

Published in final edited form as:

*J Math Psychol.* 2010 April 1; 54(2): 230–246. doi:10.1016/j.jmp.2009.12.004.

## Robust versus optimal strategies for two-alternative forced choice tasks

M. Zacksenhouse<sup>a,d</sup>, R. Bogacz<sup>b</sup>, and P. Holmes<sup>c</sup>

<sup>a</sup>Faculty of Mechanical Engineering, Technion –Israel Institute of Technology, Haifa 32000, Israel

<sup>b</sup>Department of Computer Science, University of Bristol, Bristol BS8 1UB, U.K

<sup>c</sup>Program in Applied and Computational Mathematics and Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, U.S.A

### Abstract

It has been proposed that animals and humans might choose a speed-accuracy tradeoff that maximizes reward rate. For this utility function the simple drift-diffusion model of two-alternative forced-choice tasks predicts a parameter-free optimal performance curve that relates normalized decision times to error rates under varying task conditions. However, behavioral data indicate that only  $\approx 30\%$  of subjects achieve optimality, and here we investigate the possibility that, in allowing for uncertainties, subjects might exercise robust strategies instead of optimal ones. We consider two strategies in which robustness is achieved by relinquishing performance: maximin and robust-satisficing. The former supposes maximization of guaranteed performance under a presumed level of uncertainty; the latter assumes that subjects require a critical performance level and maximize the level of uncertainty under which it can be guaranteed. These strategies respectively yield performance curves parameterized by presumed uncertainty level and required performance. Maximin performance curves for uncertainties in response-to-stimulus interval match data for the lower-scoring 70% of subjects well, and are more likely to explain it than robust-satisficing or alternative optimal performance curves that emphasize accuracy. For uncertainties in signal-to-noise ratio, neither maximin nor robust-satisficing performance curves adequately describe the data. We discuss implications for decisions under uncertainties, and suggest further behavioral assays.

### Keywords

Two-alternative forced-choice; Decision making; Robust decision making; Optimal decision making; Info-gap; Uncertainties; Drift-diffusion models

## 1 Introduction

There is a substantial literature on random walk and drift-diffusion (DD) processes as models for behavioral measures of reaction time and error rates on two-alternative forced-choice (2AFC) tasks, e.g. Stone (1960); Laming (1968); Ratcliff (1978); Ratcliff et al. (1999); Smith and Ratcliff (2004). Additionally, *in vivo* recordings in monkeys trained to respond to motion stimuli show that neural spike rates in certain oculo-motor regions evolve like sample paths of a DD process, rising to a threshold prior to response initiation (Schall, 2001; Gold and Shadlen, 2001; Roitman and Shadlen, 2002; Ratcliff et al., 2003; Mazurek et al., 2003; Ratcliff et al., 2006). Such experiments support a picture in which the state of the

<sup>d</sup>Corresponding author: mermz@tx.technion.ac.il, Telephone: +972-4-8292092; Fax: +972-4-8295711.

DD process, interpreted as a difference between accumulating, noisy evidence streams, is integrated until it reaches a confidence threshold.

The simple DD process with constant signal-to-noise ratio (SNR) is a continuum limit of the sequential probability ratio test (SPRT), which is optimal for statistically stationary tasks in the sense that, on average, it renders a decision of specified accuracy for the smallest number of observations (Wald, 1947; Wald and Wolfowitz, 1948). It therefore offers a normative theory of decision making against which human and animal behaviors can be assessed (Gold and Shadlen, 2002). Specifically, both speed and accuracy are often at a premium, and a key issue is how these requirements are balanced in a *speed-accuracy tradeoff* (SAT). Wald and Wolfowitz (1948) used a weighted sum of decision time and error rate as an objective function in their analysis of SPRT, and Edwards (1965) generalized this to model how human subjects choose decision thresholds. This was further extended and tested in Rapoport and Burkheimer (1971) and Busemeyer and Rapoport (1988), and a related theory, involving a competition between reward and accuracy (COBRA), has also been proposed (Maddox and Bohil, 1998; Bohil and Maddox, 2003). A review of these theories appears in Bogacz et al. (2006), which also describes how the DD process with constant SNR emerges from various evidence accumulator models, as parameters approach particular limits.

Gold and Shadlen (2002) then proposed that the DD model be used to derive an explicit SAT that optimizes reward rate. Holmes et al. (2005) and Bogacz et al. (2006) performed this computation and carried out 2AFC experiments on human subjects to test its predictions. As discussed below, and at greater length in Bogacz et al. (2009), the experiments reveal that, while a significant subset of 80 subjects performed near-optimally, most were substantially suboptimal, favoring accuracy over speed.

Several explanations have been advanced to account for such behavior. Physiological constraints might prevent neural circuits from integrating evidence in the optimal manner prescribed by the SPRT (cf. Bogacz et al. (2006) and see discussions below). Secondly, attention may wax and wane, leading to variable SNR (also, experimental protocols often mix stimuli of different discriminability, violating the stationarity assumed in the SPRT and in the blocked experiments of Bogacz et al. (2009)). Thirdly, intentional emphases on accuracy or speed, as in Maddox and Bohil (1998) and Bohil and Maddox (2003), can modulate the SAT at the level of individual subjects. Finally, in addition to assuming statistical stationarity, the SPRT-based optimality theory of Gold and Shadlen (2002); Holmes et al. (2005) and Bogacz et al. (2006) implicitly assumes that key parameters such as experimenter-imposed response-to-stimulus delays and SNR are precisely known to the subject.

In this paper we shall briefly discuss all these factors before focusing on the final one. We suppose that only uncertain estimates are available for key parameters and apply *robust strategies* to predict robust, rather than optimal, SATs. We find that uncertainties in inter-trial delays can account for suboptimality in the data of Bogacz et al. (2009), while uncertainties in the SNR cannot. Furthermore, we show that a robust strategy, which also involves only one free parameter, provides a better fit than a version of the COBRA theory that includes a weight for accuracy.

The paper is structured as follows. In §2 we review the DD model and derivation of the optimal SAT, describing how it can be expressed as an optimal performance curve (OPC). Robust strategies for 2AFC, which account for parameter uncertainties, are developed in §3, and used to derive the robust performance curves. In §4 we compare the predictions of the robust approaches with experimental data of Bogacz et al. (2006) and with alternative

optimization strategies, and we briefly assess the effects of parameter mis-estimates. §5 contains a discussion and proposals for future experiments to further explore our conclusions. Many proofs and other mathematical and statistical details are relegated to Appendices. Common abbreviations used in the paper are summarized in Table 1.

## 2 A drift-diffusion model and optimal performance curves

The simplest version of a DD process is described by the following stochastic differential equation:

$$dx = A dt + \sigma dW; \quad x(0) = 0, \quad (1)$$

where  $A$  denotes the drift rate and  $\sigma dW$  increments drawn from a Wiener process with standard deviation  $\sigma$  (Gardiner, 1985), also known as the Wiener process with drift (Diederich and Busemeyer, 2003). In the 2AFC context, the state variable  $x(t)$  represents a running estimate of the logarithmic likelihood ratio (Gold and Shadlen, 2002), the two stimuli produce drift rates  $\pm A$  respectively, and on each trial a choice is made when  $x(t)$  first crosses either of the predetermined thresholds  $\pm x_{th}$ . It is implicitly assumed here that stimuli are presented with equal probabilities, that  $x(0)$  is initialized midway between the thresholds at the start of each trial as in the optimal SPRT, and that drift, noise level and thresholds remain constant over each block of trials. We shall refer to Eq. (1) as a *pure* DD process, to distinguish it from the *extended* processes described in §2.1.

Average performance on a block of trials is characterized by the probability of making an error,  $p(\text{err})$ , and the *mean decision time*,  $\langle DT \rangle$  (the first passage time for (1)), which can be computed explicitly as

$$p(\text{err}) = \frac{1}{1 + \exp(2\eta\theta)} \quad \text{and} \quad \langle DT \rangle = \theta \left[ \frac{\exp(2\eta\theta) - 1}{\exp(2\eta\theta) + 1} \right], \quad (2)$$

(Gardiner, 1985; Busemeyer and Townsend, 1993), cf. (Bogacz et al., 2006, Appendix). Here the parameters  $\eta \equiv (A/\sigma)^2$  and  $\theta \equiv |x_{th}/A|$  are the SNR (having the units of inverse time), and the threshold-to-drift ratio: i.e., the passage time for the noise-free process  $x(t) = At$ . (The present notation differs from Bogacz et al. (2006) to avoid confusion with that used below to describe uncertainty.) In (2)  $\langle DT \rangle$  is that part of the *mean reaction time*  $\langle RT \rangle$  ascribed to mental processing *per se*, excluding the non-decision latency  $T_0$  required for sensory transduction and motor response, i.e.,  $\langle DT \rangle = \langle RT \rangle - T_0$ .

The formulae of Eq. (2) may be inverted to solve for the DD model parameters,  $\eta$  and  $\theta$ , in terms of the two performance parameters,  $p(\text{err})$  and  $\langle DT \rangle$ :

$$\theta = \frac{\langle DT \rangle}{1 - 2p(\text{err})} \quad \text{and} \quad \eta = \frac{1 - 2p(\text{err})}{2\langle DT \rangle} \log \left( \frac{1 - p(\text{err})}{p(\text{err})} \right). \quad (3)$$

Given a suitable objective function for performance, these explicit formulae allow us to predict a parameter-free optimal performance curve that relates  $p(\text{err})$  and  $\langle DT \rangle$ . Before describing this, we address some shortcomings of the simple DD model of Eq. (1), starting with physiological factors.

As noted in §1, the DD model gains support from neurophysiology. A decision model in wide use employs a pair of leaky, mutually-inhibiting accumulators whose states represent averaged firing rates of populations of neurons preferentially sensitive to different stimuli

(Usher and McClelland, 2001). The DD variable  $x(t)$  is the difference between these states. However, reduction to a single variable is valid only if leak and inhibition are high, and to a DD process only if they are balanced (Bogacz et al., 2006) and firing rates remain in ranges in which linearization is justified (Cohen et al., 1990). Recent studies of multi-unit spiking neuron models (Wang, 2002; Wong and Wang, 2006) question these assumptions and suggest that nonlinear models may be necessary (Roxin and Ledberg, 2008). In view of the SPRT results noted above (Wald, 1947; Wald and Wolfowitz, 1948), all such generalizations, and the extended DD models described next, are necessarily non-optimal decision makers.

## 2.1 Extensions to the DD model

In addition to the neurobiologically-motivated nonlinear accumulator models mentioned above, extended DD processes have been introduced to better account for human and animal behaviors. In particular, mean reaction times for error and correct trials typically differ, slow errors being frequently observed, whereas the pure DD process (1) predicts identical forms for the probability distributions of DTs for both trial types.

Slow errors can be produced by selecting the drift rate  $A$  for each trial from a Gaussian distribution, and fast errors result by selecting initial conditions  $x(0)$  on each trial from a uniform distribution (Ratcliff et al., 1999). Including a possible overall bias, so that  $\langle x(0) \rangle \neq 0$ , these increase the number of DD parameters from two to five. To further allow for variability, the nondecision time  $T_0$  can also be taken from a distribution (Ratcliff and Tuerlinckx, 2002). Alternatively, time-dependent drift rates  $A(t)$ , which can represent varying attention, can produce slow or fast errors (Ditterich, 2006a,b; Liu et al., 2008; Eckhoff et al., 2008), in principle expanding the model to an infinite-dimensional parameter space.

Variable drift rates (either within or between trials) seem most relevant to account for differences in the reaction time distributions of error and correct decisions. However, as noted at the end of §2.2, the resulting optimal performance curves fail to match the observed preference for accuracy over speed in the present data (Bogacz et al., 2006). Here we develop a normative theory of robust decision making, based on the pure DD model, which accounts well for this observed preference. We introduce a new factor: *the degree of uncertainty* in estimating specific model parameters. Comparing the effects on different parameters, we are able to assess which source of uncertainty may best explain the observed behavior, and to contrast it with the emphasis placed on accuracy in theories such as COBRA (Maddox and Bohil, 1998; Bohil and Maddox, 2003). Furthermore, we show that uncertainties in SNR also allow variable drift rates, and thus can account for differences in DT distributions for error and correct decisions.

## 2.2 Optimal performance curves

To determine optimal strategies for a given experimental paradigm, we must specify an objective or utility function. Here we consider the paradigm of Gold and Shadlen (2002), in which a subject is presented with a stimulus and is free to indicate a choice (by motor or oculomotor response) at any time. Correct responses are rewarded and after each response there is a fixed delay or response-to-stimulus interval (RSI), denoted  $D_{RSI}$ , before the next trial begins. This may be increased by a penalty delay  $D_{pen} \geq 0$  in the event of an error. Trials are run sequentially in blocks in which delays and stimulus discriminabilities are constant. The duration of each block, rather than the number of trials within it, is fixed, thereby presenting subjects with a clear challenge of balancing accuracy with speed (the number of trials completed in the block).

Gold and Shadlen (2002) propose that subjects seek to maximize their *average reward rates* ( $RR$ ), defined as the fraction of correct responses divided by the average time between responses:

$$RR(p(\text{err}), \langle DT \rangle; D_{RSI}, T_0, D_{pen}) = \frac{1 - p(\text{err})}{\langle DT \rangle + T_0 + D_{RSI} + p(\text{err})D_{pen}}. \quad (4)$$

Appendix A.1 shows that the mean number of correct decisions per unit time is given asymptotically by Eq. (4). For simplicity we have assumed that the non-decision time  $T_0$  and the response-to-stimulus interval  $D_{RSI}$  remain fixed during each block of trials over which the reward rate is to be maximized, but if these are variable one would simply replace them with their means  $\langle T_0 \rangle$  and  $\langle D_{RSI} \rangle$ . Since only the combination  $T_0 + D_{RSI}$  appears in (4), we define the non-decision delay  $D = T_0 + D_{RSI}$  and for future use, the total delay  $D_{tot} = D + D_{pen}$ .

