



Published in final edited form as:

Q J Exp Psychol (Hove). 2010 May ; 63(5): 863–891. doi:10.1080/17470210903091643.

Do humans produce the speed-accuracy tradeoff that maximizes reward rate?

Rafal Bogacz^{1,2}, Peter T. Hu³, Philip J. Holmes^{2,4}, and Jonathan D. Cohen^{1,3}

¹Center for the Study of Brain, Mind and Behavior, Princeton University, Princeton, NJ 08544, USA

²Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA

³Department of Psychology, Princeton University, Princeton, NJ 08544, USA

⁴Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA

Abstract

In this paper we investigate tradeoffs between speed and accuracy that are produced by humans when confronted with a sequence of choices between two alternatives. We assume that the choice process is described by the drift diffusion model, in which the speed-accuracy tradeoff is primarily controlled by the value of the decision threshold. We test the hypothesis that participants choose the decision threshold that maximizes reward rate, defined as an average number of rewards per unit of time. In particular, we test four predictions derived on the basis of this hypothesis in two behavioural experiments. The data from all participants of our experiments provide support only for some of the predictions, and on average the participants are slower and more accurate than predicted by reward rate maximization. However, when we limit our analysis to subgroups of 30-50% of participants who earned the highest overall rewards, all the predictions are satisfied by the data. This suggests that a substantial subset of participants do select decision thresholds that maximize reward rate. We also discuss possible reasons why the remaining participants select thresholds higher than optimal, including the possibility that participants optimize a combination of reward rate and accuracy or that they compensate for the influence of timing uncertainty, or both.

Keywords

drift diffusion model; reward rate; speed-accuracy tradeoff

Introduction

During decision making in natural environments, as well as in experimental settings, humans and animals can select to be either fast or accurate. When accuracy is emphasized, their decisions are slower, but when speed is emphasized, they make more mistakes. This phenomenon is known as the speed-accuracy tradeoff (Franks, Dornhaus, Fitzsimmons, & Stevens, 2003; Pachella, 1974; Wickelgren, 1977).

The existence of the speed-accuracy tradeoff can be explained within the framework of sequential sampling models of decision making (Busemeyer & Townsend, 1993; Laming, 1968; Ratcliff, 1978; Stone, 1960; Usher & McClelland, 2001; Vickers, 1970). In this paper we focus on one of these models, the drift diffusion model (DDM), which is a continuous version of the Sequential Probability Ratio Test (Wald & Wolfowitz, 1948), and has been shown to fit behavioural data from human choice tasks (Ratcliff, 2006; Ratcliff, Gomez, & McKoon, 2004; Ratcliff & Rouder, 2000; Ratcliff & Smith, 2004; Ratcliff, Thapar, & McKoon, 2003). The DDM describes choice between two alternatives. It makes the following assumptions: (i) The sensory evidence supporting the alternatives is noisy. (ii) During the decision process the difference between the evidence supporting the two alternatives is integrated over time. (iii) When this integrated difference reaches a certain positive or negative threshold value, the choice is made in favour of the corresponding alternative. In the DDM, the height of the threshold controls the speed-accuracy tradeoff: If the threshold is lower, it can be reached quicker, so decisions are faster, but they are less accurate because they are based on a smaller amount of noisy evidence. Conversely, if the threshold is higher, decisions are slower, but they are also more accurate because they are based on the integration of more evidence.

However, an important question remains unanswered: What decision thresholds do people select, especially when they are not explicitly told to emphasize speed or accuracy, or are encouraged to favor some ill-defined combination of both (as in many experimental tasks)? A number of theories addressing this question have been suggested (Busemeyer & Rapoport, 1988; Edwards, 1965; Mozer, Colagrosso, & Huber, 2002; Myung & Busemeyer, 1989; Rapoport & Burkheimer, 1971). Recently, Gold and Shadlen (2002) proposed that participants select the threshold that maximizes the reward rate (RR), defined as an average number of rewards per unit of time. How the RR can be maximized depends on the paradigm.

Gold and Shadlen (2002) considered the following experimental paradigm: a participant is presented with a choice stimulus and is free to indicate his/her choice at any time. After the response, there is a fixed interval D between the response and the onset of the next trial. If the participants' choice is correct, he/she receives a reward, and if it is an error, the response-stimulus interval D is increased by an additional penalty delay D_p . This paradigm is often used in animal studies of decision making, as it captures some key aspects of decision making in natural environments (Gold & Shadlen, 2002).

In the above paradigm, RR depends on threshold height. If the threshold is too low, the participant makes fewer correct choices and receives fewer rewards; if the threshold is too high, response durations are so long that the participant receives fewer rewards per unit of time (even if most choices are correct). As shown by Bogacz et al. (2006), there exists a unique threshold for the DDM that maximizes RR. The ability to select this optimal threshold presumably conveys an evolutionary advantage, providing an animal more rewards than its competitors.

In our previous theoretical work (Bogacz et al., 2006), we have shown how the optimal threshold depends on task parameters, such as choice difficulty and experimental delays D , D_p . We also made experimental predictions regarding error rates (ER) and reaction times (RT) that must hold if participants indeed select the threshold maximizing RR. This article describes experiments that test the following four predictions:

1. The optimal threshold is higher for longer D , and hence the participants should have lower ER and longer RT in blocks of trials with longer D .

2. The optimal threshold depends only on $D+D_p$, rather than D and D_p individually, and hence ER and mean RT should be the same on blocks of trials with equal $D+D_p$, even if D and D_p themselves differ between blocks.
3. The values of decision thresholds estimated by fitting the DDM to the distributions of RT should match the optimal values that maximize RR.
4. The mean (normalized) RT, plotted as a function of ER, should follow a particular relationship defined by the optimal performance curve (described in detail in the next section).

In Bogacz et al. (2006), we also derived optimal thresholds under the assumptions that participants not only maximize RR, but also combinations of RR and accuracy. Here, we compare predictions of these analyses with experimental data.

This article is organized as follows. In the next section we review the DDM and the behavioural predictions following the assumptions that the participants select thresholds maximizing RR or combinations of RR and accuracy. Then we present two experiments and analyze their results based on all participants, showing that they provide partial support for the theoretical predictions. Participants, on average, set their thresholds higher than predicted by RR maximization. However, for the 30-50% of participants who achieved the highest reward over the entire experiment, all theoretical predictions are confirmed by the data. We also present data describing how quickly participants learn their thresholds, and then conclude with a general discussion of the experimental results and the relation of our work to other studies. Some initial results of this work were reported by Holmes et al. (2005).

Review of the Optimal Threshold Theory

Drift Diffusion Model

We begin by briefly describing two versions of the DDM: *pure*, that captures the main features of the decision process and that is easier to analyse mathematically; and *extended*, that fits more details of behavioural data from choice tasks (for more complete reviews see Bogacz et al., 2006; Ratcliff & Smith, 2004). All symbols used are listed in Table 1 for reference.

Let $x(t)$ denote the difference between evidence supporting the first and the second alternatives accumulated until time t . The pure DDM assumes that at the beginning of the decision process there is no bias towards either alternative, so $x(0)=0$, and that when the signal appears, $x(t)$ is integrated according to the following equation (Ratcliff, 1978):

$$dx=Adt+cW. \quad (1)$$

In Equation 1, dx denotes the change in x during a small time interval dt . This change in x includes two parts: constant drift Adt representing the average increase in x during interval dt , and noise cW which has a normal distribution with mean 0 and variance c^2dt , reflecting the assumption that sensory evidence, internal processing, or both are noisy. The sign of A represents which alternative is correct. For simplicity we consider the case $A>0$ for which the first alternative is correct. The ratio of the parameters A and c represents how easy the task is (i.e., the signal-to-noise ratio).

The DDM assumes that as soon the value of x reaches a positive threshold z or a negative threshold $-z$, the choice is made in favour of the corresponding alternative. A trial is

considered as correct if the threshold corresponding to the correct alternative (i.e., the upper threshold for $A > 0$) is reached, and the trial is considered as an error when the other threshold is reached (due to noise). For the pure DDM, the expected value of ER is given by (Ratcliff, 1978):

$$ER = \frac{1}{1 + e^{-\frac{2Az}{c}}}. \quad (2)$$

The mean decision time (DT), defined as the mean time of integration before reaching a decision threshold, is given by (Ratcliff, 1978):

$$DT = \frac{z}{A} \tanh\left(\frac{Az}{c}\right). \quad (3)$$

The DDM assumes that the RT is comprised of DT and an additional interval T_0 due to non-decision processes (e.g., visual and motor), i.e.,

$$RT = DT + T_0. \quad (4)$$

The pure DDM assumes that T_0 does not differ between trials, although it may differ among subjects. It is useful to note that if the parameters of the pure DDM: A , c , z are all scaled by the same constant, the ER and DT do not change. Thus instead of considering these parameters, it is simpler to consider their ratios. For simplicity of calculations we consider the following new parameters: a normalized threshold \tilde{z} (which corresponds to DT with zero noise) and signal-to-noise ratio squared \tilde{a} :

$$\tilde{z} = \frac{z}{A}, \text{ and } \tilde{a} = \left(\frac{A}{c}\right)^2. \quad (5)$$

For these new parameters¹ Equations 2 and 3 simplify to (Bogacz et al., 2006):

$$ER = \frac{1}{1 + e^{-2\tilde{a}\tilde{z}}}, \text{ and } DT = \tilde{z} \tanh(\tilde{a}\tilde{z}). \quad (6)$$

The extended DDM differs from the pure model in three assumptions (Ratcliff & Smith, 2004). First, in the extended model T_0 may differ between trials and it is assumed it comes from a uniform distribution with range $[T_0 - s_t, T_0 + s_t]$. Second, the initial value of x is not always equal to 0, but instead is chosen randomly from a uniform distribution with range $[-s_x, s_x]$. The non-zero value of $x(0)$ may reflect participants' prior expectations about probabilities of the alternatives to be correct, or the possibility that participants prematurely start to integrate x before stimulus onset.

¹Although these new parameters are ratios, they have units: \tilde{z} has units of time and \tilde{a} has units of 1/time (Bogacz et al., 2006); and while listing their values we take units of time to be seconds.

Third, in the extended model the drift is not the same from trial to trial, but instead is chosen randomly on each trial from a normal distribution with mean A and standard deviation s_A . Trial-to-trial drift variability may reflect differences in difficulty or attention between trials. The last two forms of variability allow the extended DDM to account for the differences between RT on correct and error trials (Ratcliff & Rouder, 2000).

Reward Rate Maximization

For the sequential choice task described in the Introduction, RR is given by the ratio of the fraction of correct trials and the average duration of the trial (Gold & Shadlen, 2002):

$$RR = \frac{1 - ER}{DT + T_0 + D + ER \cdot D_p}. \quad (7)$$

Let us first examine the predictions made by RR maximization, when it is assumed that the pure DDM provides a sufficient approximation of the decision process (we will return to the extended model below). To find the threshold maximizing RR, we substitute Equations 6 into 7, calculate the derivative of RR with respect to \tilde{z} , and find that, for this derivative to vanish, the normalized threshold \tilde{z} must satisfy:

$$e^{2\tilde{z}\tilde{a}} - 1 = 2\tilde{a}(D + D_p + T_0 - \tilde{z}). \quad (8)$$

We refer to the solution of this transcendental equation as the *optimal* normalized threshold and denote it by \tilde{z}_o . Although it does not admit an explicit solution in terms of elementary functions, Equation 8 has a unique solution (Bogacz et al., 2006) which forms a basis for the predictions outlined above, and labeled to remind the reader of their contents.

Prediction $z \sim D$ —The optimal threshold increases as D lengthens. It follows intuitively that, as opportunities to receive reward become less frequent, the accuracy of each choice becomes more important. This may also be shown formally².

Prediction $z \sim (D + D_p)$ —The optimal threshold depends only on $D + D_p$, rather than on D and D_p separately, because in Equation 8 the delays D and D_p appear only as a sum. This makes the non-intuitive prediction that participants should select the same threshold for a task that has a very short intertrial interval (D) but long delay imposed after errors (D_p), as for one that has a long intertrial interval but no penalty delay. For convenience, we denote the sum of the three delays influencing the optimal threshold together by D_{total} :

$$D_{total} = D + D_p + T_0. \quad (9)$$

Prediction $z = z_o$ —The optimal threshold \tilde{z}_o satisfying Equation 8 can be found numerically (since it is known that there is only one) and directly compared against thresholds chosen by a participant. In order to perform such a comparison for experimental data collected for given D and D_p , the parameters of the pure DDM \tilde{a} , \tilde{z} and T_0 need to be obtained by fitting the model to data. We refer to the normalized threshold \tilde{z} estimated from the data as the *participant's* normalized threshold. The values of D , D_p , \tilde{a} and T_0 can then be

²By differentiating Equation 8 with respect to D and noting that the derivative of \tilde{z}_o is always positive.

substituted into Equation 8, and the equation solved numerically for \tilde{z}_o . For optimality, the normalized threshold \tilde{z}_o obtained in this way should be equal to participant's normalized threshold \tilde{z} obtained from fitting the model.

