trials was in line with their expected outcome (i.e. that the cue gives information about the *opposite* direction of the dots as experienced during the training) could modify their bias, as done by the 'c to i' group of the first experiment. We expected that the group that had the congruent training would instead exhibit a performance similar to that of the 'i to c' group of the first experiment, since they would not be able to reverse their bias, having learned a first bias that was in contrast with the expectations from the habituation manipulation. In this experiment, no neutral trials were presented.

Materials and Methods

Seven healthy subjects (4 female, mean age = 22.29, SD = 1.38) performed the second experiment. The apparatus and methods are as described for the first experiment. In this experiment, all subjects performed a congruent session after having performed the incongruent session. We manipulated the expected outcome of the first session by use of habituation. At the beginning of the experiment, subjects were presented two blocks of 80 trials and they were instructed that these trials were training trials to familiarise with the task. In this training phase, the arrow was indicating the valid or invalid direction of the trials 100% of the times. We created two groups: 4 subjects always had an invalid training while 3 subjects always

Behavioural analyses

performed a valid training.

We entered correct RTs, wrong RTs and accuracy levels, separately for each group, in three different 2 x 2 mixed-effect models with Session (incongruent, congruent), Validity (valid, invalid) and the interaction of Session and Validity as fixed factors and random effects for
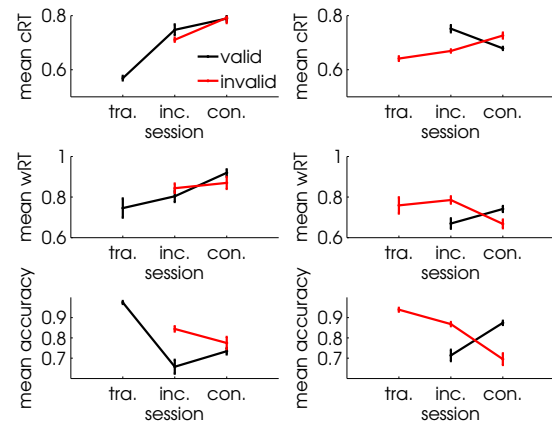

Figure 3: Mean correct RTs (first row), wrong RTs (second row) and accuracy levels (third row) for the 'c training' group (first column) and the 'i training' group (second column).

subject-specific constants and slopes, Figure 3. On correct RTs, for the 'c training' group, no significant effect was found (p>.49). On correct RTs, for the 'i training' group, the interaction between Session and Validity resulted significant [F (1, 13.43) = 7.14, p = .02]. In the incongruent session subjects were faster for the invalid (M = .67, S.E. = .05) than for the valid trials (M = .76, S.E. = .05; p=.09). On wrong RTs, for the 'c training' group, no significant effect was found (p>.49). On wrong RTs, for the 'i training' group, the interaction between Session and Validity resulted significant [F (1, 27.03) = 7.6, p = .01]. In the incongruent session, subjects had faster errors for the valid trials (M = .67, S.E. = .05) than for the invalid trials (M = .79, S.E. = .05; p=.03). On accuracy levels, for the 'c training' group, no significant effect was found (p>.11). On accuracy levels, for the 'i training' group, the interaction between Session and Validity resulted significant [F(1,12) = 8.03, p =.02]. In the congruent session subjects were less accurate for the invalid trials (M = .75, S.E. = .09), than for the valid trials (M = .90, S.E. = .04; p = .03).

Model fitting

We recovered DDM parameters for each subjects as done in the first study. Figure 4 shows how the different parameters of the DDM are affected during the task, both for the 'c training' and the 'i training' group. Wilcoxon Signed Rank tests were performed to assess whether the starting point of evidence accumulation differed from the unbiased level and whether the proportional change in drift rate (with regards to the training trials) differed from zero. Results are reported in Figure 2. The mean values of the other parameters of the DDM, that did not differ
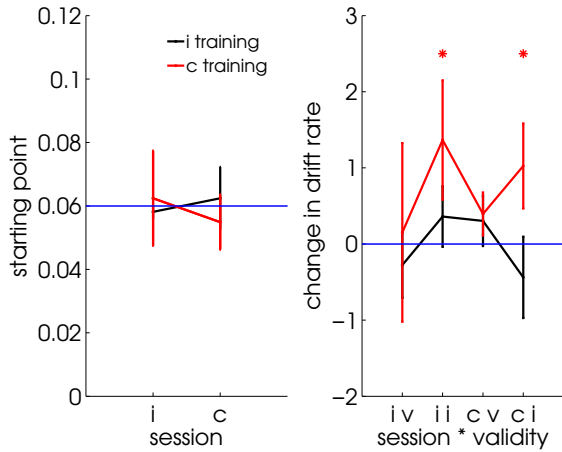
Figure 4: Effects of bias in DDM parameters. Average effects across subjects for starting point (left) and for drift rate as compared to the drift rate of the training trials (right).

between groups (p > .79), are reported in Table 2.

group	a	ter	eta	sz	st
i training	0.12	0.49	0.2	0.03	0.27
c training	0.12	0.53	0.14	0.02	0.31

Table 2: Parameter's values for the 'c training' and the 'i training' group.

Discussion

Our results show that when subjects learn a bias that is in line with their expectations, they are able to modify the bias in a second reversal session, but when they learn a bias that is not in line with their expectations, they are not able to modify the bias in a second reversal session. Only the group that firstly learned a bias in line with their expectations, either naive expectations (study 1) or controlled expectations (study 2), could become faster and more accurate for the trials for which they were slower and less accurate in the first session. A third investigation that we do not report in detail in this paper, showed that, for the 'i to c' group of the first study, even though the number of trials of the reversal session is doubled, subjects (n = 5) still cannot learn to modify the bias. In this way we rule out the hypothesis that what prevented subjects from learning the reversal bias was the length of the session that could be not long enough to give time to learn the reversal bias.