Substituting Eqs. (2) into Eq. (4) yields the reward rate that can be achieved as a function of the decision parameter  $\theta$  and the task parameters  $\eta$ ,  $D$  and  $D_{tot}$ :

$$RR(\theta; \eta, D, D_{tot}) = [\theta + D + (D_{tot} - \theta) \exp(-2\eta\theta)]^{-1}. \quad (5)$$

The response-to-stimulus interval,  $D_{RSI}$ , penalty delay,  $D_{pen}$ , and stimulus discriminability are under the experimenter's control and remain fixed throughout each block of trials. Stimulus discriminability determines the SNR,  $\eta$ , at least partially. To the extent that subjects can manipulate  $\eta$ , e.g., by increasing attention and thus decreasing the noise-variance  $\sigma^2$ , Eq. (5) implies that it should be maximized to achieve optimal  $RR$ . We therefore assume that  $\eta$  is fixed in each block (this will be relaxed when addressing uncertainties in §3). Under these conditions, the threshold-to-drift ratio,  $\theta$ , is the only adjustable parameter, so local maxima of  $RR$  occur at zeros of  $RR' / \theta$ , and the optimal threshold  $\theta_{op}$  satisfies:

$$\exp(2\eta\theta) - 1 = 2\eta(D_{tot} - \theta). \quad (6)$$

The left hand side of (6) is monotonically increasing with  $\theta$  and its right hand side is monotonically decreasing, so it has a unique solution  $\theta_{op}$  that depends only upon  $\eta$  and  $D_{tot}$ . This unique solution determines the globally optimal normalized threshold, leaving the optimal decision maker with no free decision parameter. The corresponding optimal reward rate is given by  $RR_{op} = RR(\theta_{op}; \eta, D, D_{tot})$ , and depends also on  $D$ .

Eq. (6) establishes a relationship among  $\theta$ ,  $\eta$  and  $D_{tot}$  that must hold if  $RR$  is maximized. Substituting Eqs. (3), the DD model parameters  $\theta$  and  $\eta$  can be replaced by the performance parameters  $\langle DT \rangle$  and  $p(\text{err})$  to obtain:

$$\frac{\langle DT \rangle}{D_{tot}} = \left[ \frac{1}{p(\text{err}) \log \left( \frac{1 - p(\text{err})}{p(\text{err})} \right)} + \frac{1}{1 - 2p(\text{err})} \right]^{-1}. \quad (7)$$

This *optimal performance curve* (OPC), pictured in Fig. 1 and first derived in Holmes et al. (2005) and Bogacz et al. (2006), is a unique, parameter-free relationship that describes the SAT: the tradeoff between speed ( $\langle DT \rangle / D_{tot}$ ) and accuracy ( $1 - p(\text{err})$ ) that must hold in order to maximize  $RR$  for the given task conditions. Each condition, specified by SNR  $\eta$  and

total delay  $D_{tot}$  determines a unique optimal threshold-to-drift ratio  $\theta_{op}$  given by Eq. (6), and hence a SAT that maximizes the  $RR$  for that condition. As task conditions vary – by manipulating  $\eta$  via stimulus discriminability, or  $D_{tot}$  via the RSI –  $\theta_{op}$  changes, and with it, the SAT that maximizes  $RR$ . To illustrate this, in Fig. 1(a) we mark eight points, corresponding to different SNR and  $D_{tot}$  values, on the OPC. Each of these conditions yields a different  $RR$ , as enumerated in the figure caption, *but each  $RR$  is optimal for that condition; indeed, all  $(p(\text{err}), \langle DT \rangle / D_{tot})$  pairs on the OPC correspond to optimal performances.*

The form of the OPC may be intuitively understood by observing that very noisy stimuli ( $\eta \approx 0$ ) contain little information, so that given a priori knowledge that they are equally likely, it is optimal to choose at random without examining them, giving  $p(\text{err}) = 0.5$  and  $\langle DT \rangle = 0$  (note the points for  $\eta = 0.1$  in Fig. 1(a)). At the opposite limit  $\eta \rightarrow \infty$ , noise-free stimuli are so easy to discriminate that both  $\langle DT \rangle$  and  $p(\text{err})$  approach zero (note the points for  $\eta = 100$  in Fig. 1(a)). This limit yields the highest  $RR$ s, but when SNR is finite, due to poor stimulus discriminability or a subject's inability to clearly detect a signal, making immediate decisions is not optimal. Instead, it is advantageous to accumulate the noisy evidence for just long enough to make the best possible choice (see the points for  $\eta = 1$  and  $\eta = 10$  in Fig. 1(a)).

Thresholds that differ from the optimal threshold cause sub-optimal performance, as illustrated by the diamonds in Fig. 1(a), which result from setting  $\theta$  25% above or below  $\theta_{op}$  for  $\eta = 1$  and  $D_{tot} = 2$ . In both cases the  $RR$  degrades by  $\approx 1.3\%$ .

To test whether humans can optimize their performance, two experiments were carried out using different 2AFC tasks, as detailed in Bogacz et al. (2006, 2009). In the first, 20 subjects viewed motion stimuli (Britten et al., 1993) and received 1 cent for each correct discrimination. The experiment was divided into 7-minute blocks with different penalty delays and response-to-stimulus intervals: three with  $D_{pen} = 0$  and  $D_{RSI} = 0.5\text{s}$ ,  $1.0\text{s}$ , and  $2.0\text{s}$ , respectively, and one with  $D_{pen} = 1.5\text{s}$  and  $D_{RSI} = 0.5\text{s}$ . One block of trials was run for each condition except for  $D_{pen} = 0$ ,  $D_{RSI} = 2.0\text{s}$ , for which two blocks were required to gather sufficient data. In the second experiment, 60 subjects discriminated if the majority of 100 locations on a visual display were filled with stars or empty (Ratcliff et al., 1999). The design was the same as that of the first experiment except that blocks lasted for 4 minutes, and the set of 5 blocks was repeated under two difficulty conditions. In both cases subjects were instructed to maximize their total earnings, and unrewarded practice blocks were administered before the tests began.

Since the OPC is independent of the parameters defining the DD process, and  $D_{tot}$  enters only as the denominator of the normalized decision time, data from different subjects and blocks of trials, with various SNRs and delays, can be combined in a single histogram for comparison with the OPC. For each subject and condition  $p(\text{err})$ s were computed and mean decision times  $\langle DT \rangle$  estimated by fitting the DD model to reaction time distributions, as described in Bogacz et al. (2006). The range of error rates from 0 – 50% was then divided into 10 bins and a mean normalized decision time  $\langle DT \rangle / D_{tot}$  **PH: This is correct:  $D_{tot}$  is constant on each block.** computed for each bin by averaging over all subjects and task conditions with error rates in that bin, yielding the white bars in Fig. 1 (Holmes et al., 2005; Bogacz et al., 2009). The shaded bars show results of the same analysis restricted to subgroups of subjects ranked according to their total rewards accrued over all blocks of trials and conditions. This reveals that the top 30% of subjects perform near the optimal SAT (the darkest bars lie close to the OPC), but that those with lower total scores exhibit significantly longer mean normalized decision times.



As noted in §1, this suboptimal behavior may reflect valuation of accuracy over speed, but it may also happen that subjects try to optimize their performances and fail to do so due to erroneous estimates of delays or SNR. (Since optimal thresholds depend on  $D_{tot}$  and  $\eta$  (Eq. (6)), errors in estimating them would lead to sub-optimality.) In this case, however, both under- and over-estimation would be expected to occur, yielding individual performances both below and above the OPC (see diamonds on Fig. 1(a)), in contrast with the averaged performance above the OPC in Fig. 1(b). We shall revisit this point at the end of §4, after showing in §3 that the data is captured well by assuming that subjects apply robust rather than optimal strategies and explicitly account for uncertainties in delays.

We believe that the conditions under which this data was collected, with stimulus discriminability and RSIs fixed during each block of trials, along with training sessions and emphasis on maximizing rewards, strongly encouraged subjects to improve their ability to extract signal from noise, thereby driving their decision mechanisms closer to that described by the optimal DD process. Indeed, while fits to an extended DD process with variability in  $A$  and  $x(0)$  improved on pure DD fits, detailed comparison, summarized in Appendix D, reveals that both models capture a similar proportion of the variance in subjects' thresholds. Furthermore, allowing variability in  $A$  and  $x(0)$  results in OPCs that differ qualitatively from the experimental data (Bogacz et al., 2006, Fig. 14). Specifically, trial-to-trial variability in  $A$  decreases the mean normalized decision time for a given error rate, favoring speed over accuracy, and variability in  $x(0)$  results in mean normalized decision times higher than those observed at low error rates. We therefore believe that it is reasonable to use pure DD fits in comparing the data from all 80 subjects with the OPC, and with the alternative performance curves described below.

### 2.3 Performance curves weighted for accuracy

The tendency to value accuracy over speed has been noted earlier (Myung and Busemeyer, 1989; Maddox and Bohil, 1998; Bohil and Maddox, 2003). This motivates the investigation of alternative objective functions (Holmes et al., 2005; Bogacz et al., 2006). We give two examples.

The first is a modified *reward-accuracy rate* that additionally penalizes errors, as suggested by the COBRA theory (Maddox and Bohil, 1998; Bohil and Maddox, 2003):

$$RA = RR - \frac{q p(\text{err})}{D_{tot}}. \quad (8)$$

Here  $q$  specifies the relative weight placed on accuracy, and the characteristic time  $D_{tot}$  is included so that the units of both terms are  $(\text{time})^{-1}$ . Secondly, for monetary rewards one can suppose that errors are penalized by subtraction from previous winnings, leading to the modified reward rate:

$$RR_m = \frac{(1 - p(\text{err})) - q p(\text{err})}{\langle DT \rangle + D_{tot}}. \quad (9)$$

(Here we require  $q < 1$ , so that rewards are positive.) The functions defining  $RA$  and  $RR_m$  each involve a single parameter  $q$ , and since errors are explicitly penalized in the expressions (8) and (9), no additional delay  $D_{pen}$  is included in  $D_{tot}$  here. Both reduce to  $RR$  for  $q = 0$ .

In Appendix A.2 we derive one-parameter families of OPCs for these objective functions and briefly characterize them (Eqs. (A.3–A.4) and (A.6)). We show that mean normalized decision times are well-defined for  $RA$  provided that  $q > 1.096$ , and that they all peak at  $p(\text{err}) \approx 0.1741$ , as for  $RR$ . In contrast, the OPCs of  $RR_m$  exhibit peaks that move rightward with increasing  $q$ . These modified OPCs will be assessed against the data in §4, along with predictions from robust decision making (see Fig. 6).

In the next section we develop decision strategies for the 2AFC task that are robust against uncertainties in either delays (§3.1) or SNR (§3.2). We show that they differ from the optimal strategies described above, and in §4 we argue that the data is more consistent with the use of robust decision strategies against uncertainties in delays than for any of the optimal procedures, especially for 70% of the subjects with lower total scores. In common with optimal ones, robust strategies require threshold setting, or, equivalently, the formulation of a stopping criterion for the decision procedure. We do not consider these issues here, but see Myung and Busemeyer (1989) and Busemeyer and Myung (1992) for models of criterion learning, and see Simen et al. (2006) for a neural network implementation of threshold setting in the DD model context.

### 3 Robust decisions under uncertainties

Optimal decision theory presumes that subjects maximize a relevant utility function, given the actual task parameters. However, these parameters are rarely known with accuracy and decisions must be made under uncertainties. In 2AFC tasks, as shown in §2.2, the  $RR$  depends on inter-trial delays and the SNR: quantities that may be difficult to estimate.

There are two major methodologies for handling parameter uncertainty, based on whether parameters are assumed to be stochastic or not (Ben-Tal and Nemirovski, 2008). The stochastic approach presumes that the uncertain parameters are characterized by a joint probability distribution. Given the distribution, the optimal strategy may be applied to select the decision that maximizes the expected utility. More generally, when only the class of possible distributions - but not the exact distribution - is known, the maximin strategy may be applied to select the decision that maximizes the minimum expected utility across all possible distributions (Wald, 1950). In either case, the stochastic approach requires additional information about the distribution function or the class of possible distributions.

Here we pursue a non-stochastic approach, which presumes that the parameters are deterministic but unknown (Ben-Tal and Nemirovski, 2008; Ben-Haim, 2006). First, we consider the case of bounded uncertainty, in which parameters are assumed to be bounded within a specific uncertainty set, and apply the maximin strategy to select the decision that maximizes the minimum utility (§3.1). We consider both uncertainties in delays (§3.1.1) and in SNR (§3.1.2) and show that the basic pattern of maximin performance curves for the former agrees well with the experimental data, while that for uncertainties in SNR does not.

We then consider the case in which only the family of uncertainty sets, rather than a specific one, is known. In this case, robust decisions are based on the notion of satisficing, i.e., satisfying an aspiration level of performance that is deemed sufficient. Simon (1956, 1982), who coined this term, suggested that the tendency to satisfice appears in many cognitive tasks including game playing, problem solving, and financial decisions, in which people typically do not or cannot seek optimal solution (Reber, 1995). Given an aspiration level of performance, the robust-satisficing strategy (Ben-Haim, 2006) selects the decision that achieves that performance under the largest uncertainty set as detailed in §3.2. The resulting performance curves are slightly inferior to the maximin performance curves in fitting the experimental data. We therefore outline this strategy only briefly, and discuss its potential relevance for other 2AFC experimental paradigms.



Robust decisions under parameter uncertainties can also account for trial-to-trial variations in the parameter. Trial-to-trial variations in drift rate are of special interest, given their role in explaining differences between correct and error response times (§2.2). Bounded trial-to-trial variations in drift rate are considered in §3.1.2, where, in terms of the maximin strategy, they are shown to be a special case of uncertainties in the SNR. As mentioned above, robust decisions under uncertainties in the SNR do not explain the experimental data well, so while trial-to-trial variations in drift rate may occur, they cannot alone explain the observed sub-optimal performance.

### 3.1 Maximin Performance Curves

#### 3.1.1 Uncertainties in delays

**Uncertainty set:** Uncertainties in delays may arise from poor estimation. Extensive investigations of interval timing, reviewed in Buhusi and Meck (2005), indicate that estimated intervals are normally distributed around the true duration with a standard deviation proportional to it. This is known as the scalar property of interval timing (Gibbon, 1977); it suggests that the size of the uncertainty set for the true interval increases with its duration.