Prediction DT(ER)—As we show in Appendix A, Equation 8 can be used to find the relationship between ER and DT that holds under the optimal threshold, to which we refer as the *optimal performance curve* (Bogacz et al., 2006). This relationship has the following form:

$$\frac{DT}{D_{total}} = f(ER). \quad (10)$$

The left side of the above equation expresses the ratio of the time in the trial used on decision processes to the maximum interval between the end of one decision process and the start of the next, while the right side is a function only of ER. Hence, the equation describes the relationship between ER and normalized DT as a fraction of D_{total} . This relationship, which contains only behavioral observables (DT and ER) and, in D_{total} , experimenter-determined delays and the non-decision time T_0 , is shown in black thick curves in Figure 1.

The above relationship should be satisfied for all values of signal to noise ratio \tilde{a} and for all values of total delay D_{total} . Thus the left end of the curve corresponds to very easy tasks on which the participants should be both fast and accurate. The right end of the curve corresponds to the tasks which are so difficult that the strategy maximizing RR is to guess without integrating any evidence at all. The optimal performance curve predicts that the longest normalized DT should be observed for ER around 18%, at which point DT should be equal to approximately 20% of D_{total} .

While the pure DDM is more tractable to analysis than the extended DDM, as noted above the latter has been used to account for a wider range of empirical phenomena. Therefore, we also consider whether the four predictions described above hold if it is assumed that the decision process conforms to the extended DDM.

Prediction $z \sim D$ —Numerical explorations of the extended DDM indicate that, like its pure form, the optimal threshold increases as delay D is lengthened.

Prediction $z(D+D_p)$ —Bogacz et al. (2006) have shown that, as for the pure DDM, the optimal threshold depends on D_{total} (rather than on D and D_p separately).

Prediction $z=z_o$ —The optimal threshold for the extended DDM, though different than for the pure DDM, can also be found numerically (by finding the threshold maximizing Equation 7 3), yielding alternative versions of prediction $z=z_o$.

Prediction DT(ER)—Unlike the pure DDM, an optimal performance curve independent of DDM parameters cannot be defined for the extended DDM. This is because the relationship between DT and ER depends on the variabilities s_x , s_A , as well as on D and D_p separately (Bogacz et al., 2006).

³Equation 7 includes ER and DT, for which there are no analytic expressions for the extended DDM, but we evaluated them by integrating numerically over a distributions of A , the expressions for ER and DT for a DDM with a variability of $x(0)$ (given in Bogacz et al., 2006).

Maximization of Reward Rate and Accuracy

It has been suggested that while choosing the decision threshold, participants not only optimize RR but also accuracy (Maddox & Bohil, 1998). We review two criteria in which accuracy is explicitly included (Bogacz et al., 2006). In both we assume for simplicity that $D_p=0$. The first criterion is a weighted sum of RR and accuracy, which may be written as:

$$RA = RR - \frac{q}{D_{total}} ER, \quad (11)$$

where q denotes the weight accorded to accuracy, and the second term is normalized by D_{total} to allow the units [1/time] to be consistent. The second criterion assumes that there is a penalty q for making an error⁴, and the modified RR is given by:

$$RR_m = \frac{(1 - ER) - qER}{DT + D_{total}}. \quad (12)$$

Assuming that the decision process is well approximated by the pure DDM, optimal performance curves can be derived describing the relationships between DT and ER for optimizing the threshold for the RA and RR_m criteria (see Appendix A). These have the following form:

$$\frac{DT}{D_{total}} = f(ER, q). \quad (13)$$

In both cases the right hand sides contain the additional weight parameter q , hence there exist families of optimal performance curves for different values of q . As q approaches 0, optimal performance curves for RA and RR_m converge to the optimal performance curve for RR, since in this case both criteria simplify to RR. These families are shown in Figures 1a and 1b. They are similar in that, for both criteria, DT for the optimal threshold is greater than for optimizing RR (this reflects the emphasis on accuracy). However, note also that these criteria differ in their predictions: For RA the value of q does not influence the ER corresponding to maximum DT (i.e., the position of the peak), while for RR_m increasing q moves the peak to the right. Furthermore, as q increases, the peak of the curve for RA becomes relatively narrower than for RR_m .

We note that the free parameter q in the criterion RA (or RR_m) can be adjusted so that *any single* combination of ER and DT (in a certain range) lies on the corresponding optimal performance curve. It is, however, not true that arbitrary *multiple combinations* of ER and DT must all lie on one such curve. Each criterion and choice of q predicts a curve with a particular shape, as shown in Figure 1, and multiple (ER, DT) data points may or may not follow this shape (e.g., if DT decreased monotonically across conditions as ER goes from 0 to 0.5, no curve could match the data). Hence these criteria must be assessed by comparing with ERs and DTs from multiple conditions and/or multiple participants, under the assumption that the participant(s) maximize the same criterion, with the same q , in all the conditions.

⁴ RR_m would describe reward rate in a modified paradigm in which a participant receives a unit of reward for a correct choice and loses q units of reward for an error. Thus in the standard paradigm q describes the relative negative utility of committing an error.

General Methods

We present the results of two experiments designed to test whether human decision makers set their thresholds in an optimal manner. In both experiments participants performed sequences of choices between two alternatives, but the experiments differed in the stimuli used. In the first experiment participants discriminated the direction of movement of dots (stimuli often used in decision studies in monkeys, e.g., by Shadlen and Newsome (2001)). In the second experiment participants discriminated whether the fraction of an array occupied by stars was greater or less than 50% (these stimuli were previously used by Ratcliff et al. (1999)). Both experiments were approved by the Institutional Review Panel for Human Subjects of Princeton University. Below we describe aspects common to both experiments, and later in the Methods sections of Experiments 1 and 2 we only describe aspects specific to individual experiments⁵.

Participants

These were adults recruited via announcements posted around the Princeton University campus. Participants were predominantly undergraduate and graduate students. Participants were paid 1 cent for each correct choice. To further increase motivation, participants were informed that the one who earned the most overall (in each group of 20 participants) would receive an additional prize of \$100 at the end of the experiment. All participants expressed written consent for participation.

Procedure

Participants were instructed to gain as many points as possible on a computerized two-alternative, forced-choice task by making correct choices. In each trial, a stimulus was presented and remained until the participant responded by pressing a corresponding key. After each correct response participants were informed by a short beep if the response was correct and a point was scored. No feedback was given following incorrect responses. After each response there was a delay D before presentation of the next stimulus (D was kept constant within each block, but varied across blocks). On some blocks (see below) an additional delay D_p was imposed after error responses.

Design

The length of a block of trials was limited by fixing overall block duration – rather than by fixing the number of trials completed within it – in order to enforce the importance of the speed-accuracy tradeoff. Trials were blocked by delay condition. There were four delay conditions: (1) $D=0.5s$; (2) $D=1s$; (3) $D=2s$; and (4) $D=0.5s$ and $D_p=1.5s$ (in the first three conditions $D_p=0$). Before the start of the experiment participants had three blocks of practice in which no money was paid for correct choices. Participants were informed about the number of blocks, their fixed duration, and that delays between trials (and difficulty in Experiment 2) differed between blocks, but were they not told the exact durations of the delays (or difficulty levels in Experiment 2).

After finishing the experiment, participants were asked to complete a questionnaire in which they rated (from 1 to 5) the difficulty of the experiment, the degree to which they were motivated by the one cent reward after each correct decision, and by the \$100 prize. Participants were also asked if their strategies differed between blocks (answer yes or no), and if so to describe how their strategies differed.

⁵Complete behavioural data from both experiments and the estimated parameters of DDM for all participants can be downloaded from: <http://www.cs.bris.ac.uk/home/rafal/optimal/data.html>

Experiment 1

Method

Participants—These were 20 adults (9 males and 11 females; average age: 20 years).

Stimuli and apparatus—We used the same stimuli that were used in other studies of decision making (e.g., Gold & Shadlen, 2003; Palmer, Huk, & Shadlen, 2005; Ratcliff & McKoon, 2008). The display was a field of randomly moving dots all of which appeared within a 5° circular aperture in the center of the screen. Dots were white squares 2 by 2 pixels (0.7° square) displayed against a black background, with a density of 16.7 dots/degree²/s (6 dots per frame). On each trial, a fraction of the dots moved in a single direction over time, corresponding to that trial's correct direction, while the remaining dots were randomly repositioned over time. On each frame, 11% of the dots were independently selected as the coherently-moving dots, and were shifted 0.2 deg from their position for each 40 ms (3 video frames) elapsed, corresponding to a speed of 5 deg/sec (either leftward or rightward). The remaining dots were re-plotted in random positions on each frame. The display was generated in MATLAB on a Macintosh computer using the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997) and software written by Joshua Gold.

Procedure—Participants were asked to decide whether the prevailing motion of the dots was left or right, and had to indicate their responses by pressing “M” (rightward motion) or “Z” (leftward motion) on a standard keyboard (the mapping of keys to the right and left responses was not counterbalanced across participants). Participants were required to release the key after each response in order to initiate the next trial. The current score was displayed in the center of the screen during delay intervals.

Design—The experiment consisted of 5 blocks, each lasting 7 minutes. One block of trials was run for each delay condition, except for condition $D=2s$. Two blocks were run the $D=2s$ delay condition because a single 7-minute block yielded too few trials for analysis.

Behavioral Results

Predictions $z \sim D$ and $z(D+D_p)$ stated in the Introduction were tested directly by comparing ERs and RTs on different conditions of Experiment 1 (see Figure 2). A oneway analysis of variance (ANOVA) on delay condition (with participant as a random effect) revealed significant differences between delay conditions in ERs ($F(df=3) = 5.84, p=0.0015$), but the differences between RTs did not reach significance ($F(df=3) = 2.02, p=0.12$).

Prediction $z \sim D$ —In the first three conditions ($D=0.5, D=1, D=2$) there was a significant negative correlation between D and ER (after subtracting each participant's mean ER across conditions from their ER in each condition; $r=-0.53, p<10^{-4}$), and a significant positive correlation between D and RT (after subtracting each participant's mean RT; $r=0.29, p=0.03$). Since we know that the ER and RT of the DDM are monotonic functions of the decision threshold, we can infer from Figure 2 that, on average, in the first three conditions ($D=0.5, D=1, D=2$) the participants chose higher thresholds in the longer delay conditions. This pattern is qualitatively consistent with prediction $z \sim D$.

Prediction $z(D+D_p)$ —To investigate the prediction that participants should choose the same threshold in delay conditions 3 ($D=2s$) and 4 ($D=0.5, D_p=1.5s$ so that $D_{total}=2s$), we compared the ER and RT between these conditions. Both ER ($t(df=19) = 2.23, p=0.04$) and RT ($t(df=19) = 2.49, p=0.02$) were significantly different between these conditions. This is not consistent with prediction $z(D+D_p)$. However, as discussed below, this prediction was largely satisfied for a subset of the best performing participants.

Estimating Parameters of the DDM

Testing predictions $z=z_o$ and $DT(ER)$ requires estimating parameters in the DDM for individual participants. For the purpose of comparison, we estimated parameters of both pure and extended models.

There are at least two tenable approaches to estimating the parameters of the DDM: (i) Fit a separate DDM to each participant and delay condition; or (ii) Assume that certain parameters (e.g., the non-decision part of reaction time, T_0) do not vary across conditions for a given participant, and fit the DDM with the constraint that these fixed parameters are constant across conditions. If the assumption of approach (ii) is correct, then this more constrained approach will give better estimates of parameters. Therefore, we first estimated the parameters of the pure DDM using the unconstrained method (i) and identified those parameters that did not differ systematically across conditions; we then treated these as fixed parameters using the more constrained method (ii), as described below.

For each participant and condition we used the unconstrained method to estimate the following parameters: non-decision time T_0 , signal to noise ratio \tilde{a} and decision threshold \tilde{z} (divided by drift rate)⁶.

Following Ratcliff and Tuerlinckx (2002), we divide the RT distribution into five quantiles (in units of seconds): 0.1, 0.3, 0.5, 0.7, 0.9, and denote these by RT_q^{th} and RT_q^{ex} for the pure DDM (theory) and data (experiment) respectively. For the moment, we do not distinguish between the distribution for correct and error trials, because for the pure DDM with fixed drift rate and initial condition, as employed here, their means are identical (Feller, 1968; we will return to the extended model later). We denote the error rates given by the pure DDM and observed in the experiment by ER^{th} and ER^{ex} respectively.