Similar results can be found in the literature on reversal learning showing inhibition of reversal dependant on the initial conditions of learning (Lindenblatt & Delius, 1988; Pace et al., 1980; Richman & Coussens, 1970;

Reid, 1953). For example, Newman et al. (1980) have found that, for human subjects, learning a discrimination based upon the presence of a feature is easier than learning a discrimination based upon the absence of a feature. However, when there is a reversal, so that subjects who had learned a rule based upon the presence (absence) of a feature are then switched to a new session in which they have to learn the new rule based upon the absence (presence) of the rule, only subjects in the presence-to-absence group can master the reversal while subjects in the absence-to-presence group cannot master the reversal. This same result is replicated in multiple experiments on pigeons, monkeys, rats and humans by different methods and typology of stimuli for the discrimination (Lindenblatt & Delius, 1988; Pace et al., 1980; Richman & Coussens, 1970; Reid, 1953; Newman et al., 1980). In these studies, learning discrimination to a low criterion inhibits the acquisition of the reversal rule, while learning it to a high criterion does not interfere with learning the reversal, but instead facilitates the reversal in some cases (Lindenblatt & Delius, 1988; Reid, 1953).

We believe that our results can be due to the fact that different cue-stimulus correlations have, depending on their sign, different attention-enhancing effects in the first session of the experiment that are maintained throughout the task. When this correlation is positive, the attention-enhancing effect is low since the rule is easily understood and the decision is stimulus-driven in a bottom-up fashion allowing the possibility of over-writing the rule in the second session. However, when the cue-stimulus correlation is negative, the attention-enhancing effect is high and the decision is rule-driven in a top-down fashion; this leads to a stronger learning of the difficult rule, that as a consequence, inhibits overwriting this rule with the new reversal rule. The attention-enhancing effect represents a cuing advantage for unexpected, repeated stimuli in order to learn the rule and maximise the reward during the first session. The decision-maker, right after the presentation of the cue, pays attention to the direction predicted by the unexpected stimulus (i.e. in the first study, in the opposite direction of the arrow) and this means that more information can be obtained from such directions, resulting in faster and more accurate decisions when the prediction is matched compared to when it is not matched. However, during the reversal, the decision-maker is still more biased towards the opposite direction and hence the previous advantage becomes a disadvantage. The attention-enhancing effect elicited by a first cue not in line with the expectations could be revealed by the higher drift rates that subjects of the 'i to c' group of the first study and subjects of the 'c training' group of the second study show for the invalid trials of the second session. The rate at which information is accumulated to make a decision

can be modulated by attention, and an attentional gain resulting in a signal-to-noise increase (i.e. faster RTs and higher accuracy levels) has been described in perceptual decision making (Smith et al., 2004). This interpretation is in line with the behavioural results of the first and second study. Also, our explanation is in line with experimental evidence showing that the magnitude of the difference between the expected outcome and the actual outcome results in an increase in dopamine levels that is associated with attentional orienting (Daw et al., 2006).

In the literature, two mechanisms have been proposed for the effect of bias: one affects the starting point of evidence accumulation (Mulder et al., 2012; Diederich & Busemeyer, 2006), while the second affects the rate at which evidence is accumulated (Bogacz et al., 2006; Diederich & Busemeyer, 2006). An effect of bias on the starting point is in line with that found by Mulder et al. (2012). However, together with the variation in starting point, we show that the drift rate can also be affected by a bias.

It is not surprising that subjects find pointing in the direction of the arrow as the expected outcome. Behne et al. (2012) have shown that even 12-month old infants show an understanding of communicative pointing, and Butterworth (2003) has proposed that pointing gestures spontaneously emerge from other developmental achievements playing a pivotal role during development. Pointing in the opposite direction of a stimulus of interest is rather an experimental artefact than a behaviour naturally observed.

Future studies should investigate whether the theoretical explanation proposed is consistent across different sets of stimuli and tasks, and answer the broader general question of *why*, from an ecological perspective, the inhibition of reversal in various tasks and domains characterise human and animal cognition.

References

Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds comprehension and production of pointing. *British Journal of Developmental Psychology*, *30*(3), 359–375.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review*, *113*(4), 700.

Butterworth, G. (2003). Pointing is the royal road to language for babies. *Pointing: Where language, culture, and cognition meet*, 9–33.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879.

Diederich, A., & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound-change, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics*, *68*(2), 194–207.

Lindenblatt, U., & Delius, J. D. (1988). Preventing a feature-positive effect in pigeons. *The American journal of psychology*, 193–206.

Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: a diffusion model analysis of prior probability and potential payoff. *The Journal of Neuroscience*, *32*(7), 2335–2343.

Newman, J. P., Wolff, W. T., & Hearst, E. (1980). The feature-positive effect in adult human subjects. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(5), 630.

Pace, G. M., McCoy, D., & Nallan, G. B. (1980). Feature-positive and feature-negative learning in the rhesus monkey and pigeon. *The American journal of psychology*, 409–427.

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of vision*, *5*(5), 1.

Peirce, J. W. (2007). Psychopypsychophysics software in python. *Journal of neuroscience methods*, *162*(1), 8–13.

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

Reid, L. S. (1953). The development of noncontinuity behavior through continuity learning. *Journal of experimental psychology*, *46*(2), 107.

Richman, C. L., & Coussens, W. (1970). Undertraining reversal effect in rats. *Journal of Experimental Psychology*, *86*(2), 340.

Smith, P. L., Ratcliff, R., & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: Response time distributions with masked and unmasked displays. *Vision research*, *44*(12), 1297–1320.

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*, *14*(6), 1011–1026.

Acknowledgements