According to Eq. (5), the  $RR$  is affected by both the non-decision delay  $D = D_{RSI} + T_0$  (which combines the response-to-stimulus interval and the non-decision latency) and the total delay  $D_{tot} = D + D_{pen}$  (which includes also the penalty delay), so we consider uncertainties in both of these delays. Their estimated values are referred to as the *nominal delays*, and denoted by  $\tilde{D}$  and  $\tilde{D}_{tot}$ . Accounting for the scalar property of interval timing, subjects are assumed to base their decisions on the assumption that the actual values are within the *presumed uncertainty set* given by:

$$U_p(\alpha_p; \tilde{D}, \tilde{D}_{tot}) = \{D, D_{tot} > 0: |D - \tilde{D}| \leq \alpha_p \tilde{D}, |D_{tot} - \tilde{D}_{tot}| \leq \alpha_p \tilde{D}_{tot}\}, \quad (10)$$

where the size of the uncertainty set is proportional to the nominal delay with a proportionality constant  $\alpha_p$ . We refer to  $\alpha_p$  as the *presumed level of uncertainty*, and note that it reflects the subject's assessment of how well he or she may estimate interval duration. In this section we assume that  $\alpha_p$  is known to the subject based on past experience with interval timing.

**Maximin decision strategy:** The optimal strategy of §2.2 maximizes the  $RR$  that can be achieved with the nominal delays, but ignores potential degradation in performance due to unfavorable delays. Instead, the maximin strategy focuses on the worst  $RR$  that could be obtained given the presumed level of uncertainty  $\alpha_p$ , and selects the maximin threshold  $\theta_{MM}$  that maximizes the worst  $RR$ , as defined next.

**Definition 3.1:** The maximin threshold  $\theta_{MM}$  under uncertainties in the delays is the threshold that maximizes the worst  $RR$  under the presumed uncertainty set  $U_p(\alpha_p; \tilde{D}, \tilde{D}_{tot})$  for the delays, given the presumed level of uncertainty,  $\alpha_p$ , the nominal delays,  $\tilde{D}$  and  $\tilde{D}_{tot}$ , and the SNR,  $\eta$ :

$$\theta_{MM}(\alpha_p; \eta, \tilde{D}, \tilde{D}_{tot}) = \arg \max_{\theta} \left( \min_{D, D_{tot} \in U_p(\alpha_p; \tilde{D}, \tilde{D}_{tot})} RR(\theta; \eta, D, D_{tot}) \right). \quad (11)$$

**Maximin performance curves:** The inner minimization in Eq. (11) specifies the lowest  $RR$  possible using the threshold  $\theta$ , given that delays lie within the presumed uncertainty set. This occurs when delays are longest, as detailed in Appendix B.1. Viewed as a function of

threshold  $\theta$ , the worst  $RR$  has a single maximum, which specifies the unique maximin threshold  $\theta_{MM}$ , as described in Appendix B.2. The resulting condition, Eq. (B.3), dictates a unique SAT that must hold to maximize the worst  $RR$ , as stated in the first Theorem.

**Theorem 3.1:** Maximin performance curve under uncertainties in the delays: *Maximizing the worst  $RR$  under the presumed uncertainty set given by Eq. (10), imposes the following tradeoff between the speed  $\langle DT \rangle / \tilde{D}_{tot}$  and accuracy  $(1 - p(\text{err}))$ :*

$$\frac{\langle DT \rangle}{D_{tot}} = \gamma(1 + \alpha_p) \left[ \frac{1}{p(\text{err}) \log \left( \frac{1 - p(\text{err})}{p(\text{err})} \right)} + \frac{1}{1 - 2p(\text{err})} \right]^{-1}, \quad (12)$$

where  $\gamma \equiv \frac{\tilde{D}_{tot}}{D_{tot}}$  is the ratio between the nominal and actual delays. This tradeoff is referred to as the maximin performance curve (MMPC).

**Proof:** See Appendix B.2.

The MMPCs specified by Eq. (12) are simply scaled copies of the OPC given by Eq. (7), which is the special case of a MMPC with  $\alpha_p = 0$  and  $\gamma = 1$ . Thus, the MMPCs define a one-parameter family of performance curves parameterized by the scaling factor  $\gamma(1 + \alpha_p)$ , as depicted in Fig 2. When the nominal delay is exact, i.e.,  $D_{tot} = \tilde{D}_{tot}$  and hence  $\gamma = 1$ , the resulting MMPCs are referred to as the nominal MMPCs (Fig 2, dashed curves). When the nominal delay differs from the actual value, the nominal MMPCs are further scaled by the corresponding ratio  $\gamma$ . Assuming that the actual delay is within the presumed uncertainty set, this ratio is bounded by  $(1 + \alpha_p)^{-1} \leq \gamma \leq (1 - \alpha_p)^{-1}$ . Thus, performance may range within a performance band bounded above by the extreme MMPC given by Eq. (12) with  $\gamma = (1 - \alpha_p)^{-1}$  (Fig 2, dash-dotted curves) and below by the OPC obtained for  $\gamma = (1 + \alpha_p)^{-1}$  (Fig 2, solid curve).

Since the MMPCs are scaled copies of the OPC, they all peak at the same error rate ( $\approx 17.41\%$ , see Appendix A.2). The MMPCs and performance bands capture the qualitative form of the experimental data fairly well, especially for error rates in the range 15–35%, in which mean normalized decision times peak and are widely spread. Quantitative comparisons are given in §4.

**3.1.2 Uncertainties in signal-to-noise ratios**—Apart from delays, performance is also affected by the drift rate  $A$  and noise variance  $\sigma^2$  that characterize the DD process. Uncertainty in these parameters gives rise to uncertainties in the SNR  $\eta = A^2/\sigma^2$ . Here we develop 2AFC strategies that are robust against uncertainties in SNR, but show that the resulting maximin performance curves do not match the experimental data as well as MMPCs under uncertainties in delays.

The analysis is also extended to allow for trial-to-trial variations in SNR. In particular, we focus on variations in SNR that arise from trial-to-trial variations in drift rate, since these have been important in explaining differences between correct and error reaction times. Here, however, the drift rate in each trial is assumed to be bounded within a presumed uncertainty set, instead of drawn from a Gaussian distribution (Ratcliff et al., 1999). We show that such trial-to-trial variations do not affect the performance curves, and thus cannot contribute to explaining the observed sub-optimal performance.

**Uncertainty set for fixed SNR:** First we consider the case in which the SNR is fixed but unknown. The estimated SNR is referred to as the nominal SNR, denoted by  $\tilde{\eta}$ . Given  $\tilde{\eta}$ ,

subjects are presumed to base decisions on the assumption that the actual SNR is a fixed value within the presumed uncertainty set specified by:

$$U_p(\alpha_p, \tilde{\eta}) = \{\eta > 0 : |\eta - \tilde{\eta}| \leq \alpha_p \tilde{\eta}\}. \quad (13)$$

The case in which the actual SNR may vary from trial-to-trial is considered at the end of this section.

**Sources of uncertainty in SNR:** A given level of uncertainty in the SNR may reflect different combinations of uncertainties in drift rate *and* noise variance. While neither of these factors appears explicitly in expression (5) for the reward rate, the drift rate appears implicitly in defining the threshold-to-drift ratio  $\theta = |x_{th}/A|$ . Thus, specifying  $\theta$  defines the size of the actual threshold  $x_{th}$  unambiguously only if uncertainties arise solely from uncertainties in noise variance.

Alternatively, when uncertainties in SNR are due to drift rate alone, analysis is facilitated by normalizing the threshold by noise variance and defining the threshold-to-noise ratio  $\vartheta = |x_{th}/\sigma|$ , which does specify the size of  $x_{th}$  unambiguously in this case. For simplicity we consider only these two ideal cases: uncertainties in SNR arising from noise variance with known drift rate, and uncertainties in SNR arising from drift rate, with known noise variance.

**Uncertainties in noise variance:** The maximin threshold and resulting MMPCs are developed in Appendix B.3, and depicted in Fig. 3 for two levels of presumed uncertainty. The curves exhibit a leftward shift in peak location that is not characteristic of the experimental data, and stands in contrast with the MMPCs for uncertainties in delays (see Fig. 2). We shall nonetheless revisit these MMPCs in the detailed comparison presented in §4, in order to quantitatively assess the relative likelihood of uncertainties in delays versus uncertainties in SNR (due to noise variance) in accounting for the experimental results.

**Uncertainties in drift rate:** The maximin threshold and resulting MM-PCs are developed in Appendix B.4, and depicted in Fig. 4 for two levels of presumed uncertainty. They also exhibit leftward shifts in peak values and moreover lie below the OPC for most error rates. Since these MMPCs deviate even further from the data than those of Fig. 3, they are excluded from the comparisons of §4. They are nonetheless important in evaluating the maximin strategy under trial-to-trial variations in drift rate, as described next.

**Trial-to-trial variations in drift rate:** Trial-to-trial variations in drift rate have been invoked to explain differences between correct and error reaction times (Ratcliff, 1978; Ratcliff et al., 1999; Smith and Ratcliff, 2004). Here we extend our robust analysis to account for such variations and assess their effect on performance curves. Instead of drawing drift rates from a Gaussian distribution, we assume that they belong to an infinite sequence of unknown but bounded values. Following the discussion above, variations in drift rate are formulated as trial-to-trial variations in the SNR with fixed and known noise variance. The presumed uncertainty set for an infinite sequence of SNRs is

$$U_{p,seq}(\alpha_p, \tilde{\eta}) = \{\{\eta_i\}_{i=1}^{\infty} : \eta_i > 0, |\eta_i - \tilde{\eta}| \leq \alpha_p \tilde{\eta}\}, \quad (14)$$

where  $\eta_i$  denotes the SNR on the  $i^{th}$  trial.

With known noise variance, the threshold-to-noise ratio  $\vartheta$  determines the actual threshold unambiguously. For a fixed threshold-to-noise ratio, the worst case scenario that guides the maximin strategy is independent of whether the SNR is fixed or allowed to vary from trial to trial. Hence, as long as the presumed level of uncertainty  $\alpha_p$  is the same, the worst case performance is the same, as stated in the next Theorem:

**Theorem 3.2:** Worst performance with a variable SNR: *Consider a 2AFC task with fixed threshold-to-noise ratio  $\vartheta$ . The worst performance that can occur with any sequence of bounded SNRs consistent with the presumed uncertainty set of Eq. (14) coincides with the worst performance that can occur with a fixed SNR consistent with the presumed uncertainty set of Eq. (13) at the same presumed level of uncertainty  $\alpha_p$ .*

**Proof:** See Appendix B.5.

Theorem 3.2 implies that the performance curves for the variable SNR case considered above are the same as those for the fixed SNR case depicted in Fig. 4. Since these MMPCs do not match the experimental data well, robust strategies for uncertainties in drift rate cannot account for the observed SATs, even when they include bounded trial-to-trial variations.

### 3.2 Robust-satisficing performance curves

**Robust-satisficing decision strategy**—The maximin strategy presumes that the level of uncertainty is known. Furthermore, it appeals to worst case performance, which may be very unlikely, and it yields prudent and pessimistic decisions. Robust-satisficing instead focuses on meeting a sufficient or required level of performance despite the uncertainties (Ben-Haim, 2006). It has been applied successfully to explain a range of decision making, including foraging behavior, where a critical level of food intake is required for survival (Carmel and Ben-Haim, 2005), and equity premium, in which a critical level of return is required (Ben-Haim, 2006).

As shown in this section, applying the robust-satisficing strategy to 2AFC tasks results in performance curves that differ from those associated with the maximin strategy. Under uncertainties in delays, these robust-satisficing performance curves (RSPCs) are marginally inferior to the MMPCs in describing the experimental data, while performance curves for uncertainties in SNR deviate substantially from it. Thus, we focus on robust-satisficing against uncertainties in delays. While the MMPCs do match the data better, we argue in the discussion that this may depend on the reward policy and that robust-satisficing may be relevant for explaining decision making in other 2AFC tasks. The presentation is brief, with details deferred to Appendices.

**Info-gap model**—When the level of uncertainty is unknown, only a family of uncertainty sets, rather than a **unique set**, can be specified. Motivated by the scalar property of interval timing, the presumed uncertainty set of Eq. (10) is extended to define an **unbounded, nested** family of uncertainty intervals parameterized by the level of uncertainty  $\alpha$ :

$$U(\alpha; \tilde{D}, \tilde{D}_{tot}) = \{D, D_{tot} > 0: |D - \tilde{D}| \leq \alpha \tilde{D}, |D_{tot} - \tilde{D}_{tot}| \leq \alpha \tilde{D}_{tot}\}, \forall \alpha \geq 0. \quad (15)$$

Such a structure is known as an information-gap (*info-gap*) model of uncertainty (Ben-Haim, 2006). Nesting implies that **the sets associated with larger uncertainty levels include those** associated with smaller levels, and all the sets include the nominal values for

zero uncertainty. Unbounded-ness implies that the lengths of the intervals increase without bound as the uncertainty level increases.

**Robustness and robust-satisficing threshold**—Consider a fixed threshold  $\theta$ , which achieves the nominal reward rate  $RR(\theta; \eta, \bar{D}, \bar{D}_{tot})$  of Eq (5), under the nominal delays  $\bar{D}$  and  $\bar{D}_{tot}$ . As the level of uncertainty increases, adverse delays may be encountered and the reward rate may deteriorate. Robustness assesses how far the level of uncertainty may increase without jeopardizing the required  $RR$ . Noting that a better than nominal  $RR$  cannot be guaranteed even without uncertainty, the robustness for such a requirement is defined to be zero. Otherwise, the robustness is the largest level of uncertainty under which the required  $RR$  can be guaranteed. For a fixed threshold, robustness is a non-increasing function of  $RR$ :  $RR = 0$  (if such would ever satisfy anyone!) may be achieved with infinite robustness, while better than nominal  $RR$ s have zero robustness. See Appendix B.6 for rigorous definitions and derivations.