The subplex optimization algorithm (Rowan, 1990) was used to find parameters minimizing the cost function describing the weighted difference between ERs and RT distributions of the model and from the experiment (Ratcliff & Tuerlinckx, 2002):

$$Cost = w_{ER} (ER^{th} - ER^{ex})^2 + \sum_{q=1}^5 w_{RT,q} (RT_q^{th} - RT_q^{ex})^2. \quad (14)$$

In the above equation, w 's denote the weights of the fitted statistics. We choose the weight of a given statistic close to $1 /$ (the estimated variance of this experimental statistic), as

proposed by Bogacz and Cohen (2004). In particular we take: $w_{ER} = n / ER^{av}(1-ER^{av})$ and $w_{RT,q} = n / RT_q^{av}$, where ER^{av} and RT_q^{av} are ER^{ex} and RT_q^{ex} averaged across all delay conditions⁷. This averaging across conditions is done to avoid dividing by 0 in blocks in which the participant did not make any errors, and also because the differences in ER^{ex} and

⁶We estimate the ratios \tilde{a} and \tilde{z} rather than the original parameters of the DDM A , c , and z , because it is not possible to estimate A , c , and z uniquely (see Section Drift Diffusion Model). One could also fit different ratios (e.g. A/c and z/c), but we use \tilde{a} and \tilde{z} for consistency with the theoretical treatment, in which these ratios best simplified the mathematical analysis. Although we estimate \tilde{z} ($= z/A$) rather than z , we will still be able to determine whether a participant chooses the optimal threshold z_o : recall that Equation 5 implies that the optimal normalized threshold $\tilde{z}_o = z_o/A$, thus if the estimated \tilde{z} is equal to \tilde{z}_o , this also implies that $z = z_o$.

⁷The variance of ER can be estimated as follows: Assume, for a given condition, that a participant has a probability of making an error equal to p_e and there are n trials in this condition. Then the experimental error rate ER^{ex} comes from the binomial distribution with mean p_e and variance $p_e(1-p_e)/n$. For simplicity we estimate p_e as ER^{av} . The variance of RT_q^{ex} can be estimated using a method proposed by Maritz and Jarrett (1978), but here for simplicity we assume that this variance is proportional to the mean of RT_q^{ex} and inversely proportional to n , thus we simply take $w_{RT,q} = n/RT_q^{av}$.

RT_q^{ex} across conditions for single participants are small in comparison to differences between participants.

Using the unconstrained method described above, we fitted the pure DDM to all conditions of all 20 participants. The ANOVA did not reveal significant differences in either T_0 ($F(df=3) = 0.7, p=0.55$) or \tilde{a} ($F(df=3) = 0.92, p=0.44$) between delay conditions. To further investigate the differences in T_0 and \tilde{a} between delay conditions, we performed paired t-tests between all pairs of conditions, and none of the tests were significant ($p > 0.1$ for all tests). Therefore, in the constrained estimation method we assumed that a given participant has the same T_0 and \tilde{a} for all conditions. Accordingly, for a given participant we estimated six parameters: non-decision time T_0 , signal to noise ratio \tilde{a} , and four normalized decision thresholds \tilde{z} (one for each of the four delay conditions). We found parameters minimizing the following cost function (analogous to that of Equation 14).

$$Cost = \sum_{d=1}^4 \left(w_{ER} (ER_d^{th} - ER_d^{ex})^2 + \sum_{q=1}^5 w_{RT,q} (RT_{d,q}^{th} - RT_{d,q}^{ex})^2 \right). \quad (15)$$

In the above equation ER and RT denote the error rates and reaction time indexed additionally by the delay conditions. The parameters of the pure DDM found in this way were used in the analyses in the following sections.

To estimate parameters of the extended DDM from the data we used the DMAT toolbox (Vandekerckhove & Tuerlinckx, 2007) that finds parameters that maximize a (multinomial) likelihood function (Ratcliff & Tuerlinckx, 2002). In fitting the extended DDM we followed our treatment of the pure DDM by constraining all parameters except the decision threshold to be constant across conditions for any given participant.

Fits of the DDM

Figure 3 compares the fits of the pure and the extended DDMs to data from a representative participant (see caption for the explanation of the quintile probability plots). This participant produced different RTs for error and correct trials (note that open circles are in different positions in the left and right parts of the panels). As could be expected this difference was captured by the extended DDM, but not by the pure DDM which always produces the same RT distribution for error and correct trials (note that lines in Figure 3a are symmetric). Nevertheless, the pure DDM was able to capture the speed-accuracy tradeoff produced by the participant: Note that the lines in the right part of Figure 3a are increasing, which indicates that as the accuracy increases, the RT also increases. Thus, the pure DDM was able to fit quite well this tradeoff on correct trials (constituting a great majority of trials).

As could be expected, across the participants, there was a strong correlation between the parameter values estimated by fitting the pure DDM and the corresponding parameters estimated by fitting the extended DDM, as shown in Figure 4 (for T_0 , $r = 0.67$; for \tilde{a} , $r = 0.61$; for \tilde{z} , $r = 0.92$). Such strong correlations have been reported before (Wagenmakers, van der Maas, & Grasman, 2007). Nevertheless, Figure 4 also shows that there are systematic differences between parameters, namely the pure DDM underestimates \tilde{a} and overestimates \tilde{z} , in agreement with observations of Wagenmakers et al. (2007)⁸. Overall, however, our

⁸Wagenmakers et al. fitted DDMs with an assumption that parameter c is fixed. They observed that the pure DDM underestimates parameter A in comparison to the extended model. Since the definitions of \tilde{a} and \tilde{z} involve ratios of A , see Equations 5, it implies that the pure DDM underestimates \tilde{a} (proportional to A^2) and overestimates \tilde{z} (inversely proportional to A).

findings suggest that the simpler pure DDM is a useful tool in the analysis of the speed-accuracy tradeoff.

Decision Thresholds

Prediction $z=z_o$ —This is tested in Figure 5 which compares the estimated values of participants' thresholds (estimated from the DDM) with the optimal ones (found as described in Section Reward Rate Maximization), for the pure and extended DDM.

Figures 5a and 5b show data for a sample participant, whose thresholds are very close to the optimum values. To quantify the similarity between the normalized participant's and optimal thresholds, we do not use Pearson's correlation as it can give high values even if the mean of one set of thresholds is higher than of another (as long as there is a linear relationship between them). Thus to test Prediction $z=z_o$ Appendix B defines a correlation measure r_1 that assumes that the two sets of thresholds have the same mean. In Figures 5a and 5b these correlations are for the pure DDM: $r_1 = 0.89$ ($p = 0.07$), and for the extended DDM: $r_1 = 0.87$ ($p = 0.09$)⁹.

Figures 5c and 5d show the data for all 20 participants. While not all participants set their thresholds close to the optimal values, nevertheless there is a strong correlation between normalized participants' and optimal thresholds for the pure DDM: $r_1 = 0.44$ ($p < 10^{-5}$), and for the extended DDM: $r_1 = 0.62$, ($p < 10^{-5}$)¹⁰.

Figures 5e and 5f compare the participants' and optimal normalized thresholds for the four delay conditions, each averaged across all 20 participants. On average, participants appear to set their thresholds higher than the optimal values, an effect that is significant in the short delay conditions ($D=0.5s$ and $D=1s$) for the pure DDM and in all conditions for the extended DDM.

Comparing the left and right columns of panels in Figure 5 suggests that fitting the pure and the extended DDM reveals very similar relationships between the participants' and optimal thresholds. This similarity can be traced to the strong correlation between those parameter values estimated by fitting the pure DDM and those estimated by fitting the extended DDM (Figure 4).

In summary, while the optimal threshold theory explains a significant proportion of the variance in participants' thresholds, on average they tend to set their thresholds to values higher than optimal for maximizing RR – especially in the short delay conditions. We will return to this observation below.

Experiment 2

Method

Participants—These were 60 adults (30 males and 30 females). In addition to payments described in General Methods, participants were also paid \$8 for participation. Participants in Experiment 2 did not take part in Experiment 1.

⁹Note that since we found no significant differences in signal to noise ratios across conditions within participants (see Section Estimating Parameters of the DDM), it is reasonable to assume that drift A is also the same in all delay conditions, and all thresholds in Figure 5a are therefore normalized by the same value of drift. The plot comparing non-normalized thresholds of this participant with non-normalized optimal thresholds would look exactly the same as Figure 5a, but both scales of the axes would be divided by the value of the drift. So the non-normalized thresholds of this participant are also very close to the optimal values (and there is the same correlation $r_1=0.89$ between the participants' and optimal non-normalized thresholds).

¹⁰Note however that this does not imply that there is equal correlation between non-normalized participants' and optimal thresholds, because Figure 5c includes data from all participants and each participant has thresholds normalized by a different value of drift; unfortunately it is more difficult to say what the correlation would be between the non-normalized thresholds.

Stimuli and apparatus—Experiment 2 used the same stimuli that were used in a previous study of the DDM by Ratcliff et al. (1999). Participants were presented with a 10×10 grid in the upper left corner of a VGA monitor, subtending a visual angle of 4.30° horizontally and 7.20° vertically. Random cells within the grid were filled with asterisks; others were empty. On each trial, participants had to decide if the majority of locations in the grid were empty or filled with asterisks. The grid and asterisks appeared as light characters against a dark background, were presented with high brightness and contrast and were clearly visible. The VGA monitors were driven by a PC computer, and the stimuli were displayed using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). The number of displayed asterisks was either 40 or 60 during blocks referred to as *easy*, and 47 or 53 during blocks referred to as *difficult*.

Procedure—Participants had to indicate their choices by pressing “M” or “Z” key on a standard keyboard, and the mapping of keys to the “low” and “high” responses was counterbalanced across participants. The score was continuously displayed in the center of the screen. In the first series of 20 participants, some participants held one of the buttons down continuously on some blocks of trials (which resulted in $RT = 0$, and ER close to 50%). To prevent such behavior, the remaining participants were required to release both buttons in order to initiate the next trial.

Design—The experiment consisted of 10 blocks lasting 4 minutes each. For each delay condition there was an easy block and a difficult block, except for condition $D=2s$ for which there were four blocks (two easy and two difficult) to collect a sufficient number of trials for analysis.

Behavioral Results

A two-way ANOVA (delay \times difficulty) revealed that ER was significantly influenced by difficulty ($F(df=1) = 103, p < 10^{-4}$) and delay ($F(df=3) = 3.79, p = 0.01$), but not by their interaction ($F(df=3) = 0.67, p = 0.55$). Similarly, RT was significantly influenced by both difficulty ($F(df=1) = 870, p < 10^{-4}$) and delay ($F(df=3) = 9.01, p < 10^{-4}$), but not by their interaction ($F(df=3) = 0.52, p = 0.62$). Since there was no interaction between difficulty and delay conditions, in order to visualize the overall effect of delays, Figure 6 shows ER and RT averaged across difficulty conditions. The behavioral data from Experiment 2 follow the same pattern as in Experiment 1.

Prediction $z \sim D$ —In the first three conditions ($D=0.5, D=1, D=2$) longer delay was associated with lower ER and higher RT (correlation between D and participants' ER after subtraction of mean ER for given participant and difficulty condition was $r=-0.32, p < 10^{-4}$; correlation between D and participants' RT after subtraction of mean RT for given participant and difficulty condition was $r=0.31, p < 10^{-4}$). Hence, in the first three conditions ($D=0.5, D=1, D=2$) the longer the delay, the higher threshold participants chose, in agreement with prediction $z \sim D$.

Prediction $z(D+D_p)$ —To test the prediction that participants should choose the same threshold in delay condition 3 ($D=2s$) and 4 ($D=0.5, D_p=1.5s$), we compared the ER and RT between these conditions. ER did not differ significantly ($t(df=119) = 0.43, p=0.66$) but RT did significantly differ between these conditions ($t(df=119) = 2.88, p=0.005$). The latter finding (for RT) is not consistent with prediction $z(D+D_p)$. However, as discussed below, this prediction was largely satisfied for a subset of the best performing participants.

Estimating Parameters of the DDM

We only fit the pure DDM to the data from Experiment 2 because the data were too few to constrain the extended DDM (involving three more parameters); there were fewer trials per condition in Experiment 2 than in Experiment 1 (4 minutes instead of 7 minutes). More importantly, many participants made very few or no errors in the easy conditions of Experiment 2 and information on RT distribution on error trials is necessary to accurately estimate the parameters of the extended DDM. We note, however, that the pure and extended DDM produced highly-correlated parameters (Figure 4), and they provided similar information on the relationship of participants' and optimal thresholds (observed in Figure 5). Hence we feel that it is worthwhile to use the pure DDM in the analysis of threshold setting in Experiment 2.

The parameters of the pure DDM were estimated using the same method as in Experiment 1. Namely, for every participant and difficulty condition we estimated: T_0 , \tilde{a} and four normalized decision thresholds \tilde{z} (different values of both T_0 and \tilde{a} were estimated for each level of difficulty since these parameters differed significantly between difficulty conditions).