Given a sub-optimal required reward rate  $R_r$ , the robust-satisficing strategy selects the threshold that provides the largest robustness. The resulting RSPCs are derived in Appendix B.7 (see Eq. (B.28)). Representative RSPCs are depicted in Fig. 5 for zero penalty delay  $D_{pen} = 0$  (i.e.,  $D = D_{tot}$ ), and three different levels of normalized required performance  $q \equiv (R_r^{-1} - D)/D_{tot}$ . All the RSPCs peak at  $p(err) \approx 13.52\%$ : to the left of the OPCs and MMPCs that peak at  $p(err) \approx 17.41\%$ . Since the RSPCs are evaluated for  $D_{pen} = 0$ , they must be evaluated against experiments without penalty delays. This is done in §4, where it is shown that the leftward shift slightly degrades the data fit (see Fig. 6 and Table 4).

The robust satisficing strategy is relevant only when the required reward rate is sub-optimal, i.e.,  $R_r < RR_{op}(\eta; \bar{D}, \bar{D}_{tot})$ . In Appendix B.7 it is shown that this constrains the RSPCs to lie above the OPC. Since the experimental data also primarily lies above the OPC, this is the main region of interest. Nevertheless, as noted in Appendix B.7, the extension of RSPCs below the OPC is also relevant and the complete RSPCs are used for the analysis of §4.

### 3.3 Summary of robust strategies

We have presented two robust strategies, maximin and robust-satisficing, and derived the associated performance curves under different sources of uncertainty. Performance curves strongly depend on the source of uncertainty and appear to qualitatively match the data of Fig. 1 best for uncertainties in delays, with MMPCs seeming marginally superior to RSPCs (Figs. 2 and 5). Uncertainties in SNR do not appear as relevant to the present data, particularly when they arise from uncertainties in drift rate (Fig. 4) rather than noise variance (Fig. 3).

## 4 Comparisons with experimental data

In this section we assess quantitative fits of robust performance curves to the behavioral data. Following the qualitative analyses of §3 (summarized in §3.3) we focus on the three best candidates: MMPCs and RSPCs for uncertainties in delays, and MMPCs for uncertainties in noise variance. For comparison, we also assess the OPCs for the alternative objective functions introduced in §2.3.

### Experimental data

We limit our comparison to data from blocks with no penalty delays  $D_{pen} = 0$  (i.e.,  $D_{tot} = D$ ), because OPCs for  $RR_m$  and  $RA$  are only defined for  $D_{pen} = 0$ , and RSPCs for uncertainty in  $D$  involve the ratio  $D/D_{tot}$ , which differs between blocks with  $D_{pen} = 0$  and  $D_{pen} > 0$ .

As noted in §2, the SAT differs among subjects with different total scores, and hence performance curves must be compared with data from different groups of subjects separately. In particular, some subjects with the lowest scores had decision times an order of magnitude higher than others (see the bottom panel of Fig. 6), and we therefore analyze the lowest 10% of subjects separately. We then split the remaining 90% into 3 equal groups, but upon finding no significant difference between mean normalized decision times of the 30–60% and 60–90% groups (paired t-test across 10 ranges of error rates:  $p > 0.35$ ), we henceforth pooled the data from these two groups. Normalized decision times for the resulting middle 60% and for the top 30% groups are shown in the upper panels of Fig. 6. The number of experimental conditions falling into each bin (i.e. the number of data points averaged to obtain each mean decision time) is given in Table 2.

### Fitting method

Except for the parameter-free OPC for  $RR$ , each performance curve includes one free parameter, denoted by  $q$  or  $\alpha$ . We estimated these parameters via the maximum likelihood method, using the standard assumption that normalized decision times are normally distributed around the performance curve being fitted, with variance estimated from the data. As noted in §2.3, the OPC for  $RA$  is not defined for weight parameters  $q > 1.096$  (cf. inequality (A.5) of Appendix A.2) and the OPC for  $RR_m$  is not defined for  $p(\text{err})$ s close to 0.5 if  $q > 1$ , so in fitting these functions we restricted to  $q \leq 1.096$  and  $q \leq 1$  respectively. Further comments appear below and in Appendix C.

### Results

Fig. 6 shows the best-fitting performance curves for the three groups, and parameter values estimated for each performance curve are given in Table 3. For each family of performance curves, the corresponding parameter  $\alpha$  or  $q$  increases as the total score decreases, being lowest for the top 30% and highest for the bottom 10% of subjects. This is consistent with natural interpretations of the various strategies, as follows: (i) for MMPCs it implies that poorer performers have higher presumed uncertainty in  $D$  or SNR; (ii) for RSPCs it implies that they require lower levels of reward, and (iii) for the  $RR_m$  and  $RA$  criteria it implies that they place higher emphasis on accuracy. Note that the best fit for the  $RA$  theory with the lowest 10% is with  $q = 1.096$ , at the constraint limit of Eq. (A.5), and that the corresponding OPC has become concave on either side of its sharp peak.

Comparison of likelihoods of data given different performance curves reveals that the MMPC with uncertainty in  $D$  fits the data best for all three groups of subjects. Table 4 shows the ratios of the likelihood of the data given the MMPC with uncertainty in  $D$ , with likelihoods given each of the other performance curves. For the top 30% of subjects all curves fit comparably well, except the OPC for  $RR$ . Differences in fit qualities increase significantly for the middle and bottom groups, and are also high when considering the data from *all* subjects, as summarized in the last column of Table 4. These indicate that the data is several orders of magnitude more likely given the MMPC with uncertainty in  $D$ , than given the OPC for either  $RR$  or  $RR_m$  or the MMPC with uncertainty in SNR. It also shows that the data is over 13 (or 43) times more likely given the MMPC with uncertainty in  $D$ , than the RSPC with uncertainty in  $D$  (or the OPC for  $RA$ ). In summary, the MMPC with uncertainty in  $D$  provides the best fit for all three subgroups, followed closely by the RSPC for  $D$  and the OPC for  $RA$ .

At least two effects influence the reliability of the likelihood ratio estimates. First, the test assumes that experimental decision times are normally distributed around the performance curve being fitted. We assessed this in Appendix C, concluding that there is no evidence for non-Gaussianity in the majority of bins, although up to 3 out of the 10 distributions in each



group are significantly non-Gaussian, being skewed toward long DTs. Second, likelihood ratios may depend on the way that subjects are split into groups, since we implicitly assume a homogeneous strategy within each group. To investigate this we fitted data from six additional splits, ranging from all 80 individuals treated singly, to two groups of 40 subjects each. As described in Appendix C, while different splits lead to different likelihood ratios, the relative ordering of performance curves remains essentially the same as that of Table 4. We therefore believe that the conclusions drawn above and developed in §5, are sound.

We end this section with a brief discussion of the alternative hypothesis noted in §2.2, that subjects may try to optimize, but do so using erroneous parameter estimates. As pointed out there, both over- and under-estimation of thresholds would be expected in such a case, yielding individual performances above and below the OPC. We tested this by computing, for each subgroup of subjects, the fraction of individual performances below the OPC. For the top 30% of subjects this is 50.6%, and for the remaining subgroups 25.9% and 0% respectively. Although the individual subject data shows substantial scatter (not shown here), this finding supports our contention that while the top group of subjects may indeed try to optimize, the remaining 70% intentionally employ higher thresholds.

## 5 Discussion

We have developed an approach to two-alternative forced choice (2AFC) tasks that accounts for potential uncertainties in experimental conditions, specifically, estimated delays and signal-to-noise ratio. Two strategies – maximin and robust satisficing – were proposed. The former assumes that subjects maximize guaranteed performance under a presumed uncertainty level; the latter assumes that they maximize the uncertainty under which a required level of performance can be guaranteed. We compared these strategies with optimal procedures based on reward rate and modified reward rates weighted for accuracy, and found that the maximin strategy with uncertainties in delays predicts performance curves that best match behavioral data from a group of 80 subjects (Fig. 6). Performance curves predicted by the robust satisficing strategy with uncertainties in delays, and optimization of a modified reward-accuracy rate (RA) are the closest competitors. Additional experiments are needed to further assess which strategy is more consistent with human decision making, as discussed at the end of the section.

The robust strategies result in performance bands around the nominal performance curves. These bands explain well the range of mean normalized decision times observed for a given error rate. We also considered the alternative hypothesis that subjects may try to optimize, but with erroneous parameter estimates. We found it inconsistent with the observation that most subjects tended to perform above rather than below the OPC.

Accounting for uncertainty in the parameters of the pure DD model provides an alternative to the probabilistically-extended DD model that includes trial-to-trial parameter variations (Smith and Ratcliff, 2004). We show that uncertainty in SNR can derive from trial-to-trial variations in drift rate, albeit bounded ones, and thus can also account for differences between response times for correct and error trials. However, unlike those for uncertainties in delays, we find that the resulting performance curves do not match well the experimental data (nor do performance curves for Gaussian-distributed drift rates (Bogacz et al., 2006)). This does not mean that trial-to-trial variations in SNR are absent, but it may imply that the presumed uncertainty in the SNR is much smaller than the presumed uncertainty in delays.

It is notable that uncertainties in delays result in performance curves that fit the data well, while uncertainties in SNR do not do so. A major source of the former may be attributed to the scalar property of interval timing (Gibbon, 1977; Buhusi and Meck, 2005). Psychophysical experiments suggest that elapsed time estimates are distributed normally

around the true value, with a standard deviation proportional to it. Thus the sizes of confidence intervals around an estimated experiential delay are proportional to the delay itself. This relationship is captured by the presumed uncertainty set of §3.1, and info-gap model of §3.2. In contrast to potential uncertainties in estimating time intervals, human subjects seem to be very sensitive to SNR levels. For example, Luijendijk (1994) found that the perception threshold for noise in images was as much as 27dB below signal level. This suggests that humans can accurately assess SNR, although further, more direct, experimental verification would be required.

The underlying assumption in this paper is that subjects either optimize or satisfice by setting their decision thresholds based on estimates of intertrial delays and SNR, as modeled by a DD process. In a related study, Simen et al. (2006) proposed a neural network model for rapid threshold adjustment based on estimates of reward rate  $RR$  that are updated as trials proceed. Their model relies on prior, long-term, learning of an approximate linear relationship between maximum  $RR$  and threshold that is independent of SNR over an appropriate SNR range. It too implicitly assumes the ability to estimate time intervals, since the discrete rewards following correct responses must be converted into  $RR$ , and it also suggests that SNR may not play as important a role as delays.

As noted in §3, satisficing was first introduced by Simon (1956, 1982) to account for decisions motivated by a level of aspiration, rather than optimization. Optimization has been heavily criticized due to its dependence on unrealistic amounts of knowledge and simplifying assumptions (Gigerenzer, 2001). Instead, decision making based on bounded rationality (Gigerenzer and Goldstein, 1996; Gigerenzer, 2001, 2002) stresses the importance of proper cues for evaluating different alternatives and heuristics for making fast and frugal decisions. The Take-The-Best heuristic, for example, suggests focusing on the best cue that differentiates between the alternatives. In the context of the DD model Eq. (1), this may suggest monitoring the state variable  $x(t)$  as the cue; but the issue of determining the threshold is left open.

We propose **two kinds of** additional experiments to investigate whether subjects satisfice or optimize in 2AFC: (i) direct investigation of the effects of induced uncertainties, and (ii) correlation of performance in 2AFC with accuracy in interval timing. Uncertainties could be induced by varying SNRs **randomly in each block of trials, although this** violates the assumption of statistical stationarity on which the optimality proofs for the SPRT and DD model rely. **However, instead of keeping  $D_{RSI}$  constant on each block, RSIs could be drawn from a distribution with a fixed mean  $\bar{D}_{RSI}$ , thus keeping the mean  $\bar{D}_{tot}$  constant on blocks.  $\bar{D}_{RSI}$  and the variance of the distribution could then be changed from block to block and the effects on performance** assessed in comparison with the different decision strategies.

Alternatively, a subject's standard deviation in interval timing,  $\sigma_I$ , could be measured directly in separate time-estimation trials, and then correlated with performance on 2AFC tasks. The maximin strategy suggests that the normalized decision time  $\langle DT \rangle / D_{tot}$  is proportional to  $1 + \alpha_p$ . Assuming that the presumed uncertainty  $\alpha_p$  is proportional to  $\sigma_I$ , there should be a correlation (across subjects) between  $\langle DT \rangle / D_{tot}$ 's in 2AFC experiments and average errors in interval timing. One could also correlate timing ability with performance on a deadline choice task with substantial penalties for failures to respond before the deadline. Unlike the free response paradigm, poor timers are predicted to respond prematurely, and hence *faster and less accurately*, under such conditions (Frazier and Yu, 2007): in the opposite direction to their suboptimal slower and more accurate free response behavior.

The robust satisficing strategy is relevant when a critical performance level must be met, while the maximin strategy is appropriate for optimizing performance under an expected level of uncertainty. The current results indicate that the majority of subjects in 2AFC tasks seem to follow maximin rather than robust satisficing or optimal performance curves. It would also be interesting to investigate whether subjects change their strategy under different conditions, e.g., do they resort to robust satisficing when the success of the whole block (and the consequent net reward) depends on achieving a specific level of performance. This question could be addressed by additional experiments in which a fixed reward is given when outperforming a preset required level, instead of the current paradigm in which the reward increases linearly with the number of correct responses.

As noted in §1, the pure DD model is analytically tractable, yielding explicit, parameter-free OPCs. It supports a normative decision theory that can explain sub-optimal experimental data by allowing analytical derivation of one-parameter models for comparison with the optimal strategy that maximizes reward rate. The current analysis suggests that robust strategies against uncertainties in delays best explain the sub-optimal results in the present data. Nevertheless, additional potential sources of nonoptimality identified in §2.1 should also be investigated.

## Acknowledgments

This work was supported by PHS grants MH58480 and MH62196 (Cognitive and Neural Mechanisms of Conflict and Control, Silvio M. Conte Center). MZ was supported by the Abramsom Center for the Future of Health and by the fund for the promotion of research at the Technion, and RB was supported by EPSRC grant EP/C514416/1. The authors thank Yakov Ben-Haim, Jonathan Cohen, Pat Simen, Angela Yu and the members of the Conte Center modeling group for numerous helpful suggestions.