Decision Thresholds

Prediction $\mathbf{z}=\mathbf{z}_o$ —Figures 7a and 7b compare the estimated values of participants' normalized thresholds with the optimal ones. The correlation between the participants' and the optimal normalized thresholds is $r_1 = 0.67$ ($p < 10^{-5}$) in the easy condition, and $r_1 = 0.25$ ($p = 10^{-4}$) in the difficult condition.

Figures 7c and 7d compare the participants' and the optimal normalized thresholds averaged across all 60 participants. Again, participants tend to set their thresholds to higher than optimal values. The difference between participants' and optimal thresholds is significant in all conditions, except the easy condition in delay condition 4 ($D = 0.5$, $D_p = 1.5$).

Overall, the optimal threshold theory explains a significant proportion of variance in participants' thresholds, but participants consistently set their thresholds higher than the optimal values.

Performance of Participants with High Reward Scores

The match of empirical results to theory in Experiments 1 and 2 was short of perfect. Here, we explore factors that may have influenced how closely participants set their thresholds to the optimal value. We focus on the results of Experiment 2, for which there were enough participants to conduct reliable analyses.

To quantify the match between participants' thresholds and the theory, for each participant we compute a *theory match error* defined as the Euclidian distance between the participant's thresholds in the eight conditions of Experiment 2 and the optimal thresholds for those conditions:

$$E = \sqrt{\sum_{i=1}^8 (\tilde{z}_i - \tilde{z}_{o,i})^2}. \quad (16)$$

We found that this measure was not correlated with participants' responses to any of the questions in the debriefing questionnaire. However, it was strongly correlated ($r = -0.71$) with *reward score* (total amount of money earned), as shown in Figure 8a. This may not

seem surprising: The theory identifies thresholds that maximize RR, and so those participants who approximate these thresholds should be expected to earn the highest rewards. The theory match error of Equation 16 was weakly dependent on gender (two tailed unpaired t-test $t(df=58) = 2.13, p < 0.04$), i.e. on average males chose thresholds closer to values maximizing RR than females. We will come back to a possible interpretation of this dependence in General Discussion.

Let us now examine whether the performance of those participants who achieved reward scores above the median satisfies predictions $z(D+D_p)$ and $z=z_o$ that were not satisfied when all participants were considered.

Prediction $z(D+D_p)$ —Figures 8b and 8c show the resulting ER and RT for each delay condition averaged over this subset of participants. Paired t-tests showed that ER and RT did *not* differ significantly between delay conditions 3 ($D = 2s$) and 4 ($D = 0.5s, D_p = 1.5s$) (ER: $t(df=57) = 0.32, p = 0.75$; RT $t(df=57) = 0.16, p = 0.87$).¹¹ These findings are consistent with predictions of the optimal threshold theory: threshold should only depend on $D + D_p$.

Prediction $z=z_o$ —Figures 8d and 8e show the mean (normalized) participants' and optimal thresholds, averaged across the same set of participants, for the easy and difficult conditions. In the easy condition, the mean participants' threshold does not differ significantly from optimal in *any* of the delay conditions (Figure 8d). In the difficult condition, the mean participants' threshold is significantly different from optimal only in delay condition $D = 0.5$ (Figure 8e). Thus, participants with higher reward scores set their thresholds much closer to the optimal value than other participants (compare with Figures 7c and 7d).

Empirical Test of Relationships between ER and DT

Relationship Predicted by Maximization of RR

Prediction DT(ER)—This describes the relationship between ER and DT that holds for optimal performance across various task parameters. Here we examine how well this relationship describes actual human performance. The relationship is expressed in Equation 10, which relates ERs to the normalized DT (i.e., DT as a fraction of D_{total}). Therefore, for both experiments, we calculated the normalized DT for each participant in each task condition as: $(RT - T_0) / D_{total}$, where RT was the participant's mean for that condition, and T_0 was the non-decision part of the RT estimated for that participant from the pure DDM as described earlier. Each of these normalized DTs was associated with a corresponding ER for that participant in that condition. Figure 9a plots DTs from all experimental conditions as a function of ER for a sample participant whose DTs lie in the vicinity of values predicted by the optimal performance curve (see black curve in Figure 9a).

For the majority of participants however, the match between the normalized DT and the optimal performance curve is not as good as shown in Figure 9a. For each participant of Experiment 2, we computed the correlation between the normalized DTs for 8 experimental conditions and the normalized DT predicted by the optimal performance curve. Figure 9b shows the histogram of these correlations. Although these correlations are usually positive, they are statistically significant only for 12 participants.

To evaluate how well the optimal performance curve describes the data averaged across participants, we divided the possible range of ER (0-50%) into ten equal intervals and, for each interval, calculated the mean normalized DT over all participants and conditions with

¹¹However, both ER and DT did vary significantly for all other pairwise comparisons between conditions, with p ranging from $p = 0.018$ to $p < 10^{-4}$ (without correcting for multiple comparisons).

ERs falling into that interval. The results of this analysis are shown in Figure 9c. Participants' mean normalized DTs are higher than those predicted by the optimal performance curve for RR (black curve in Figure 9c). This is consistent with the results of the analyses in the Section Decision Thresholds (see Figures 5e, 5f, 7c, and 7d) indicating that, in aggregate, participants set their thresholds higher than is optimal to earn maximum reward. Note, however, that the predicted relationship provides a good qualitative description of the relationship between ER and DT. In addition to showing the same inverted-U shape, participants have the highest normalized DT for blocks in which their ERs range from 15% to 20%, bracketing the predicted value of 18%.

As demonstrated above, there is a strong correlation between participants' reward scores and how closely their thresholds match the theoretical optima. Therefore, we compared the optimal performance curves with the data from participants with different levels of reward score separately. In particular, some participants with the lowest reward scores had DTs an order of magnitude higher than other participants, and thus we separately analyzed the 10% of participants with the lowest RRs¹² (their DTs are shown in the right panel of Figure 9d). We split the remaining 90% of participants into 3 equal groups according to RR, but there was no significant difference in normalized DTs between the 30-60% and 60-90% groups (paired t-test across 10 bins of ER: $p > 0.2$), hence we pooled the data from these groups (constituting the middle 60% of participants). The normalized DTs for the remaining top 30% group and the middle 60% group are shown in left and middle panels of Figure 9d.

The left panel of Figure 9d indicates that the predictions of the optimal performance curve for RR match the data from the participants with reward scores in the top 30% quantitatively, and indeed their performance is statistically indistinguishable from the theoretical predictions except at one ER interval¹³. The other two groups of participants (middle 60% and bottom 10%) shown in Figure 9d have significantly higher normalized DTs than predicted by RR maximization.

Relationships Predicted by Maximization of RA and RR_m

We investigated whether the optimal performance curves predicted by maximization of the RA and RR_m criteria (combining RR and accuracy) can quantitatively capture the relationship between the ER and DT of the middle 60% and the bottom 10% of participants. Both performance curves include a free parameter q describing the emphasis placed on accuracy. For both curves, q was estimated via the maximum likelihood method, using the standard assumption that normalized DTs are normally distributed around the performance curve being fitted with a variance estimated from the data. Additionally, in fitting the curve for RA, we constrained $q \leq 1.096$, and in fitting the curve for RR_m we constrained $q \leq 1$ (see Appendix A). The values of estimated parameters are given in Table 2. Note that the groups of participants with lower performance have higher estimated values of q , which could be interpreted as greater emphasis on accuracy.

Figure 9e shows the best fitting performance curves for the three groups of participants. Both performance curves for RA and RR_m are able to describe the DTs of the middle 60% of participants. In case of the bottom 10%, the curve for RA fits the data better than the curve for RR_m , as the latter predicts that the longest normalized DT should be obtained for

¹²Two participants from Experiment 1, and six participants from Experiment 2.

¹³For each participant and condition we took the empirically observed ER and used Equation 18 to calculate the theoretically predicted optimal normalized DT. Then, for each of the ER intervals, we used paired t-tests to compare the distribution of experimentally observed mean normalized DTs with the distribution of predicted values. The difference was significant only in the first ER interval [0-5%], $p = 0.003$, and non-significant, $p > 0.05$ for all other ER intervals). The reason for the difference in the interval with the highest ER may be that the first interval contained many blocks with ER = 0, and Equation 18 predicts that for these blocks DT = 0, but participants had DT > 0.

higher ERs, which is inconsistent with the data. These observations are quantified in Table 3, which lists the ratios of the likelihoods of the data given different curves. The first row shows that the data from each group of participants are more likely given the curve for RA than given the curve for RR_m , in particular the data from the bottom 10% are more than 43 times more likely given RA than RR_m curves. Table 3 also shows that the curves for RA and RR_m provide much better fit to the data than the curve for RR for the bottom 70% of participants.

In summary, although the relationship between ER and normalized DTs observed in the experiments for the middle and the bottom groups of participants cannot be described quantitatively by the optimal performance curve for RR, it can be described by the performance curve based on the assumption that participants maximize RA, i.e., the weighted difference between RR and ER. This suggests that the “conservatism” of these participants may reflect a greater emphasis on accuracy than is optimal for maximizing reward rate. In the General Discussion we consider alternative account of this apparent emphasis on accuracy.

Adjustments in Decision Thresholds

In this brief section we investigate how quickly participants adjust their thresholds after block onset. Maximization of RR requires higher decision thresholds for longer D , and to visualize how participants adjust thresholds depending on D , Figure 10 shows differences between RTs in conditions $D = 2$ and $D = 0.5$. These differences are averaged over 5 trials in a given bin, over all blocks, and over all participants of Experiments 1 and 2. Figure 10a shows that RT differences develop and stabilize within the first ~20 trials for participants with reward scores above the median. RT differences for the remaining participants are shown in Figure 10b, but unfortunately the data is too noisy to identify when they stabilize. We will return to the issue of learning decision thresholds in the General Discussion.

General Discussion

Summary of Results

In the Introduction we stated four predictions that should be satisfied if participants select thresholds maximizing RR. We performed two experiments testing these predictions. The support provided by experimental data for our predictions is summarized in Table 4. In general, the predictions were only partially consistent with the data from all participants, but virtually all of them were satisfied by data from the participants with highest reward scores.

Prediction $z \sim D$ —The prediction that the increase in response-stimulus interval D should increase the decision threshold, and thus decrease ER and increase RT, was satisfied in the behavioural data averaged over all participants in both experiments.

Prediction $z(D+D_p)$ —The prediction that the threshold (and thus ER and RT) should depend only on $D + D_p$, rather than D and D_p individually, was not satisfied in the behavioural data averaged over all participants. However, the prediction was satisfied in the data averaged over the participants with reward score higher than the median.

Prediction $z=z_o$ —The prediction that the threshold estimated from fitting the DDM should be equal to the optimal threshold was only partially satisfied by the data from all participants. There was a significant correlation between the participants' and optimal thresholds, but in each of the experimental conditions the average participants' thresholds were higher than optimal. However, the estimated thresholds of participants of Experiment 2

with reward scores higher than the median was not significantly different from the optimal thresholds in 7 out of 8 experimental conditions.

Prediction DT(ER)—The predicted relationship between DT and ER was only qualitatively satisfied by the experimental data from all participants, for whom the experimental DTs were higher than optimal. However, the data from the top 30% participants with highest reward scores satisfied the predicted relationship quantitatively.

Taken together, the above results suggest that, on average, participants set threshold higher than required to maximize RR, but that a significant proportion of participants indeed select a speed-accuracy tradeoff that approximately maximizes RR. One could ask how surprising it is that the participants with highest RR follow the predictions of our theory, since our theory assumes RR maximization. Note however, that the predictions were also based on the assumption that the choice process can be described by the DDM. Different models of information processing during choice would require different speed-accuracy tradeoffs to maximize RR. However, our data suggest that a significant proportion of participant select a decision threshold required by the DDM to maximize RR, thus our results also provide a support for the DDM.

Pure Drift Diffusion Model

It is important to note that predictions $z \sim D$, $z(D+D_p)$, $z=z_o$ were made by both the pure and extended DDM while the prediction $DT(ER)$ could only be formulated on the basis of the pure DDM. The fact that a significant fraction of participants conformed to this prediction suggests that the pure DDM is a useful approximation of the decision process between two alternatives.

It is interesting to ask why the predictions of the pure DDM matched well with the behaviour of many participants despite large systematic differences in the parameter values estimated by the pure and extended DDM (Figure 4). A possible reason may be that the maximization of RR under pure and extended DDM predicts similar speed-accuracy tradeoffs for given task parameters (difficulty and D_{total}) (Bogacz et al., 2006). We have previously shown that introducing the variability of drift slightly decreases DT predicted for given ER, while the variability of starting point increases the predicted DT (see Figure 14 in Bogacz et al., 2006). Since the effects of the two forms of variability on predicted DT are in opposite directions, it is possible that the speed-accuracy tradeoff predicted by the extended DDM is similar to that predicted when both forms of variability are ignored.

Relationship to Other Studies on Selection of Speed-accuracy Tradeoff

A parallel study (Simen, Buck, Hu, Holmes, & Cohen, submitted) has also tested some of the predictions of our previous theoretical work (Bogacz et al., 2006), but its focus was different. It tested predictions regarding tasks in which one of the alternatives is either more probable or more rewarded: in such situations accuracy needs to be sacrificed to maximize reward even to greater extent than in the paradigm considered here (Bogacz et al., 2006; Maddox, 2002). Simen et al. (submitted) also tested predictions $z \sim D$ and $z=z_o$, but not predictions $z(D+D_p)$ ¹⁴ and $DT(ER)$. Furthermore, they collected substantially more data from 9 individual participants, allowing more precise fits of the extended DDM, while here we gather less data from each of many more (80) participants, allowing analysis of factors influencing participants' choices of decision threshold. Importantly, the results of Simen et al. (submitted) support those of the present paper. Specifically, Simen et al. also observed

¹⁴In the experiment of Simen et al. (submitted) the delays between response and the onset on the next trial were the same after correct and incorrect responses (i.e. $D_p=0$) in all blocks.

that participants' thresholds increased with D (prediction $z \sim D$ satisfied), and found that there was a high correlation between participants' and optimal thresholds, but on average participants set their thresholds above the optimal values (prediction $z = z_o$ partially satisfied).

Edwards (1965) proposed an alternative theory suggesting that participants choose decision thresholds that minimize the weighted sum $RT + qER$ (where q is a parameter determining the relative weight of ER, cf. Equations 11 and 12). Busemeyer and Rapoport (1988) designed an experiment in which the participants were explicitly required to minimize this cost function in order to maximize their rewards, and they found that the participants chose decision thresholds close to those minimizing this function. Their results and ours taken together suggest that participants are able to adjust their decision thresholds to increase reward in *various* experimental paradigms. Such an ability could come from an adaptive reinforcement learning mechanism aiming to maximize reward rate or reward in any task. We will return to this issue below.

Why do Some Participants Select Thresholds Higher than Optimal?

We considered the possibility that some participants select thresholds higher than optimal because they maximize a weighted combination of RR and accuracy, and we showed that this theory is able to predict the shape of the relationship between ER and DT. This theory has been widely applied to data from perceptual discrimination tasks in which stimuli belonging to two classes differ on a continuous perceptual dimension (e.g., length of a line). In this context a discrimination threshold corresponds to a value on this dimension such that stimuli below and above this value belong to two different classes. If one of the alternatives is more rewarded, the discrimination threshold maximizing reward is different from that maximizing accuracy. The theory that assumes maximization of a combination of reward and accuracy quantitatively accounts for the discrimination threshold used by participants, and for the observed dependence of the discrimination threshold on accuracy (Maddox, 2002; Maddox & Bohil, 1998).

Other explanations for why participants select thresholds higher than optimal have also been suggested. Recently Zackenhause et al. (submitted) investigated the assumption that participants are unable to precisely estimate the duration of the response-stimulus interval D and the value of signal to noise ratio \tilde{a} , and analyzed predictions of four theories proposing that they seek to achieve performances robust to the uncertainties in D and \tilde{a} . One of these theories assumes that participants maximize a guaranteed level of performance under a presumed level of uncertainty in estimation of interval D . They show that the data in Figure 9e can be fit slightly better by a performance curve based on this assumption than the performance curve for RA.

Alternatively, thresholds higher than optimal are also predicted under the assumption that participants try to maximize RR using an adaptive threshold adjustment procedure (e.g., driven by a reinforcement learning mechanism): In previous work we have shown that if participants do not have a priori knowledge of the threshold that maximizes RR, and therefore must discover this through adaptive approximation, on average they will achieve a higher RR by overestimation of the optimal threshold than by underestimation. This is based on the observation that, for the DDM, there is a greater drop off in RR for thresholds below the optimum than above it (Bogacz et al., 2006). Timing uncertainty (contributing to variability in reward rate estimation for a given threshold) could further contribute to this effect. This would force participants to optimize a distribution of thresholds which, for the reasons stated above, would have a mean value that is greater than the single optimal threshold.

Thresholds higher than optimal can also be expected under the standard economic assumption that the participants discount future rewards, i.e. they value a present reward more than the same reward in the future. If so, then participants value the reward on the current trial more than rewards on succeeding trials, and to increase the probability of reward on the current trial, they increase the threshold (Yakov Ben-Haim, personal communication, 20th July 2007). Adjudicating between these alternative hypotheses will require further research.

Motivational Effects on Threshold Learning

It has been demonstrated in a perceptual discrimination task that if participants receive rewards for correct choices and no penalty for errors, then the possibility of winning a prize for best performing participants tends to move participants' discrimination thresholds towards those maximizing reward (Markman, Baldwin, & Maddox, 2005). Thus it is possible that the reward of \$100 in our experiments also encouraged participants to set decision thresholds closer to those maximizing RR, and may partially explain why so many participants were nearly optimal.

Motivational effects of the reward for best participants could also explain the gender effects in Figure 8a. Namely, it is possible that the reward of \$100 in our experiments was on average more motivating for males than females, since males have been reported to be more motivated by competition (Gneezy & Rustichini, 2004) and have a higher tendency to gamble (Desai, Maciejewski, Pantalon, & Potenza 2005).

Influence of Task Difficulty on Match between Optimal and Participants' Thresholds

Figure 7 shows that there is a better match between optimal and participants' thresholds in the easy task condition than in the difficult one, and a number of speculations are possible about why this is the case. First, for a difficult task or long D_{total} (or for both), the optimal threshold may be high and correspond to a firing rate of decision neurons above the range in which their firing rates are linear functions of input. Since our theory is based on the simple linear DDM, it cannot predict the value of the optimal threshold outside of the linear range. Simple phase plane methods show that drift rates will slow down as firing rates exceed the linear range (Brown & Holmes, 2001), yielding longer integration times for a given level of accuracy. Attempting to capture these effects with a (constant-drift) DDM may result in the higher-than-optimal threshold predictions. The analysis required to investigate this is beyond the scope of this article.

At the same time, for a very difficult task and *short* D_{total} , the optimal threshold may be very low, corresponding to near random responding. However, participants may feel uncomfortable producing random responses. There may also be a computational disadvantage in keeping the threshold fixed near zero. If the information content of the input increases (e.g., the task becomes easier), this would go undetected (since no integration is occurring). Thus, higher level considerations (e.g., knowledge about non-stationarity of the environment) may also influence threshold setting, and drive participants to raise thresholds above zero (at least from time to time), providing another possible explanation for thresholds $z > z_0$.

It is also possible that if the task is very difficult, participants may explore strategies other than simple diffusion-type accumulation of information. For example, in Experiment 2 they may have counted the asterisks. Indeed, one participant reported adopting such a strategy in the questionnaire: "If it is hard to identify, I will roughly count each row." This participant produced very long RTs – up to 9.1s. Two participants reported a different change in strategy: "On the tougher ones [...] I started looking for empty spaces instead of stars." A

consideration of such alternate strategies is also beyond the range of this article, although it is certainly an important target for cognitive and neuroscientific research and indeed has received much attention in cognitive science (e.g., Newell & Simon, 1988). In any event, the engagement of such alternate strategies would of course weaken the relationship between observed performance and that predicted by the DDM. However, this observation does suggest an intriguing possibility: If the DDM, because of its intrinsic simplicity, and its optimality for simple decision making, is highly conserved, then it should provide a progressively more accurate account of decision making performance for organisms that have progressively simpler cognitive architectures, and therefore have less access to more complex strategies. Similarly, it should be possible to promote optimal performance in human participants by interfering with higher cognitive functions subserving complex decision making strategies.

Learning of Decision Thresholds

The observation that some participants were able to find decision thresholds close to the optimal values raises the question of how this was accomplished. It is most likely that they determine the threshold using an adaptive procedure: Different values of threshold are set, their effect on a criterion is observed, and the threshold is adjusted to optimize the criterion. Several models of such adaptation have been proposed (Erev, 1998; Myung & Busemeyer, 1989; Simen et al., 2006).

Myung and Busemeyer (1989) tested the predictions of threshold adaptation models in the experiment in which the participants were required to minimize a weighted difference between ER and RT. In their experiment the effects of learning were observed on a much longer time-scale than we observed: Their analysis suggested that the threshold was converging to the optimal value within 100-200 trials, while we observed much faster convergence (within ~20 trials). There are a number of factors that can contribute to this difference. First, Myung and Busemeyer plotted the precise threshold value on each trial (which was allowed by their experimental design), rather than noisy functions of thresholds, hence from this point of view their analysis was more precise. On the other hand, our study generated substantially more data: Myung and Busemeyer calculated the value of threshold on a given trial as the average over only 6 learning episodes for one condition, and 4 episodes for the other condition. By contrast in Figure 10a we calculated the difference in RT between delay conditions on a given trial as the average over 700 learning episodes (20 participants of Experiment 1 \times 5 blocks + 60 participants of Experiment 2 \times 10 blocks).

Furthermore, it can be speculated that we observed faster threshold convergence because RR is a more ecologically relevant criterion than the weighted difference between ER and RT, and because humans' threshold adjustment mechanisms are tuned to maximize RR. Recently Simen et al. (2006) proposed such a specialized mechanism that is able to rapidly find the threshold maximizing RR. Future work could compare the predictions of different models of threshold adaptation with experimental data.

Acknowledgments

This work was supported by EPSRC grant EP/C514416/1 and PHS grants MH58480 and MH62196 (Cognitive and Neural Mechanisms of Conflict and Control, Silvio M. Conte Center). We thank Joshua Gold for providing us with the program to generate the "moving dots" stimuli, and for discussions, Patrick Simen for reading the manuscript and very useful comments, Eric Shea-Brown and Jeff Moehlis for advice on the planning and analysis of the experimental data, and Miriam Zackenhause, Yakov Ben-Haim, and Paul Howard-Jones for discussion.

Appendix A. Optimal Performance Curves

Here we show how to obtain the relationship between ER and DT that holds for thresholds that optimize RR. First, note that Equations 6 can be solved for \tilde{a} and \tilde{z} (cf. Wagenmakers et al., 2007):

$$\tilde{a} = \frac{1 - 2ER}{2DT} \log\left(\frac{1 - ER}{ER}\right), \quad \text{and} \quad \tilde{z} = \frac{DT}{1 - 2ER}. \quad (17)$$

Substituting Equations 17 into 8 and rearranging terms, we obtain the optimal performance curve for RR (Bogacz et al., 2006):

$$\frac{DT}{D_{total}} = \left(\frac{1}{ER \ln \frac{1-ER}{ER}} + \frac{1}{1-2ER} \right)^{-1}. \quad (18)$$

The curves for RA and RR_m can be obtained in an analogous way and are given by Equations 19 and 20 respectively (Bogacz et al., 2006):

$$\frac{DT}{D_{total}} = \frac{E - 2q - \sqrt{E^2 - 4q(E+1)}}{2q}, \quad \text{where} \quad E = \frac{1}{ER \ln \frac{1-ER}{ER}} + \frac{1}{1-2ER}, \quad (19)$$

$$\frac{DT}{D_{total}} = (1+q) \left(\frac{\frac{1}{ER} - \frac{q}{1-ER}}{\ln \frac{1-ER}{ER}} + \frac{1-q}{1-2ER} \right)^{-1}. \quad (20)$$

Note that for the normalized decision time to be well defined in Equation 19 the weight q must satisfy:

$$q \leq \min \left\{ \frac{E^2}{4(E+1)} \right\} = \frac{E_{min}^2}{4(E_{min}+1)}. \quad (21)$$

Numerically, we find that the lowest value E can take is $E_{min} \approx 5.224$, implying that $q \leq 1.096$ (Zackenhause et al., submitted). Similarly, for the normalized decision time to be non-negative in Equation 20 the weight must satisfy $q \leq 1$.

Appendix B. Correlation Measure Assuming Equal Means

Pearson's correlation coefficient between two samples x_1, \dots, x_n , and y_1, \dots, y_n is calculated as:

$$r = \frac{Av((x - Av(x))(y - Av(y)))}{\sqrt{Av((x - Av(x))^2) Av((y - Av(y))^2)}} \quad (22)$$

where Av denotes the average. This coefficient assumes that the two random variables generating the samples have different means – note that in Equation 22 $Av(x)$ and $Av(y)$ are estimated separately. If we assume that the two variables have equal means, we can estimate the mean from both sets of samples $m = Av(x, y)$, and use it in the definition of the modified correlation coefficient:

$$r_1 = \frac{Av((x - m)(y - m))}{\sqrt{Av((x - m)^2)Av((y - m)^2)}} \quad (23)$$

The significance of a particular value R of r_1 for a given sample size n was found using the following Monte-Carlo simulation. In each of 10,000 iterations, n pairs of samples were generated from independent normal standard distributions, and r_1 was computed. The significance was taken as a fraction of iterations in which $|r_1| > |R|$ (corresponding to a two-tailed test).