## References

- Abramowitz, M.; Stegun, I., editors. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover Publications; New York: 1972.
- Akaike H. Likelihood of a model and information criteria. *J Econometrics*. 1981; 16:3–14.
- Ben-Haim, Y. Information Gap Decision Theory: Decisions under Severe Uncertainty. 2. Academic Press; New York: 2006.
- Ben-Tal A, Nemirovski A. Selected topics in robust convex optimization. *Mathematical Programming*. 2008; 112:125–158.
- Bogacz R, Brown E, Moehlis J, Holmes P, Cohen J. The physics of optimal decision making: A formal analysis of models of performance in two alternative forced choice tasks. *Psych Rev*. 2006; 113(4): 700–765.
- Bogacz R, Hu P, Cohen J, Holmes P. Do humans produce the speed-accuracy tradeoff that maximizes reward rate? *Quart J Exp Psych*. 2009 In press.
- Bohil C, Maddox W. On the generality of optimal versus objective classifier feedback effects on decision criterion learning in perceptual categorization. *Memory & Cognition*. 2003; 31(2):181–198.
- Britten K, Shadlen M, Newsome W, Movshon J. Responses of neurons in macaque MT to stochastic motion signals. *Visual Neurosci*. 1993; 10:1157–1169.
- Buhusi C, Meck W. What makes us tick? functional and neural mechanisms of interval timing. *Nature*. 2005; 6:755–765.
- Bussemeyer J, Myung I. An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *J Exp Psych General*. 1992; 121(2):177–194.
- Bussemeyer J, Rapoport A. Psychological models of deferred decision making. *J Math Psych*. 1988; 32:91–134.
- Bussemeyer J, Townsend J. Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*. 1993; 100:432–459. [PubMed: 8356185]

- Carmel Y, Ben-Haim Y. Info-gap robust-satisfying model of foraging behavior: Do foragers optimize or satiate? *American Naturalist*. 2005; 166(5):633–641.
- Cohen J, Dunbar K, McClelland J. On the control of automatic processes: A parallel distributed processing model of the Stroop effect. *Psychological Review*. 1990; 97(3):332–361. [PubMed: 2200075]
- Conover, W. *Practical Nonparametric Statistics*. Wiley; New York: 1980.
- Cox, D. *Renewal Theory*. Methuen, London, monographs on applied probability and statistics. 1962.
- Diederich A, Busemeyer J. Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. *J Math Psych*. 2003; 47:304–322.
- Ditterich J. Evidence for time-variant decision making. *European J of Neurosci*. 2006a; 24:3628–3641.
- Ditterich J. Stochastic models of decisions about motion direction: Behavior and physiology. *Neural Networks*. 2006b; 19:981–1012. [PubMed: 16952441]
- Eckhoff P, Holmes P, Law C, Connolly P, Gold J. On diffusion processes with variable drift rates as models for decision making during learning. *New J of Physics*. 2008; 10:1088/1367–2630/10/1/015006.
- Edwards W. Models for statistics, choice reaction times, and human information processing. *J Math Psych*. 1965; 2:312–329.
- Frazier, P.; Yu, A. *Advances in Neural Information Processing Systems*. Neural Information Processing Systems Foundation; 2007. Sequential hypothesis testing under stochastic deadlines. downloadable from <http://books.nips.cc/nips20.html>
- Gardiner, C. *Handbook of Stochastic Methods*. 2. Springer; New York: 1985.
- Gibbon J. Scalar expectancy theory and Weber's law in animal timing. *Psych Rev*. 1977; 84(3):279–325.
- Gigerenzer, G. Decision making: nonrational theories. In: Smelser, N.; Bates, P., editors. *International Encyclopedia of the Social and Behavioral Sciences*. Vol. 4. MIT Press; Cambridge, MA: 2001. p. 3304–4409.
- Gigerenzer, G. The adaptive toolbox. In: Gigerenzer, G.; Selten, R., editors. *Bounded Rationality: the adaptive toolbox*. MIT Press; Cambridge, MA: 2002.
- Gigerenzer G, Goldstein D. Reasoning the fast and frugal way: Models of bounded rationality. *Psych Rev*. 1996; 103:650–669.
- Gold J, Shadlen M. Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Science*. 2001; 5 (1):10–16.
- Gold J, Shadlen M. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*. 2002; 36:299–308. [PubMed: 12383783]
- Holmes P, Shea-Brown E, Moehlis J, Bogacz R, Gao J, Aston-Jones G, Clayton E, Rajkowski J, Cohen J. Optimal decisions: From neural spikes, through stochastic differential equations, to behavior. *IEICE Transactions on Fundamentals on Electronics, Communications and Computer Science*. 2005; E88A(10):2496–2503.
- Judge, G.; Hill, R.; Griffiths, W.; Lutkepohl, H.; Lee, T-C. *Introduction to the Theory and Practice of Econometrics*. Wiley; New York: 1988.
- Laming, D. *Information Theory of Choice-Reaction Times*. Academic Press; New York: 1968.
- Liu Y, Holmes P, Cohen J. A neural network model of the Eriksen task: Reduction, analysis, and data fitting. *Neural Computation*. 2008; 20 (2):345–373. [PubMed: 18045022]
- Luijendijk H. Practical experiment on noise perception in noisy images. *Proc Soc of Optical Engineering*. 1994; 2166:2–8.
- MacKay, D. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press; Cambridge, U.K: 2003.
- Maddox W, Bohil C. Base-rate and payoff effects in multidimensional perceptual categorization. *J Exp Psych*. 1998; 24(6):1459–1482.
- Mazurek M, Roitman J, Ditterich J, Shadlen M. A role for neural integrators in perceptual decision making. *Cerebral Cortex*. 2003; 13 (11):891–898.
- Myung I, Busemeyer J. Criterion learning in a deferred decision making task. *Amer J Psych*. 1989; 102(1):1–16.

- Rapoport A, Burkheimer G. Models for deferred decision making. *J Math Psych.* 1971; 8:508–538.
- Ratcliff R. A theory of memory retrieval. *Psych Rev.* 1978; 85:59–108.
- Ratcliff R, Cheria A, Segraves M. A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two choice decisions. *J Neurophysiol.* 2003; 90:1392–1407. [PubMed: 12761282]
- Ratcliff R, Hasegawa Y, Hasegawa R, Smith P, Segraves M. Dual-diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *J Neurophysiol.* 2006; 97:1756–1774. [PubMed: 17122324]
- Ratcliff R, Tuerlinckx F. Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonom Bull Rev.* 2002; 9:438–481.
- Ratcliff R, Van Zandt T, McKoon G. Connectionist and diffusion models of reaction time. *Psych Rev.* 1999; 106 (2):261–300.
- Reber, A. Practical Nonparametric Statistics. The Penguin Dictionary of Psychology. Penguin Books; London, U.K: 1995.
- Roitman J, Shadlen M. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci.* 2002; 22 (21):9475–9489. [PubMed: 12417672]
- Roxin A, Ledberg A. Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Computational Biology.* 2008; 4(3):10.1371/journal.pcbi.1000046.
- Schall J. Neural basis of deciding, choosing and acting. *Nature Reviews in Neuroscience.* 2001; 2:33–42.
- Simen P, Cohen J, Holmes P. Rapid decision threshold modulation by reward rate in a neural network. *Neural Networks.* 2006; 19:1013–1026. [PubMed: 16987636]
- Simon H. Rational choice and the structure of environments. *Psych Rev.* 1956; 23:129–138.
- Simon, H. Models of bounded rationality. MIT Press; Cambridge, MA: 1982.
- Smith P, Ratcliff R. Psychology and neurobiology of simple decisions. *Trends in Neurosci.* 2004; 27 (3):161–168.
- Stone M. Models for choice-reaction time. *Psychometrika.* 1960; 25:251–260.
- Usher M, McClelland J. On the time course of perceptual choice: The leaky competing accumulator model. *Psych Rev.* 2001; 108:550–592.
- Wald, A. Sequential Analysis. Wiley; New York: 1947.
- Wald, A. Statistical Decision Functions. John Wiley & Sons, Inc; 1950.
- Wald A, Wolfowitz J. Optimum character of the sequential probability ratio test. *Ann Math Statist.* 1948; 19:326–339.
- Wang XJ. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron.* 2002; 36:955–968. [PubMed: 12467598]
- Wong K, Wang XJ. A recurrent network mechanism of time integration in perceptual decisions. *J Neurosci.* 2006; 26 (4):1314–1328. [PubMed: 16436619]
- Zacksenhouse M, Nemets S, Lebedev M, Nicolelis M. Robust-satisficing linear regression: performance/robustness tradeoff and consistency criterion. *Mechanical Systems and Signal Processing.* 2009 In Press. 10.1016/j.ymssp.2008.09.008.

## Appendices: Proofs and mathematical details

### A Reward rate and optimal performance

#### A.1 Reward rate

##### Lemma A.1

Let the random variable  $S_t$  denote the number of correct decisions in time  $t$ . As  $t \rightarrow \infty$ , the mean number of correct decisions per unit time  $\langle S_t \rangle / t$  approaches the RR defined by Eq. (4).

## Proof

Consider the DD model of Eq. (1), which is re-initialized immediately after crossing the threshold. The resulting passage times define a sequence of points in time, and hence may be regarded as a realization of a point process. Furthermore, since the first passage time is independent of the previous passage times, the resulting point process can be described as a renewal process (Cox, 1962), i.e., a point process in which intervals between the points are the realizations of a random variable  $T$  whose pdf depends only on the time since the last point (i.e., last passage time). Let  $N_t$  be the random variable indicating the number of passages in time  $t$ . For a renewal process the asymptotic distribution of  $N_t$  for large  $t$  is normal with mean  $\langle N_t \rangle = t/\langle T \rangle$  (Cox, 1962, Eq. (3.3.3)) (this relies on the asymptotic normal distribution of the sum of the first  $r$  renewals for large  $r$ ). So, for large  $t$ ,  $t = \langle T \rangle \langle N_t \rangle$ .

Let  $p(\text{corr})$  be the probability that a decision is correct. Then, for  $n$  decisions the mean number of correct decisions is:  $\langle S_t | N_t = n \rangle = p(\text{corr}) n$ ; and the mean number of correct decisions is:

$$\langle S_t \rangle = \sum_{n=0}^{\infty} \langle S_t | N_t = n \rangle \Pr[N_t = n] = p(\text{corr}) \sum_{n=0}^{\infty} n \Pr[N_t = n] = p(\text{corr}) \langle N_t \rangle. \quad (\text{A.1})$$

Hence, asymptotically as  $t \rightarrow \infty$  we have

$$\frac{\langle S_t \rangle}{t} = \frac{p(\text{corr}) \langle N_t \rangle}{\langle T \rangle \langle N_t \rangle} = \frac{p(\text{corr})}{\langle T \rangle}. \quad (\text{A.2})$$

The proof is completed by noting that the decision times form a renewal process too, where each decision interval equals the corresponding first passage time plus processing and delay intervals. Substituting  $p(\text{corr}) = 1 - p(\text{err})$  and  $\langle T \rangle = \langle DT \rangle + T_0 + D_{RSI} + D_{pen} p(\text{err})$  in Eq. (A.2) results in Eq. (4).

## A.2 Families of OPCs for alternative objective functions RA and RR<sub>m</sub>

Employing Eqs. (3) and an analogue of (6) as in §2.2, the expression (8) for  $RA$  may be maximized to yield a family of OPCs parameterized by  $q$ :

$$\frac{\langle DT \rangle}{D_{tot}} \frac{\mathcal{E} - 2q - \sqrt{\mathcal{E}^2 - 4q(\mathcal{E} + 1)}}{2q}, \quad (\text{A.3})$$

$$\text{where } \mathcal{E} = \left[ \frac{1}{p(\text{err}) \log \left( \frac{1-p(\text{err})}{p(\text{err})} \right)} + \frac{1}{1-2p(\text{err})} \right] \quad (\text{A.4})$$

is the reciprocal of the OPC formula (7). In Lemma A.2 below we prove that  $\mathcal{E}$  has a unique minimum (and hence that the OPC has a unique peak) at  $p(\text{err}) \approx 0.1741$ . It follows that the mean normalized decision times given by Eq. (A.3) are non-negative real numbers provided that the weight satisfies the following inequality:

$$q \leq \min \left\{ \frac{\mathcal{E}^2}{4(\mathcal{E} + 1)} \right\} = \frac{\mathcal{E}_{min}^2}{4(\mathcal{E}_{min} + 1)}. \quad (\text{A.5})$$



Numerically, we find  $\varepsilon_{\min} \approx 5.224$ , implying that  $q \approx 1.096$ . Moreover, differentiating the expression in (A.3-A.4) with respect to  $p(\text{err})$ , we find that its critical points coincide with the minimum of  $\varepsilon$  for all  $q$  satisfying (A.5), implying that the OPCs for  $RA$  all peak at  $p(\text{err}) \approx 0.1741$ , as for  $RR$ .

A similar computation for the objective function (9) for  $RR_m$  leads to the following OPC family:

$$\frac{\langle DT \rangle}{D_{tot}} = (1+q) \left[ \frac{\frac{1}{p(\text{err})} - \frac{q}{1-p(\text{err})}}{\log\left(\frac{1-p(\text{err})}{p(\text{err})}\right)} + \frac{1-q}{1-2p(\text{err})} \right]^{-1}, \quad (\text{A.6})$$

whose mean normalized decision times are non-negative provided that  $q \geq 1$ . Unlike the OPCs for  $RA$ , the maxima of the family (A.6) move rightward with increasing  $q$  (Fig. 6). Eqs. (A.3) and (A.6) both reduce to (7) for  $q = 0$ .

### Lemma A.2

The function  $\varepsilon$  of Eq. (A.4) has a unique minimum  $\varepsilon \approx 5.224$  at  $ER \approx 0.1741$  and  $\varepsilon$  approaches  $+\infty$  as  $ER \rightarrow 0$  and as  $ER \rightarrow 0.5$ .