References

- Akaike H. Likelihood of a model and information criteria. *Journal of Econometrics*. 1981; 16:3–14.
- Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks. *Psychological Review*. 2006; 113:700–765. [PubMed: 17014301]
- Bogacz R, Cohen JD. Parameterization of connectionist models. *Behavioral Research Methods, Instruments, and Computers*. 2004; 36:732–741.
- Brainard DH. The Psychophysics Toolbox. *Spatial Vision*. 1997; 10:433–436. [PubMed: 9176952]
- Brown E, Holmes P. Modeling a simple choice task: stochastic dynamics of mutually inhibitory neural groups. *Stochastics and Dynamics*. 2001; 1:159–191.
- Busemeyer JR, Rapoport A. Psychological models of deferred decision making. *Journal of Mathematical Psychology*. 1988; 32:91–134.
- Busemeyer JR, Townsend JT. Decision field theory: A dynamic-cognitive approach to decision making in uncertain environment. *Psychological Review*. 1993; 100:432–459. [PubMed: 8356185]
- Desai RA, Maciejewski PK, Pantalon MV, Potenza MN. Gender Differences in Adolescent Gambling. *Annals of Clinical Psychiatry*. 2005; 17:249–258. [PubMed: 16402759]
- Edwards W. Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology*. 1965; 2:312–329.
- Erev I. Signal detection by human observers: a cut-off reinforcement learning model of categorization decisions under uncertainty. *Psychological Review*. 1998; 105:280–298. [PubMed: 9669925]
- Feller, W. An introduction to probability theory and its applications. New York: Wiley; 1968.
- Franks NR, Dornhaus A, Fitzsimmons JP, Stevens M. Speed versus accuracy in collective decision making. *Proceedings of the Royal Society, B*. 2003; 270:2457–2463.
- Gneezy U, Rustichini A. Gender and Competition at a Young Age. *The American Economic Review*. 2004; 94:377–381.
- Gold JI, Shadlen MN. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*. 2002; 36(2):299–308. [PubMed: 12383783]
- Gold JI, Shadlen MN. The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *Journal of Neuroscience*. 2003; 23:632–651. [PubMed: 12533623]
- Holmes P, Shea-Brown E, Moehlis J, Bogacz R, Gao J, Aston-Jones G, et al. Optimal decisions: From neural spikes, through stochastic differential equations, to behavior. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*. 2005; 88:2496–2503.
- Laming, DRJ. Information theory of choice reaction time. New York: Wiley; 1968.
- Laming DRJ. Choice reaction performance following an error. *Acta Psychologica*. 1979; 43:199–224.

- Maddox WT. Toward a unified theory of decision criterion learning in perceptual categorization. *Journal of the Experimental Analysis of Behaviour*. 2002; 78:567–595.
- Maddox WT, Bohil CJ. Base-rate and payoff effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 1998; 24:1459–1482.
- Maritz JS, Jarrett RG. A Note on Estimating the Variance of the Sample Median. *Journal of the American Statistical Association*. 1978; 73:194–196.
- Markman AB, Baldwin GC, Maddox WT. The interaction of payoff structure and regulatory focus in classification. *Psychological Science*. 2005; 16:852–855. [PubMed: 16262768]
- Masson MEJ, Loftus GR. Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*. 2003; 57:203–220. [PubMed: 14596478]
- Mozer, MC.; Colagrosso, MD.; Huber, DH. A rational analysis of cognitive control is a speeded discrimination task. In: Dietterich, T.; Becker, S.; Ghahramani, Z., editors. *Advances in Neural Information Processing Systems*. Vol. 14. Cambridge MA: MIT Press; 2002. p. 51–57.
- Myung IJ, Busemeyer JR. Criterion learning in a deferred decision making task. *American Journal of Psychology*. 1989; 102:1–16.
- Newell, A.; Simon, HA. The theory of human problem solving. In: Collins, AM.; Smith, EE., editors. *Readings in cognitive science: A perspective from psychology and artificial intelligence*. San Mateo, CA: Morgan Kaufmann; 1988. p. 33–51.
- Pachella, RG. The interpretation of reaction time in information-processing research. In: Kantowitz, B., editor. *Human information processing: tutorial in performance and recognition*. Hillsdale, NJ: Erlbaum; 1974. p. 41–82.
- Palmer J, Huk AC, Shadlen MN. The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*. 2005; 5:376–404. [PubMed: 16097871]
- Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*. 1997; 10:437–442. [PubMed: 9176953]
- Rapoport A, Burkheimer GB. Models for deferred decision making. *Journal of Mathematical Psychology*. 1971; 8:508–538.
- Ratcliff R. A theory of memory retrieval. *Psychol Rev*. 1978; 83:59–108.
- Ratcliff R. Modeling response signal and response time data. *Cognitive Psychology*. 2006; 53:195–237. [PubMed: 16890214]
- Ratcliff R, Gomez R, McKoon G. A diffusion model account of the lexical decision task. *Psychological Review*. 2004; 111:159–182. [PubMed: 14756592]
- Ratcliff R, McKoon G. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput*. 2008; 20(4):873–922. [PubMed: 18085991]
- Ratcliff R, Rouder JN. A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance*. 2000; 26:127–140. [PubMed: 10696609]
- Ratcliff R, Smith PL. Comparison of sequential sampling models for two-choice reaction time. *Psychological Review*. 2004; 111(2):333–367. [PubMed: 15065913]
- Ratcliff R, Thapar A, McKoon G. A diffusion model analysis of the effects of aging on brightness discrimination. *Perception and Psychophysics*. 2003; 65:523–535. [PubMed: 12812276]
- Ratcliff R, Tuerlinckx F. Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*. 2002; 9:438–481. [PubMed: 12412886]
- Ratcliff R, Van Zandt T, McKoon G. Connectionist and diffusion models of reaction time. *Psychological Review*. 1999; 106:261–300. [PubMed: 10378014]
- Rowan, T. *Functional Stability Analysis of Numerical Algorithms*. University of Texas; Austin: 1990.
- Shadlen MN, Newsome WT. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol*. 2001; 86(4):1916–1936. [PubMed: 11600651]
- Simen PA, Buck C, Hu P, Holmes P, Cohen JD. Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. submitted.
- Simen PA, Cohen JD, Holmes P. Rapid decision threshold modulation by reward rate in a neural network. *Neural Networks*. 2006; 19:1013–1026. [PubMed: 16987636]

- Stone M. Models for choice reaction time. *Psychometrika*. 1960; 25:251–260.
- Usher M, McClelland JL. The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev*. 2001; 108(3):550–592. [PubMed: 11488378]
- Vandekerckhove J, Tuerlinckx F. Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*. 2007; 14:1011–1026. [PubMed: 18229471]
- Vickers D. Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*. 1970; 13:37–58. [PubMed: 5416868]
- Wagenmakers EJ, van der Maas HLJ, Grasman RPPP. An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*. 2007; 14:3–22. [PubMed: 17546727]
- Wald A, Wolfowitz J. Optimum character of the sequential probability ratio test. *Annals of Mathematical Statistics*. 1948; 19:326–339.
- Wickelgren WA. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*. 1977; 41:67–85.
- Zackenhause M, Holmes P, Bogacz R. Robust versus optimal strategies for two-alternative forced choice tasks. submitted.

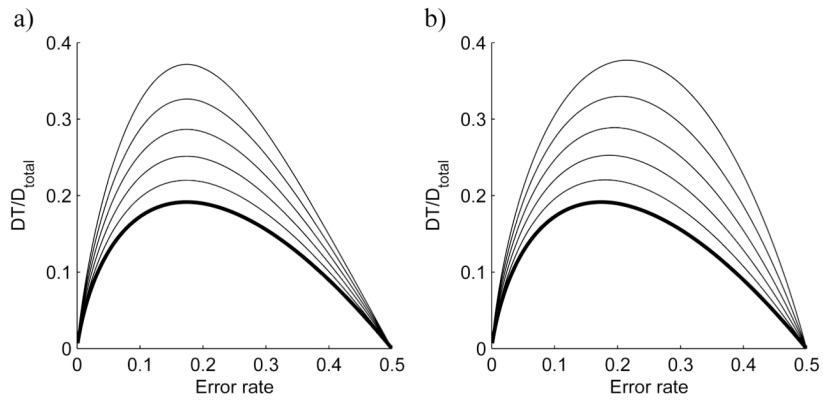


Figure 1. Optimal performance curves. Horizontal axes show the error rate and vertical axes show the decision time (DT) normalized by total delay D_{total} . The thick line (identical in both panels) is the optimal performance curve for the reward rate (RR). The thin lines show the generalized optimal performance curves for reward accuracy (panel a) and modified RR (panel b). Each curve corresponds to a different value of q ranging from 0 (when it reduces to the optimal performance curve for RR, shown in the thick lines) to 0.5 (top curves) in steps of 0.1.

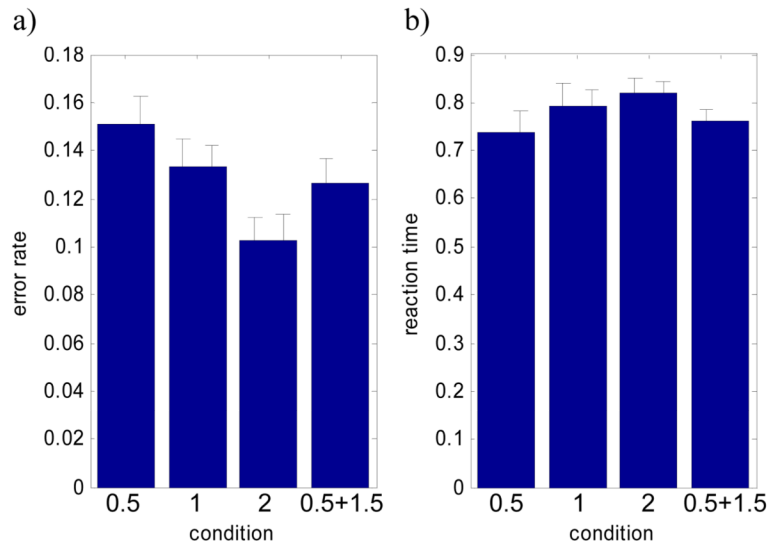


Figure 2.

Average error rate (ER) and reaction time (RT) in units of seconds for all twenty participants of Experiment 1 and all delay conditions. Delay conditions are indicated on x-axes – labels correspond to conditions (1) $D=0.5s$; (2) $D=1s$; (3) $D=2s$; and (4) $D=0.5s, D_p=1.5s$. Error bars consider comparisons between two adjacent conditions (inspired by Masson & Loftus, 2003). For example, in panel a, error bar for condition $D=0.5$ and the left error bar for condition $D=1$ consider comparison between these two conditions. The height of the error bars corresponds to standard error of the *differences* between ERs (or RTs in panel b) in the two conditions. Specifically, the difference is first calculated for each participant, and then the standard error of the differences is calculated across participants; both error bars are equal to this standard error, so they have equal heights. These error bars have standard interpretation: if two adjacent bars are different from each other by more than approximately two heights of corresponding error bars, then ERs (or RTs) are significantly different according to the paired t-test.

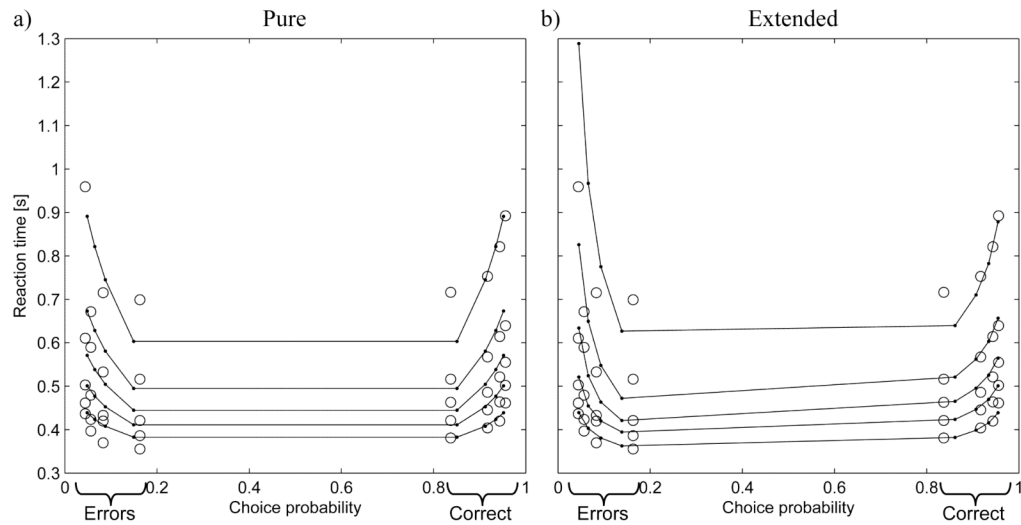


Figure 3.