### Proof

The limits at  $ER = 0$  and  $0.5$  are obtained directly for the second term in  $\varepsilon$ , and via L'Hôpital's rule for the term involving logarithms. Before computing the derivative to prove that there is a unique minimum, it is convenient to change variables by defining

$$y = \frac{1-p(\text{err})}{p(\text{err})} \Rightarrow p(\text{err}) = \frac{1}{1+y}; \quad (\text{A.7})$$

note that as  $p(\text{err})$  rises from  $0$  to  $0.5$ ,  $y$  falls monotonically from  $+\infty$  to  $1$ . We then have

$$\frac{d\varepsilon}{dy} = \frac{d}{dy} \left[ \frac{y+1}{\log(y)} + \frac{y+1}{y-1} \right] = \frac{\log(y)-1-y^{-1}}{[\log(y)]^2} - \frac{2}{(y-1)^2}. \quad (\text{A.8})$$

Setting Eq. (A.8) equal to zero and rearranging, we find that critical points of  $\varepsilon$  occur at the roots of

$$\log(y)-y^{-1} = 1 + 2 \left[ \frac{\log(y)}{y-1} \right]^2, \quad (\text{A.9})$$

but in fact the solution is unique, since the left hand side of (A.9) is strictly increasing from  $-1$  to  $\infty$  and its right hand monotonically decreases from  $2$  to  $1$ , as  $y$  goes from  $1$  to  $\infty$ . The former claim is easy to check, and the latter follows from computation of the derivative of  $[\log(y)/(y-1)]$ :

$$\frac{d}{dy} \left[ \frac{\log(y)}{y-1} \right] = \frac{1-y^{-1}-\log(y)}{(y-1)^2}. \quad (\text{A.10})$$

This expression is clearly negative for all  $y > e$ , and the following series expansion shows that it is in fact negative for all  $y > 1$  and zero only at  $y = 1$ :

$$\log(y) = \sum_{j=1}^{\infty} \frac{1}{j} \left( \frac{y-1}{y} \right)^j = 1 - y^{-1} + \sum_{j=2}^{\infty} \frac{1}{j} \left( \frac{y-1}{y} \right)^j, \quad (\text{A.11})$$

see Abramowitz and Stegun, eds. (1972, Formula 4.1.25). Numerical solution of Eq. (A.9) yields the estimates of the lemma.

## B Robust performance

### B.1 Extreme performance under uncertainties in delays

The worst (minimal) reward rate that may occur under the presumed uncertainty set  $U_p(\alpha_p; \tilde{D}, \tilde{D}_{tot})$  for the delays, is denoted by:

$RR_{MIN}(\alpha_p; \theta, \eta, \tilde{D}, \tilde{D}_{tot}) \equiv \min_{D, D_{tot} \in U_p(\alpha_p; \tilde{D}, \tilde{D}_{tot})} RR(\theta; \eta, D, D_{tot})$ . According to Eq. (5), the  $RR$  is minimized when the actual delays are the longest possible within the presumed uncertainty set. Given the presumed uncertainty set of Eq. (10), the longest possible delays are  $D = \tilde{D}(1 + \alpha_p)$  and  $D_{tot} = \tilde{D}_{tot}(1 + \alpha_p)$ , so:

$$RR_{MIN}(\alpha_p; \theta, \eta, \tilde{D}, \tilde{D}_{tot}) = [\theta + \tilde{D}(1 + \alpha_p) + [\tilde{D}_{tot}(1 + \alpha_p) - \theta] \exp(-2\eta\theta)]^{-1} \quad (\text{B.1})$$

We note that when there are no delays, the reward rate, denoted by  $RR_0$ , is  $RR_0(\theta, \eta) \equiv [\theta(1 - \exp(-2\eta\theta))]^{-1}$ .

Similarly, the best performance (maximum  $RR$ ) that can be obtained is denoted as:

$RR_{MAX}(\alpha_p; \theta, \eta, \tilde{D}, \tilde{D}_{tot}) \equiv \max_{D, D_{tot} \in U_p(\alpha_p; \tilde{D}, \tilde{D}_{tot})} RR(\theta; \eta, D, D_{tot})$ . According to Eq. (5),  $RR$  is maximized when the actual delays are the shortest, i.e., when  $D = \max(0, \tilde{D}(1 - \alpha))$  and  $D_{tot} = \max(0, \tilde{D}_{tot}(1 - \alpha))$ . For  $\alpha = 1$ , we have:

$$RR_{MAX}(\alpha_p \leq 1; \theta, \eta, \tilde{D}, \tilde{D}_{tot}) = [\theta + \tilde{D}(1 - \alpha_p) + [\tilde{D}_{tot}(1 - \alpha_p) - \theta] \exp(-2\eta\theta)]^{-1} \quad (\text{B.2})$$

The case  $\alpha = 1$  includes the favorable condition of zero delays, so  $RR_{MAX}(\alpha_p = 1; \theta, \eta, \tilde{D}, \tilde{D}_{tot}) = RR_0(\theta, \eta)$ .

### B.2 Maximin performance curves under uncertainties in delays

Here we prove Theorem 3.1. Specifically, we develop the maximin performance curves (MMPC) given that the uncertainties in the delays are described by the presumed uncertainty set of Eq. (10) with the presumed level of uncertainty  $\alpha_p$ .

#### Proof of Theorem 3.1

The inner minimization in Eq. (11) specifies the worst reward rate that might be obtained using the threshold  $\theta$ , given the uncertainty set of Eq. (10) and the presumed level of uncertainty  $\alpha_p$ . The worst (minimal) reward rate  $RR_{MIN}$ , is derived in Appendix B.1, and given by Eq. (B.1). The maximin threshold  $\theta_{MM}$  should maximize Eq. (B.1). Differentiating the  $RR_{MIN}(\alpha_p; \theta, \eta, \tilde{D}, \tilde{D}_{tot})$  given by Eq. (B.1) with respect to  $\theta$  and setting this derivative to zero results in the following condition for the maximin threshold  $\theta_{MM}$ :

$$\exp(2\eta\theta_{MM}) - 1 = 2\eta[\tilde{D}_{tot}(1 + \alpha_p) - \theta_{MM}]. \quad (\text{B.3})$$

It is straightforward to verify that when this condition holds, the second derivative is negative, and hence condition Eq. (B.3) defines a unique maximum for the worst reward rate,  $RR_{MIN}$ .

Finally, substituting Eq. (3) in Eq. (B.3), and setting  $\gamma \equiv \frac{\tilde{D}_{tot}}{D_{tot}}$  result in the *maximin performance curve* (MMPC) of Eq. (12)

### B.3 Maximin Performance curves for uncertainties in noise variance

Here we consider uncertainties in the SNR that arise from uncertainties in the noise variance. Thus, the uncertainty set for the SNR, specified by Eq. (13) is assumed to reflect uncertainties in the noise variance, while the drift rate  $A$  in each experimental block is assumed fixed and known. In this case, determining the threshold-to-drift ratio  $\theta$  unambiguously defines the threshold  $x_{th} = A\theta$  for the DD process of Eq. (1).

#### Definition B.1

Maximin threshold-to-drift ratio under uncertainties in the SNR: *The maximin threshold-to-drift ratio  $\theta_{MM}$  is the one that maximizes the worst RR under the presumed level of uncertainty  $\alpha_p$  in the SNR, given the delays and nominal SNR:*

$$\theta_{MM}(\alpha_p; \tilde{\eta}, D, D_{tot}) = \arg \max_{\theta} \left( \min_{\eta \in U_p(\alpha_p; \tilde{\eta})} RR(\theta; \eta, D, D_{tot}) \right). \quad (\text{B.4})$$

#### Theorem B.1

Maximin threshold-to-drift ratio under uncertainties in SNR: *Given the presumed uncertainty set given by Eq. (13) with  $\alpha_p \leq 1$ , the maximin threshold-to-drift  $\theta_{MM}$  satisfies:*

$$\exp(2\tilde{\eta}(1 - \alpha_p)\theta_{MM}) - 1 = 2\tilde{\eta}(1 - \alpha_p)(D_{tot} - \theta_{MM}), \quad (\text{B.5})$$

#### Proof

The inner minimization in Eq. (B.4) specifies the worst reward rate that may occur under the presumed uncertainty set  $U_p(\alpha_p; \tilde{\eta})$  for the SNR, and is denoted by:

$RR_{MIN}(\alpha_p; \theta, \tilde{\eta}, D, D_{tot}) \equiv \min_{\eta \in U_p(\alpha_p; \tilde{\eta})} RR(\theta; \eta, D, D_{tot})$ . According to Eq. (5), the minimization of the  $RR$  with respect to  $\eta$  depends on the sign of  $D_{tot} - \theta$ . We note that for the optimal threshold, Eq. (6) implies that  $D_{tot} - \theta_{op} > 0$ .

First, we consider thresholds that satisfy  $D_{tot} - \theta > 0$ , for which  $RR$  is minimized at the lowest SNR, i.e.,  $\eta_L = \max\{0, (1 - \alpha_p)\tilde{\eta}\}$ , so for  $\alpha_p \leq 1$ :

$$RR_{MIN}(\alpha_p \leq 1; \theta, \tilde{\eta}, D, D_{tot}) = [\theta + D + (D_{tot} - \theta)\exp(-2\tilde{\eta}(1 - \alpha_p)\theta)]^{-1}. \quad (\text{B.6})$$

The case  $\alpha_p = 1$  includes the worst condition of zero SNR. The resulting performance  $RR_{MIN}(\alpha_p = 1; \theta, \tilde{\eta}, D, D_{tot}) = [D + D_{tot}]^{-1}$ , corresponding to instant response at chance level (i.e., with  $p(\text{err}) = 0.5$  cf. Eq. (2)), is referred to as *chance performance* and denoted by

$RR_{chance} \equiv [D + D_{tot}]^{-1}$ . Chance performance can be achieved even under infinite uncertainty, and is thus of no interest.

Concentrating on  $\alpha_p = 1$ , we note that for either  $\theta = 0$  or  $\theta = D_{tot}$ , the worst reward rate is the chance level:  $RR_{MIN}(\alpha_p = 1 : \theta = 0, \tilde{\eta}, D, D_{tot}) = RR_{MIN}(\alpha_p = 1 : \theta = D_{tot}, \tilde{\eta}, D, D_{tot}) = RR_{chance}$ . Furthermore, at  $\theta = 0$ , the derivative of  $RR_{MIN}(\alpha_p = 1 : \theta, \tilde{\eta}, D, D_{tot})$  with respect to  $\theta$  is positive, and hence there is at least a local maximum of Eq. (B.6) for thresholds in the range  $0 < \theta < D_{tot}$ . Differentiating Eq. (B.6) with respect to  $\theta$ , and setting the derivative to zero, indicates that  $\theta_{MM}$  must satisfy Eq. (B.5). It is straightforward to verify that  $RR_{MIN}(\alpha_p = 1 : \theta_{MM}, \tilde{\eta}, D, D_{tot})$  is a maximum by evaluating the second derivative at  $\theta_{MM}$  that satisfies Eq. (B.5).

For large thresholds that satisfy  $D_{tot} - \theta < 0$ , the  $RR$  is minimized at the highest SNR, i.e.,  $\eta_H = (1 + \alpha_p)\tilde{\eta}$ , so:

$$RR_{MIN}(\alpha_p : \theta, \tilde{\eta}, D, D_{tot}) = [\theta + D + (D_{tot} - \theta)\exp(-2\tilde{\eta}(1 + \alpha_p)\theta)]^{-1}. \quad (B.7)$$

Noting that the last term is always negative, and that since  $\alpha_p = 0$ , the exponent is always less than unit, i.e.,  $(\exp(-2\tilde{\eta}(1 + \alpha_p)\theta) < 1)$ , we conclude that  $RR_{MIN}(\alpha_p : \theta, \tilde{\eta}, D, D_{tot}) < [\theta + D + (D_{tot} - \theta)]^{-1} = RR_{chance}$ . As argued above, sub-chance performance levels, i.e.,  $RR < RR_{chance}$ , can be achieved even under infinite uncertainty, and are thus of no interest.

When the presumed level of uncertainty is zero, Eq. (B.5) or the maximin threshold-to-drift ratio reduces to Eq. (6) for the optimum threshold.

Assuming that the nominal SNR is the actual SNR ( $\eta = \tilde{\eta}$ ), the nominal MMPC for a presumed level of uncertainty  $\alpha_p = 1$  in the SNR, can be derived by substituting Eq. (3) in Eq. (B.5):

$$\frac{\langle DT \rangle}{D_{tot}} = [1 - 2p(\text{err})] \left[ \frac{\left( \frac{1 - p(\text{err})}{p(\text{err})} \right)^\chi - 1}{\chi \log \left( \frac{1 - p(\text{err})}{p(\text{err})} \right)} + 1 \right]^{-1}, \quad (B.8)$$

where  $\chi \equiv 1 - \alpha_p$ . Under uncertainties in the SNR, the MMPCs are not scaled version of the OPC (Eq. (7)), but reduce to it when the uncertainty vanishes ( $\chi = 1$ ).

If the SNR varies within the range consistent with the presumed level of uncertainty, i.e.,  $\eta \in [\tilde{\eta}(1 - \alpha_p), \tilde{\eta}(1 + \alpha_p)]$  for  $\alpha_p = 1$ , the maximin performance bands are bounded below (for  $\eta = \tilde{\eta}(1 - \alpha_p)$ ) by the OPC (Eq. (7)), and above (for  $\eta = \tilde{\eta}(1 + \alpha_p)$ ) by Eq. (B.8) with

$$\chi = \frac{1 - \alpha_p}{1 + \alpha_p}.$$

## B.4 Maximin performance curves under uncertainties in drift rate

Here we consider uncertain drift rates. In this case, specifying the threshold-to-drift ratio  $\theta \equiv |x_{th}/A|$  does not specify the process threshold  $x_{th}$  unambiguously, since the latter depends on the uncertain drift rate. Instead we introduce the threshold-to-noise ratio  $\vartheta \equiv |x_{th}/\sigma|$ . Assuming that the noise variance  $\sigma^2$  is fixed and known, determining the threshold-to-variance ratio  $\vartheta$  unambiguously defines the process threshold  $x_{th}$  for the drift-diffusion process of Eq. (1). Note that  $\vartheta = \theta \sqrt{n}$ , so the quantities describing performance:  $p(\text{err})$  and  $\langle DR \rangle$  in Eq. (2) and  $RR$  in Eq. (5), can be expressed in terms of  $\vartheta$  and  $\eta$  in place of  $\theta$  and  $\eta$ . In particular, the reward rate can be expressed as a function of  $\eta$  and  $\vartheta$ .

$$RR(\vartheta; \eta, D, D_{tot}) = \left[ \frac{\vartheta}{\sqrt{\eta}} + D + (D_{tot} - \frac{\vartheta}{\sqrt{\eta}}) \exp(-2\sqrt{\eta}\vartheta) \right]^{-1}. \quad (\text{B.9})$$

### Definition B.2

Maximin threshold-to-noise ratio under uncertainties in SNR: *The maximin threshold-to-noise ratio  $\vartheta_{MM}$  is the threshold that maximizes the worst RR under the presumed uncertainty set  $U_p(\alpha_p, \tilde{\eta})$  for the SNR, given the level of uncertainty  $\alpha_p$ , the delays,  $D$  and  $D_{tot}$ , and the nominal SNR,  $\tilde{\eta}$ :*

$$\vartheta_{MM}(\alpha_p; \tilde{\eta}, D, D_{tot}) = \arg \max_{\vartheta} \left( \min_{\eta \in U_p(\alpha_p, \tilde{\eta})} RR(\vartheta; \eta, D, D_{tot}) \right). \quad (\text{B.10})$$

### Theorem B.2

Maximin threshold-to-noise ratio under uncertainties in SNR: *Given the presumed uncertainty set for the SNR given by Eq. (13), with a presumed level of uncertainty  $\alpha_p \leq 1$ , the maximin threshold-to-noise  $\vartheta_{MM}$  satisfies:*

$$\exp(2\sqrt{\tilde{\eta}(1-\alpha_p)}\vartheta_{MM}) - 1 = 2\tilde{\eta}(1-\alpha_p)(D_{tot} - \frac{\vartheta_{MM}}{\sqrt{\tilde{\eta}(1-\alpha_p)}}), \quad (\text{B.11})$$

### Proof

For a fixed  $\vartheta$ , the reward rate given by Eq. (B.9) is an increasing function of  $\eta$ , and hence its minimum occurs at the lowest SNR consistent with the presumed uncertainty  $\alpha_p \leq 1$ , i.e.,  $\eta_L = (1 - \alpha_p)\tilde{\eta}$ . The maximin threshold-to-noise variance, defined by Eq. (B.10), is obtained by substituting the lowest SNR in Eq. (B.9), differentiating with respect to  $\vartheta$ , and equating to zero.

As in Appendix B.3, the case  $\alpha_p = 1$  includes the worst case of zero SNR, which results in chance performance, which is of no further interest.

Given  $\tilde{\eta}$ , the threshold-to-noise ratio is related to the threshold-to-drift ratio by  $\vartheta = \theta \sqrt{\tilde{\eta}}$ . Thus, the expression for  $\theta$  in Eq. (3) can be used to express  $\vartheta_{MM}$ , and consequently the condition in Eq. (B.11), in terms of the two performance parameters,  $p(\text{err})$  and  $\langle DT \rangle$ , and the nominal SNR,  $\tilde{\eta}$ . Assuming that the actual SNR equals the nominal SNR ( $\eta = \tilde{\eta}$ ), the later can also be expressed in the terms of the performance parameters to derive the nominal MMPC for a fixed level of presumed uncertainty  $\alpha_p \leq 1$  in the SNR:

$$\frac{\langle DT \rangle}{D_{tot}} = \sqrt{\chi}(1 - 2p(\text{err})) \left[ \frac{\left( \frac{1-p(\text{err})}{p(\text{err})} \right) \sqrt{\chi} - 1}{\sqrt{\chi} \log \left( \frac{1-p(\text{err})}{p(\text{err})} \right)} + 1 \right]^{-1}, \quad (\text{B.12})$$

where, as in Eq. (B.8),  $\chi \equiv 1 - \alpha_p$

If the SNR varies within the range consistent with the presumed level of uncertainty, i.e.,  $\eta \in [\tilde{\eta}(1 - \alpha_p), \tilde{\eta}(1 + \alpha_p)]$  for  $\alpha_p \leq 1$ , the maximin performance bands are bounded by the OPC (Eq. (7), obtained for  $\eta = \tilde{\eta}(1 - \alpha_p)$ ) and by Eq. (B.12) with  $\chi = \frac{1-\alpha_p}{1+\alpha_p}$  (for  $\eta = \tilde{\eta}(1 + \alpha_p)$ ).

## B.5 Performance under uncertain and variable drift rate

Here we prove Theorem 3.2, which states that for a fixed threshold-to-noise ratio, the worst performance with a sequence of possibly variable SNRs, within the presumed uncertainty set defined by Eq. (14), is the same as that for a fixed SNR within the presumed uncertainty set defined by Eq. (13), under the same level of presumed uncertainty  $\alpha_p$ .

### Proof of Theorem 3.2

Consider a 2AFC task where the SNR at the  $i^{\text{th}}$  trial is the  $i^{\text{th}}$  element in an infinite sequence  $\{\eta_i\}_{i=1}^{\infty}$  consistent with the presumed uncertainty set of Eq (14). Given the presumed level of uncertainty  $\alpha_p$  the lowest SNR consistent with that uncertainty set is given by  $\eta_w = \max(0, (1 - \alpha)\tilde{\eta})$ .

Let  $p(\eta: \vartheta)$  be the probability of error in a 2AFC with a fixed SNR  $\eta$  using the threshold-to-noise ratio  $\vartheta$ . Noting that  $\eta\vartheta = \sqrt{\eta}\vartheta$ , it follows from Eq (2), that:

$$p(\eta: \vartheta) = \frac{1}{1 + \exp(2\sqrt{\eta}\vartheta)}. \quad (\text{B.13})$$

When the SNR varies from trial-to-trial according to a specific sequence  $\{\eta_i\}_{i=1}^{\infty}$  the probability of error in a 2AFC experiment of fixed duration is a random variable  $P_{\text{err}}$ , which depends on the number of trials in the experiment. Let  $P(n)$  denote the probability of having  $n$  trials in the fixed duration of the experiment, the mean  $P_{\text{err}}$  (see Section A.4.3 of Bogacz et al. (2006), and Gold and Shadlen (2002)) is:

$$\langle P_{\text{err}}(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle = \sum_{n=0}^{\infty} \frac{P(n)}{n} \left( \sum_{i=1}^n p(\eta_i: \vartheta) \right). \quad (\text{B.14})$$

Since  $p(\eta: \vartheta)$  given in Eq. (B.13) is a decreasing function of  $\eta$ , the mean  $P_{\text{err}}$  can only increase by replacing each  $\eta_i$  with  $\eta_w$ :

$$\langle P_{\text{err}}(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle \leq \sum_{n=0}^{\infty} \frac{P(n)}{n} \left( \sum_{i=1}^n p(\eta_w: \vartheta) \right) = p(\eta_w: \vartheta), \quad (\text{B.15})$$

which is the probability of error with a fixed SNR  $\eta_w$ .

Similarly, let  $\langle DT(\eta: \vartheta) \rangle$  be the mean response time in a 2AFC with a fixed SNR  $\eta$  using the threshold-to-noise ratio  $\vartheta$ . Noting that  $\theta = \vartheta / \sqrt{\eta}$ , it follows from Eq (2), that:

$$\langle DT(\eta: \vartheta) \rangle = \frac{\vartheta}{\sqrt{\eta}} \frac{\exp(2\sqrt{\eta}\vartheta) - 1}{\exp(2\sqrt{\eta}\vartheta) + 1}, \quad (\text{B.16})$$

which is a decreasing function of  $\eta$  (as proved in Lemma B.1 at the end of this Appendix).

The mean response time with the sequence of SNRs  $\{\eta_i\}_{i=1}^{\infty}$  is (Gold and Shadlen, 2002):

$$\langle DT(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle = \sum_{n=0}^{\infty} \frac{P(n)}{n} \left( \sum_{i=1}^n \langle DT(\eta_i: \vartheta) \rangle \right). \quad (\text{B.17})$$



Since  $\langle DT(\eta_i: \vartheta) \rangle$  given by Eq. (B.16) is a decreasing function of  $\eta$ , the mean response time with a sequence of SNRs  $\langle DT(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle$  can only increase by replacing each  $\eta_i$  with  $\eta_w$ :

$$\langle DT(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle \leq \sum_{n=0}^{\infty} \frac{P(n)}{n} \left( \sum_{i=1}^n \langle DT(\eta_w: \vartheta) \rangle \right) = \langle DT(\eta_w: \vartheta) \rangle, \quad (\text{B.18})$$

which is the mean response time with a fixed SNR  $\eta_w$ .

The  $RR$ , given by the ratio between the probability of correct response and the average time between responses:

$$RR(\{\eta_i\}_{i=1}^{\infty}: \vartheta) = \frac{1 - \langle P_{\text{err}}(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle}{\langle DT(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle + D + (D - D_{\text{tot}}) \langle P_{\text{err}}(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle}, \quad (\text{B.19})$$

is a decreasing function of both  $\langle P_{\text{err}}(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle$  and  $\langle DT(\{\eta_i\}_{i=1}^{\infty}: \vartheta) \rangle$ . According to Eqs (B.15) and (B.18), both of these terms are maximized simultaneously when the SNR is fixed at its worst level,  $\eta_w$ . Hence, when the SNR may vary from trial-to-trial according to a sequence  $\{\eta_i\}_{i=1}^{\infty}$  consistent with the presumed uncertainty set of Eq (14), the worst performance is obtained when  $\eta_i = \eta_w$ . This performance is the same as the worst performance for a fixed SNR within the presumed uncertainty set of Eq. (13), for the same level of uncertainty  $\alpha_P$ .

### Lemma B.1

The function  $\langle DT(\eta|\vartheta) \rangle = \frac{\vartheta \exp(2\sqrt{\eta}\vartheta) - 1}{\sqrt{\eta} \exp(2\sqrt{\eta}\vartheta) + 1}$  is a decreasing function of  $\eta$

### Proof

The derivative of  $\langle DT(\eta|\vartheta) \rangle$  is given by:

$$\begin{aligned} \frac{\partial \langle DT(\eta|\vartheta) \rangle}{\partial \eta} &= \vartheta \left[ \frac{2\vartheta \exp(2\sqrt{\eta}\vartheta) \sqrt{\eta} (\exp(2\sqrt{\eta}\vartheta) + 1) - (\exp(2\sqrt{\eta}\vartheta) - 1)(\exp(2\sqrt{\eta}\vartheta) + 1 + 2\vartheta \exp(2\sqrt{\eta}\vartheta))}{\eta (\exp(2\sqrt{\eta}\vartheta) + 1)^2} \right] \\ &= \vartheta \left[ \frac{-\exp(4\sqrt{\eta}\vartheta) + 4\sqrt{\eta}\vartheta \exp(2\sqrt{\eta}\vartheta) + 1}{\eta (\exp(2\sqrt{\eta}\vartheta) + 1)^2} \right] \end{aligned} \quad (\text{B.20})$$

Using a Taylor series to expand the exponent and noting that for  $n > 3$ ,  $2n < 2^n$ , we may conclude that  $2y \exp(y) + 1 < \exp(2y)$ . Substituting  $y = 2\sqrt{\eta}\vartheta$  indicates that the above derivative is negative, thereby completing the proof.

## B.6 Robustness to uncertainties in delays

This section defines and derives the robustness of a 2AFC task for achieving the required reward rate under uncertainties in the delays.

### Definition B.3

Robustness of a 2AFC task to uncertainties in delays: *Consider a 2AFC task with uncertainties in the delays specified by an info-gap model. Given a threshold  $\theta$ , the robustness  $\hat{\alpha}$  is the greatest level of uncertainty in the info-gap model  $U(\alpha, \tilde{D}, \tilde{D}_{\text{tot}})$  for which the required reward rate  $R_r = RR(\theta: \eta, \tilde{D}, \tilde{D}_{\text{tot}})$  can be guaranteed:*

$$\widehat{\alpha}(R_r, \theta; \eta, \tilde{D}, \tilde{D}_{tot}) = \max \left\{ \alpha \left| \left( \min_{D, D_{tot} \in U(\alpha, \tilde{D}, \tilde{D}_{tot})} RR(\theta, \eta, D, D_{tot}) \right) \geq R_r \right. \right\}. \quad (\text{B.21})$$

The upper limit on the required reward rate,  $RR(\theta; \eta, \tilde{D}, \tilde{D}_{tot})$ , is the nominal reward rate that can be achieved using the threshold  $\theta$  with the nominal delays,  $\tilde{D}$  and  $\tilde{D}_{tot}$ , and the SNR,  $\eta$ , i.e., Eq. (5) with the nominal delays. The robustness for achieving a better than nominal reward rate is zero.

The robustness depends on the structure of the info-gap model. Given the specific info-gap model of Eq. (15), the robustness against uncertainties in the delays can be derived as specified in the next Theorem.

### Theorem B.3

Robustness of a 2AFC task to uncertainties in the delays: *Consider a 2AFC task with uncertainties in the delays specified by the info-gap model of Eq. (15). Given the nominal delays,  $\tilde{D}$  and  $\tilde{D}_{tot}$ , and the SNR,  $\eta$ , the robustness for achieving the required performance  $R_r$  with the threshold  $\theta$  is given by:*

$$\widehat{\alpha}(R_r, \theta; \eta, \tilde{D}, \tilde{D}_{tot}) = \max \left\{ 0, \frac{1/R_r - \theta(1 - \exp(-2\eta\theta))}{\tilde{D} + \tilde{D}_{tot} \exp(-2\eta\theta)} - 1 \right\}. \quad (\text{B.22})$$

### Proof

The internal minimization in Eq. (B.21) is derived in Appendix B.1 for the presumed level of uncertainty, i.e., for  $\alpha = \alpha_p$  and expressed in Eq. (B.1). For a general level of uncertainty  $\alpha$ , Eq. (B.1) implies that the internal inequality in Eq. (B.21) can be expressed as:

$$\theta + \tilde{D}(1 + \alpha) + [\tilde{D}_{tot}(1 + \alpha) - \theta] \exp(-2\eta\theta) \leq R_r^{-1} \quad (\text{B.23})$$

The robustness  $\hat{\alpha}$  is the level of uncertainty  $\alpha$  for which equality holds, and thus given by Eq. (B.22).