Quantile probability plots showing the fits of the pure (a) and the extended (b) drift diffusion model (DDM) to behavioural data from a sample participant. Open circles correspond to the behavioural data. The horizontal axes indicate the probability of error in the left parts of the panels, and probability of correct choice in the right parts (see labels on x-axes). In each part, the four columns of circles correspond to the four delay conditions in the experiment. In each column of circles, the vertical positions of the five circles indicate the quantiles of reaction times. Small filled circles visualise the corresponding predictions for the DDM. These are connected by lines to make the patterns they create more visible. For clarity, error bars with confidence intervals for quantiles of reaction times are not shown here. They are plotted for the same participant in Figure 5 of Bogacz et al. (2006), which shows that the confidence intervals are very large for the error trials (up to 1.48s for the 0.9 quantile in $D=0.5$, $D_p=1.5$ condition) because of the small number of such trials. The estimated parameters of the models (for noise parameter fixed at $c=0.1$): a) pure DDM: $T_0=0.346$, $A=0.219$, $z_1=0.0398$, $z_2=0.0535$, $z_3=0.0610$, $z_4=0.0682$; b) extended DDM: $T_0=0.372$, $s_f=0.084$, $m_A=0.344$, $s_A=0.152$, $s_x=0.044$, $z_1=0.0503$, $z_2=0.0592$, $z_3=0.0687$, $z_4=0.0816$.

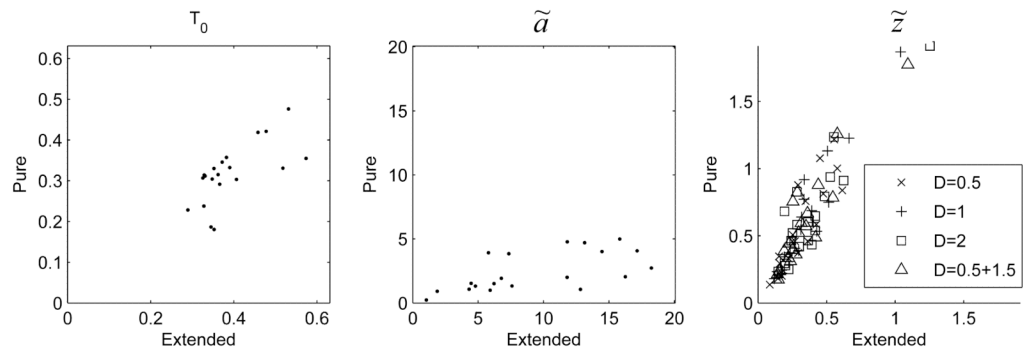


Figure 4.

Comparison of parameters T_0 , \tilde{a} and \tilde{z} estimated by fitting pure (vertical axes) and extended (horizontal axes) drift diffusion model. In the left and the central panels dots correspond to individual participants; in the right panel different experimental delay conditions are indicated by different symbols (see legend).

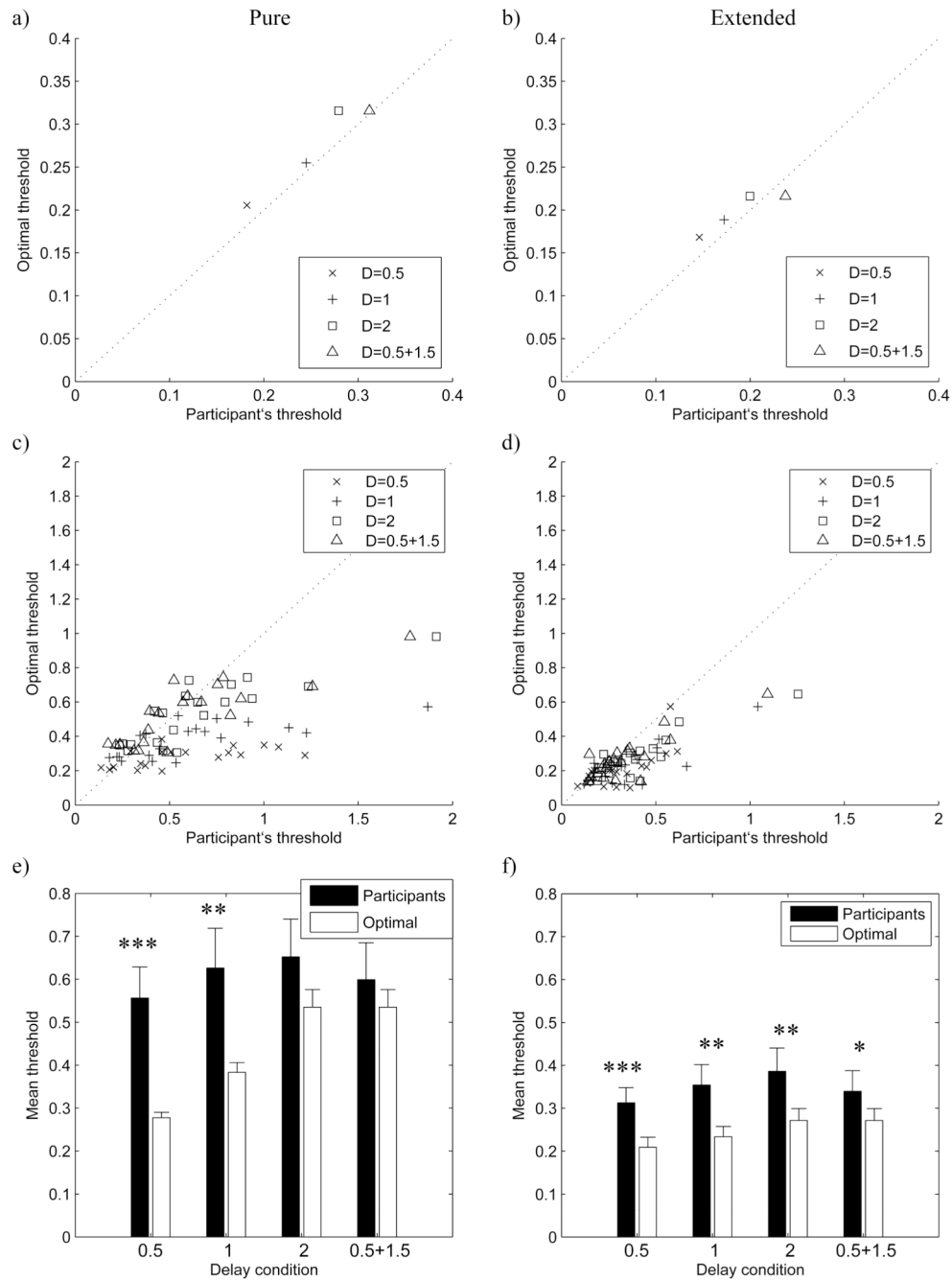


Figure 5. Comparison of participants' and optimal thresholds in Experiment 1. The left column of panels (a, c, e) compares estimates and predictions of the pure drift diffusion model (DDM), while the right column of panels (b, d, f) – of the extended DDM. In panels a-d, horizontal axes correspond to participants' normalized thresholds, the optimal normalized thresholds are along the vertical axes, and the dotted line is the identity line. Different experimental delay conditions are indicated by different symbols (see legend). Note different scales between rows of panels. Panels a and b show a sample participant with the best fit; panels c and d show data of all 20 participants. Panels e and f show the mean participants' (black bars) and optimal (white bars) thresholds averaged across participants for different delay

conditions (indicated on the horizontal axis). Error bars show standard error (there are error bars on bars corresponding to optimal thresholds, because different participants have different optimal thresholds). Stars indicate the level of significance of the difference between participants' and optimal thresholds (paired t-test): one star denotes $p < 0.05$, two stars denote $p < 0.01$, three stars denote $p < 0.001$.

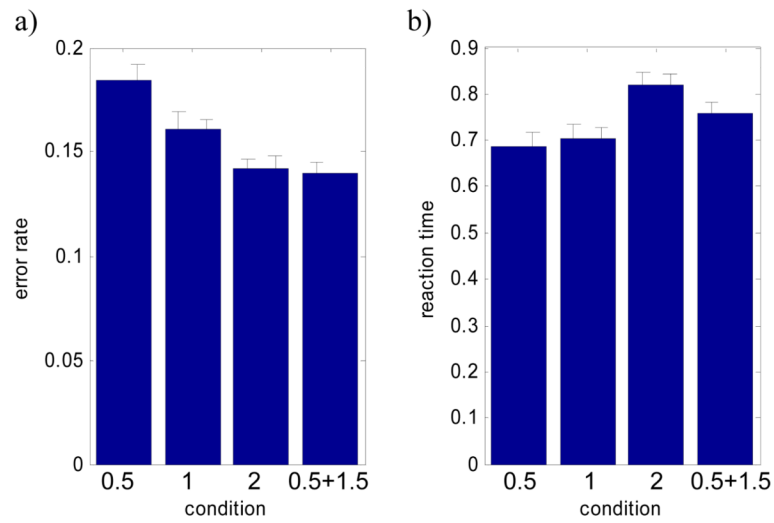


Figure 6. Mean error rate and mean reaction time (in seconds) for all 60 participants in Experiment 2 for each delay condition averaged across difficulty conditions. Error bars as in Figure 2.

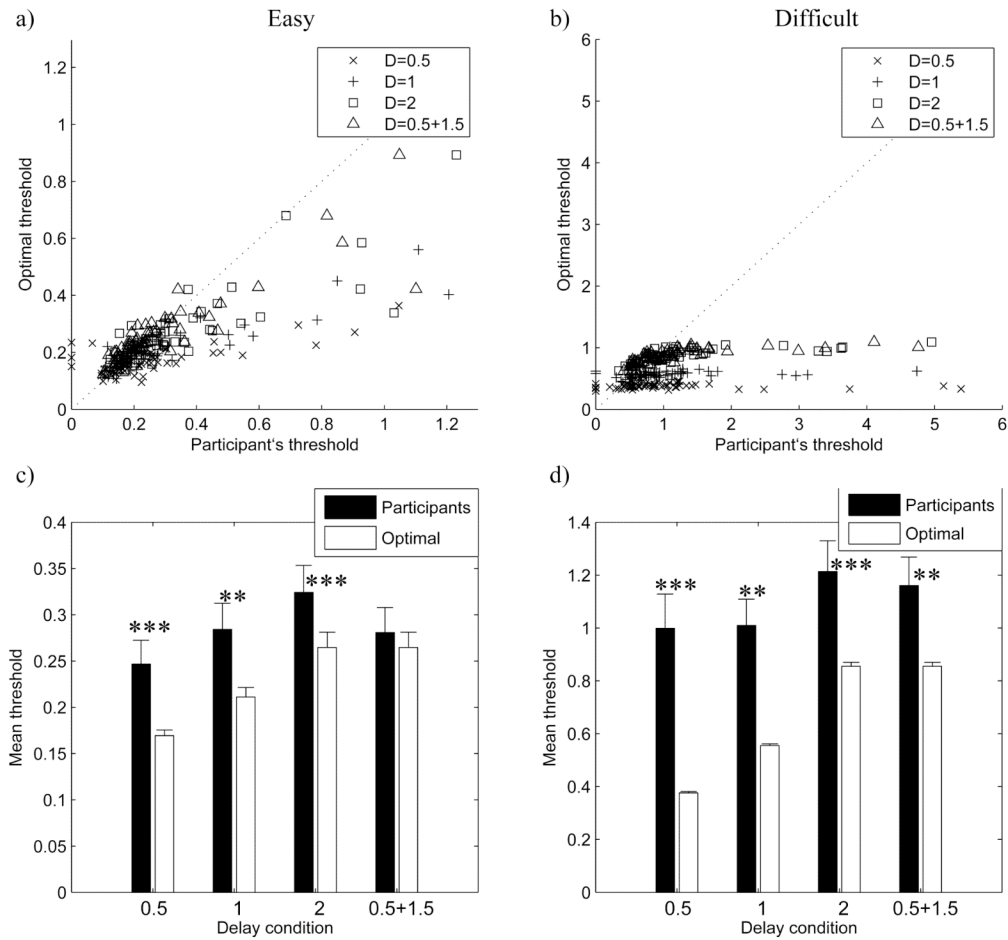


Figure 7. Comparison of participants' and optimal thresholds in Experiment 2. Left panels (a and c) show findings for the easy condition, right panels (b and d) for the difficult condition. In panels a and b, participants' normalized thresholds are shown on horizontal axes, optimal normalized thresholds on the vertical axes, and the dotted line is the identity line. Different experimental delay conditions are indicated by different shapes (see legend). Note different scales. Panels c and d show the mean participants' (black bars) and optimal (white bars) thresholds averaged across participants for different delay conditions (indicated on the horizontal axis). Error bars show standard error, and stars indicate the level of significance of the difference between participants' and optimal thresholds (paired t-test): one star denotes $p < 0.05$, two stars denote $p < 0.01$, three stars denote $p < 0.001$.

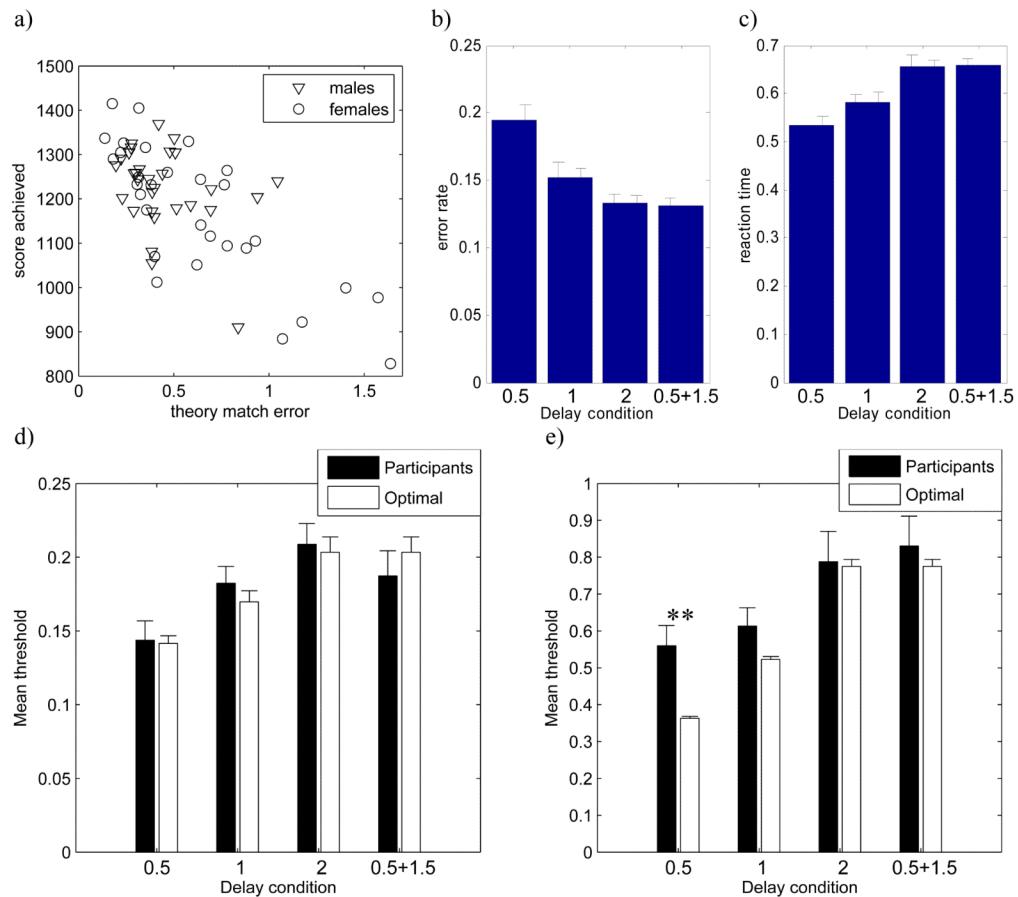


Figure 8.

Performance of participants with reward scores higher than the median reward score for Experiment 2. a) Relationship between theory match error (x-axis) and achieved reward score (y-axis). Each symbol corresponds to one participant; genders as indicated in key. b, c) Mean error rate and reaction time (in units of seconds) averaged across difficulty conditions. Error bars as in Figure 2. d, e) mean participants' (black bars) and optimal (white bars) normalized thresholds averaged across participants for different delay conditions (indicated on the horizontal axis). Panel d shows data for easy conditions; panel e – for difficult conditions. Error bars show standard error, and stars indicate the level of significance of the difference between participants' and optimal thresholds (paired t-test): one star denotes $p < 0.05$, two stars denote $p < 0.01$, three stars denote $p < 0.001$.

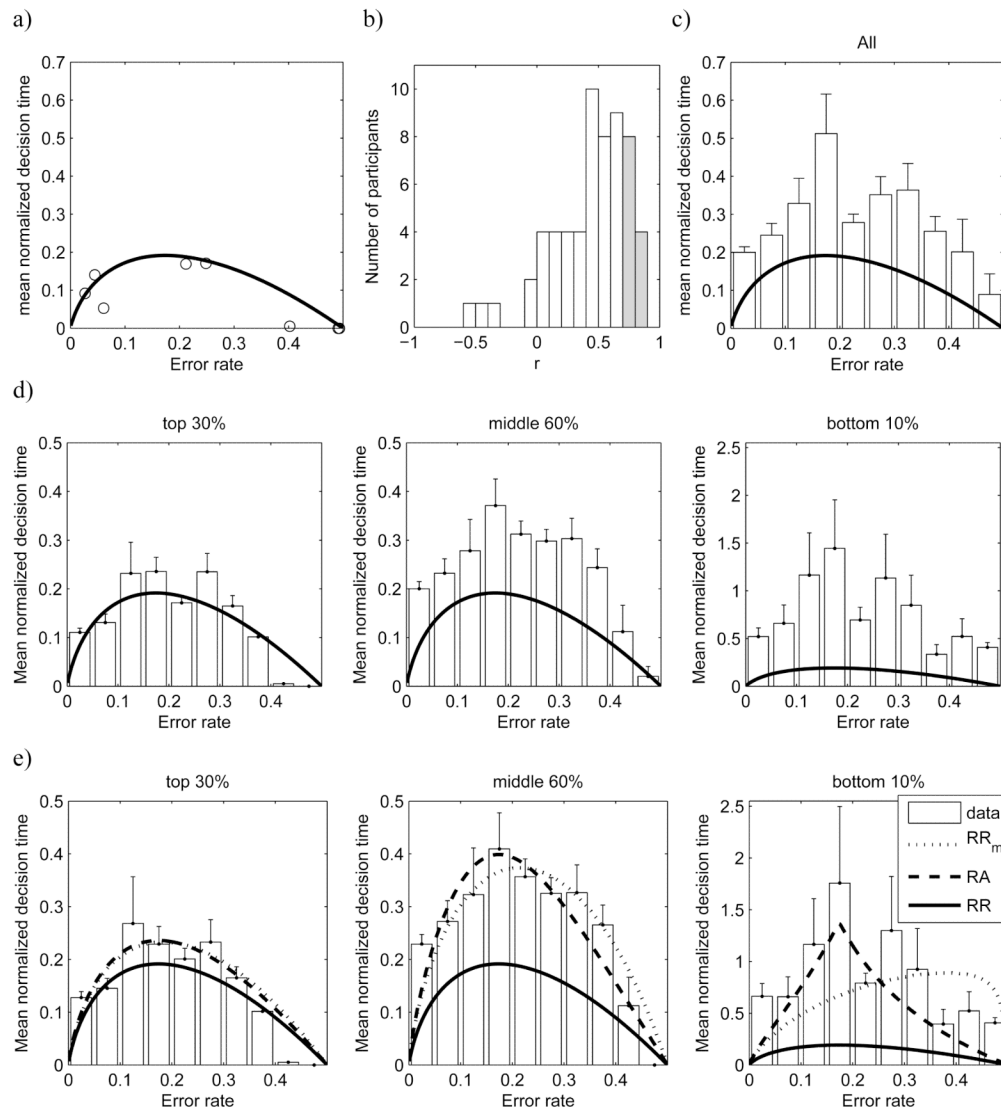


Figure 9.

Optimal performance curves. In all panels except b, horizontal axes show the error rate (ER) and vertical axes show the normalized decision time (DT), i.e. DT divided by total delay D_{total} . Note the differences in scales between panels. Solid black curves show the theoretical prediction based on the assumption that participants maximize the reward rate. a) Normalized DTs as a function for ER plotted for one participant of Experiment 2. Each circle corresponds to one experimental condition. For two conditions the normalized DTs and ERs were very close, so two circles overlap (in the bottom right corner of the panel). b) Histogram of the correlations between normalized DT and the normalized DT predicted by the optimal performance curve. Each correlation is calculated for one participant of Experiment 2. Shaded bars indicate the correlations that were statistically significant ($p < 0.05$). In panels c-e, the bars show the mean normalized DT averaged across participants' conditions with ERs falling into the given interval (from Experiments 1 and 2). The error bars indicate standard error. There are no error bars for some bins, as there was just one condition falling into this bin, or DTs for participants' conditions falling into this bin were all very close to 0. c) The bars are based on all participants. d) The participants are divided into three groups on the basis of their reward scores and bars in each of the three panels are

based on the data from the corresponding group of participants (see titles of the panels). e) Data from conditions with $D_p=0$ based on subgroups of participants as in panel d. Best fitting optimal performance curves for RA and RR_m are also shown (see legend).

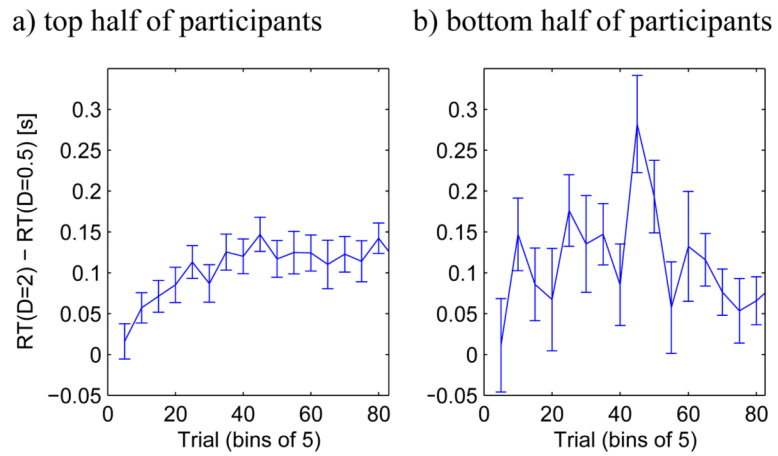


Figure 10.

Changes in participants' performance within single blocks of trials. In all panels horizontal axes show the trial number within a block – each point corresponds to a bin of 5 trials. Vertical axis shows the difference in the mean reaction time (RT) between delay condition $D = 2$, and $D = 0.5$ averaged over: 5 trials in a given bin and both difficulty conditions. Error bars indicate standard error. Panel a shows data averaged across participants of Experiments 1 and 2 with reward scores higher than the median for each of experiment, and panel b shows data averaged across participants with reward scores lower than the median.

Table 1

Symbols used in the optimal threshold theory.

Symbol	Description
A	Drift rate
\tilde{a}	Signal-to-noise ratio (squared)
c	Magnitude of noise
D	Delay between response and next stimulus
D_p	Additional penalty delay on error trials
DT	Decision Time
D_{total}	$D + D_p + T_0$
ER	Error Rate
q	Relative weight of accuracy
RA	Weighted difference between RR and ER
RR	Reward Rate
RR_m	Modified RR
RT	Reaction Time
s_A	Variability of A
s_t	Variability of T_0
s_x	Variability of $x(0)$
T_0	Part of RT connected with non-decision processes
$x(t)$	Integrated difference between evidence supporting two alternatives
z	Decision threshold
\tilde{z}	Normalized threshold
\tilde{z}_o	Optimal normalized threshold

Table 2

The estimated values of parameter q of the optimal performance curves for RA and RR_m for three groups of participants sorted by reward score.

Criterion	Top 30%	Middle 60%	Bottom 10%
RA	0.15	0.55	1.096
RR_m	0.14	0.49	0.98

Table 3

Ratios of the likelihoods of experimental data given the pairs of criteria listed in the left columns for groups of participants listed in top row. To account for the fact that the performance curves for RA and RR_m have one free parameter and the curve for RR does not, the bottom two rows include in brackets the ratio of the likelihood of the data given the two curves divided by $\exp(1)$, as suggested by Akaike (1981).

Criteria compared	Top 30%	Middle 60%	Bottom 10%	All
RA / RR_m	1.15	4.78	43.16	237.25
RA / RR	42.65 (15.69)	$> 10^{20}$	$> 10^5$	$> 10^{26}$
RR_m / RR	36.98 (13.60)	$> 10^{20}$	$> 10^4$	$> 10^{25}$

Table 4

Summary of the support for the predictions stated in the Introduction in the data from Experiments 1 and 2.

Prediction	All participants	Participants with highest RR
1. z increases with D	Satisfied	Satisfied
2. z depends on $D + D_p$	Not satisfied	Satisfied
3. $z = z_o$	Partially: z correlated with z_o but $z > z_o$	Satisfied
4. Optimal relationship between ER and DT	Partially: qualitative match but DTs higher than optimal	Satisfied