Eqs. (5) and (B.22) imply that the robustness for achieving the nominal performance is zero, i.e.,  $\hat{\alpha}(R_r = RR(\theta; \eta, \tilde{D}, \tilde{D}_{tot}), \theta; \eta, \tilde{D}, \tilde{D}_{tot}) = 0$ . In particular, the optimal nominal performance  $RR(\theta_{op}; \eta, \tilde{D}, \tilde{D}_{tot})$  (i.e., Eq. (4) for the optimal threshold under nominal delays) has zero robustness. Only lower reward rates  $R_r < RR(\theta; \eta, \tilde{D}, \tilde{D}_{tot}) = RR(\theta_{op}; \eta, \tilde{D}, \tilde{D}_{tot})$ , can be attained robustly. Robustness to uncertainties requires relinquishing performance.

## B.7 Robust satisficing under uncertainties in delays

This section specifies the robust-satisficing strategy and derives the resulting performance curves under uncertainties in the delays.

Eq. (B.21) implies that sub-optimal reward rates have positive robustness to uncertainties in the delays. The robust-satisficing strategy seeks the threshold  $\theta_{RS}$  that provides maximum robustness for required reward rates that are sub-optimal, as defined below:

**Definition B.4**

Robust-satisficing threshold under uncertainties in the delays: *Consider a 2AFC task with SNR,  $\eta$ , and nominal delays,  $\tilde{D}$  and  $\tilde{D}_{tot}$ . Given the required sub-optimal performance,  $R_r$ , the robust-satisficing threshold,  $\theta_{RS}$ , maximizes the robustness for achieving  $R_r$  under uncertainties in delays:*

$$\theta_{RS}(R_r; \eta, \tilde{D}, \tilde{D}_{tot}) = \arg \max_{\theta} \hat{\alpha}(R_r, \theta; \eta, \tilde{D}, \tilde{D}_{tot}). \quad (\text{B.24})$$

Robust-satisficing under uncertainties in the delays imposes a strict condition on the threshold  $\theta_{RS}$  as stated in the next Theorem.

**Theorem B.4**

Robust-satisficing threshold under uncertainties in the delays: *Consider a 2AFC task with uncertainties in the delays specified by the info-gap model of Eq. (15). Given the nominal delays,  $\tilde{D}$  and  $\tilde{D}_{tot}$ , the SNR,  $\eta$ , and the sub-optimal required reward rate  $R_r < RR_{op}(\eta; \tilde{D}, \tilde{D}_{tot})$ , the robust satisficing threshold  $\theta_{RS}$  satisfies:*

$$2\eta\tilde{D}_{tot}R_r^{-1} = [2\eta\theta_{RS} + \exp(2\eta\theta_{RS}) - 1]\tilde{D} + [2\eta\theta_{RS} - \exp(-2\eta\theta_{RS}) + 1]\tilde{D}_{tot}. \quad (\text{B.25})$$

**Proof**

For a fixed level of required performance  $R_r$ , consider the interval of threshold values for which the nominal reward rate is higher than required, i.e., threshold values for which  $R_r < RR(\theta, \eta; \tilde{D}, \tilde{D}_{tot})$ . This is a single (connected) interval since  $RR$  as a function of the threshold has a single global maximum, as noted following Eq. (6). Within this interval the robustness is positive and given by the second term in Eq. B.22. The boundaries of the interval are defined by  $R_r = RR(\theta, \eta; \tilde{D}, \tilde{D}_{tot})$  - at which the robustness is zero; or by  $\theta = 0$  at the lower boundary. In the later case, the derivative of the robustness with respect to the threshold at  $\theta = 0$  is positive. Since the robustness within the interval is higher than at the boundaries, there must be at least one maximum within this interval. The derivative of the right term in Eq. B.22 with respect to  $\theta$  can be expressed as

$$\frac{\partial \hat{\alpha}(R_r, \theta; \eta, \tilde{D}, \tilde{D}_{tot})}{\partial \theta} = \frac{g(\theta, \eta; R_r, \tilde{D}, \tilde{D}_{tot}) \exp(-2\eta\theta)}{(\tilde{D} + \tilde{D}_{tot} \exp(-2\eta\theta))^2}, \quad \text{where} \quad (\text{B.26})$$

$$g(\theta, \eta; R_r, \tilde{D}, \tilde{D}_{tot}) = \frac{2\eta\tilde{D}_{tot}}{R_r} - [2\eta\theta + \exp(2\eta\theta) - 1]\tilde{D} - [2\eta\theta - \exp(-2\eta\theta) + 1]\tilde{D}_{tot}. \quad (\text{B.27})$$

Hence, extrema of  $\hat{\alpha}(R_r, \theta; \eta, \tilde{D}, \tilde{D}_{tot})$  as a function of  $\theta$  satisfy  $g(\theta, \eta; R_r, \tilde{D}, \tilde{D}_{tot}) = 0$ . It is straightforward to verify that when the latter equality holds, the second derivative of the robustness is negative. Thus  $g(\theta, \eta; R_r, \tilde{D}, \tilde{D}_{tot}) = 0$  defines a unique maximum and, according to Eq. (B.27), the robust-satisficing threshold that satisfies this equality must obey Eq. (B.25).

Substituting Eq. (3) in Eq. (B.25), specifies the *robust-satisficing performance curve* (RSPC) that describes the tradeoff between speed ( $\langle DT \rangle / \tilde{D}_{tot}$ ) and accuracy ( $1 - p(\text{err})$ ) that must hold in order to robust-satisfice the required performance  $R_r$ :

$$\frac{\langle DT \rangle}{\tilde{D}_{tot}} = \frac{R_r^{-1}}{\tilde{D}_{tot}} \left[ \frac{\tilde{D}/\tilde{D}_{tot} + p(\text{err})(1 - \tilde{D}/\tilde{D}_{tot})}{p(\text{err})(1 - p(\text{err})) \log \left( \frac{1 - p(\text{err})}{p(\text{err})} \right)} + \frac{\tilde{D}/\tilde{D}_{tot} + 1}{1 - 2p(\text{err})} \right]^{-1}. \quad (\text{B.28})$$

Note that this expression does not reduce to a scaled OPC (7) even when the actual delays equal the nominal values ( $D = \tilde{D}$  and  $D_{tot} = \tilde{D}_{tot}$ ), or when there is no penalty delay ( $\tilde{D}_{tot} = \tilde{D}$ ). The RSPCs have a different shape.

The RSPCs depend on the ratio  $\tilde{D}/\tilde{D}_{tot}$ , so for comparison with experimental data, this ratio must be determined or related to the actual ratio  $D/D_{tot}$ . When no penalty delay is used,  $D_{pen} = 0$ , and the ratio is one. The corresponding RSPCs are specified by Eq. (B.28) with  $\tilde{D}/\tilde{D}_{tot} = 1$ :

We note that the constraint on the required performance  $R_r < RR_{op}(\eta : \tilde{D}, \tilde{D}_{tot})$  implies that the RSPCs lie above the OPC. However, it is possible to extend Eqs. (B.25) and (B.28) below the OPC, as briefly outlined here. While better than optimal reward rates cannot be guaranteed, they may nevertheless occur under favorable conditions, which are made increasingly possible as the level of uncertainty increases. The minimum level of uncertainty which makes it possible to get a desired reward rate is defined as the opportuneness, and the opportunity facilitating strategy selects the threshold with minimum opportuneness for the desired reward rate. The resulting performance curves follow the above equations below the OPC. Since this region is only marginally important for the current paper (the RSPCs in Fig. 6 are mostly above the OPC) the opportunity facilitating strategy is not developed further in this paper, (but see Ben-Haim (2006) and Zacksenhouse et al. (2009) for more details).

## C The likelihood ratio test and the influence of subgroup sizes

Here we provide further details on the reliability of the statistics used in §4. We first address the assumption that the data are normally distributed. For each subgroup of subjects and each  $p(\text{err})$  bin, we tested the distribution of normalized decision times for Gaussianity using the Jarque-Bera Judge et al. (1988) and Lilliefors Conover (1980) tests, obtaining the following numbers of significantly non-Gaussian  $p(\text{err})$  bins for both tests, at a significance level of 0.05. Among the top 30% and 30–60% groups: 3 bins in each group; among the 60–90% group: 2 bins; among the bottom 10% group no significantly non-Gaussian bins were found. When the 30–60% and 60–90% groups are combined into a single group as in §4, the number of significantly non-Gaussian  $p(\text{err})$  bins rises to 6, probably due to the fact that some subjects in this large group employ different decision strategies, a point that we address further below. (There are 250 data points in this middle group, compared with 125 in the top 30% and 42 in the bottom 10% groups, cf. Table 2.)

When comparing two nested models, i.e., a complex model and a simple model to which the complex one can be reduced, the likelihood ratio test can be used to determine whether the data is significantly less likely under the simple rather than the complex model. This is the case when comparing the MMPC with uncertainty in  $D$  with the OPC for  $RR$ , to which the former reduces by fixing  $\alpha = 0$ . The likelihood ratio test shows that the MMPC with uncertainties in  $D$  provides a significantly better description of the data for the top 30% of subjects ( $p < 0.01$ ), and for both the other groups ( $p \approx 0$ ). Except for the OPC for  $RR$ , all the other PCs are based on non-nested models with a single free parameter each. The corresponding likelihood ratio between the MMPC with uncertainties in  $D$  and each of these models (first four rows Table 4) provides a reasonable criterion for assessing whether the former explains the data better. A more rigorous comparison using the Bayesian approach

would require additional assumptions on prior probabilities of the parameters  $q$  and  $\alpha$ : see, e.g. MacKay (2003).

As noted in §4, in fitting performance curves to data from subgroups of subjects divided according to total rewards accrued, we implicitly assume that all members of each given subgroup employ a common decision strategy. Here we probe the validity of this assumption by considering six additional ways of splitting the data. We first fitted all performance curves to the data from each subject separately and computed the average likelihood ratio for each pair of curves, as a product of likelihood ratios for the 80 individuals, obtaining the entries in the first column of Table C.1. This split avoids the assumption of a common strategy, but the ratios are unreliable because individual subject data are very noisy and some bins contain as few as 3 data points, leading to overfitting. This is reflected in the fact that the entries of column 1 differ substantially from those of the other columns in Table C. 1.

**Table C.1**

Likelihood ratios of the data given MMPC with uncertainty in  $D$  and given other performance curves (rows), for different splits of subjects into groups (columns). The higher the likelihood ratio, the less likely is it that the data could be generated by the corresponding model, in comparison to the MMPC for  $D$ . See Table 1 for abbreviations.

Performance curve	Individual	20 groups	10 groups	5 groups	4 groups	2 groups
MMPC for SNR	$> 10^{13}$	$> 10^{10}$	$> 10^9$	$> 10^{10}$	$> 10^{10}$	$> 10^8$
RSPC for $D$	1.96	42.00	29.26	8.28	7.86	9.61
OPC for $RR_m$	471.54	$> 10^5$	$> 10^5$	522.09	316.03	59.16
OPC for $RA$	$> 10^6$	482.05	70.52	3.33	8.82	3.96
OPC for $RR$	$> 10^{59}$	$> 10^{33}$	$> 10^{29}$	$> 10^{26}$	$> 10^{24}$	$> 10^{20}$

This analysis was repeated five times to produce the remaining columns of Table C.1, by successively dividing the subjects (sorted by total rewards accrued) into 20, 10, 5, 4, and 2 groups, with equal numbers in each group (4, 8, 16, 20, and 40 respectively). Although the precise values of the ratios depend on the number of groups, the relative ranking remains almost completely consistent. All ratios in Table C.1 exceed 1, implying that the data are most likely under MMPC with uncertainty in  $D$  in all cases. The next best fits are provided by RSPC with uncertainty in  $D$  and OPC for  $RA$ , their order depending on the split. Somewhat further behind comes OPC for  $RR_m$ , trailed by MMPC with uncertainty in SNR, and finally the parameter-free OPC for  $RR$ .

## D Comparison of pure and extended DD models for experiment 1

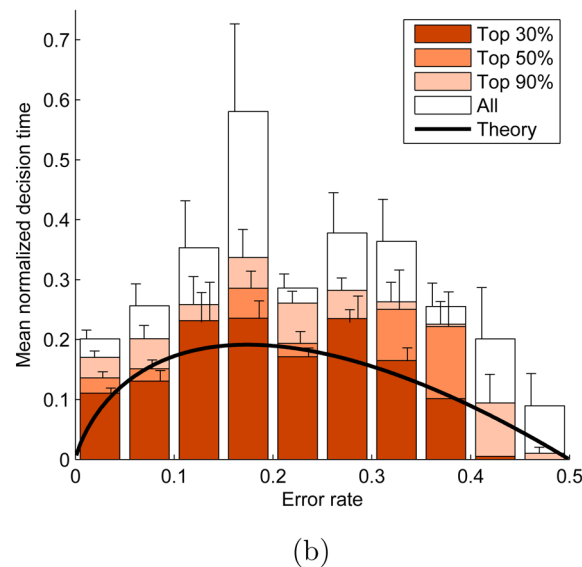
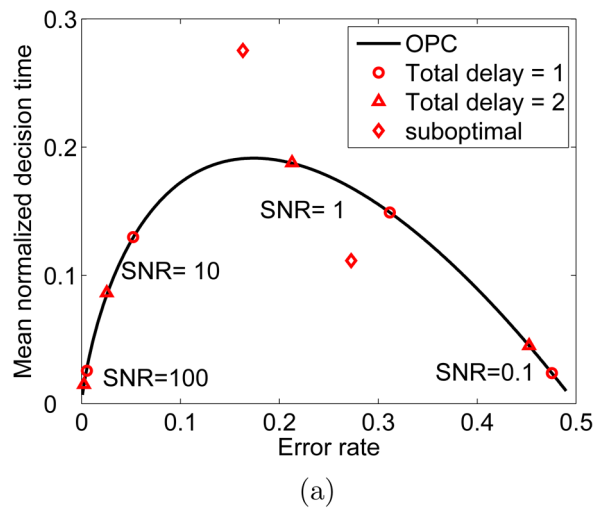
As explained in Bogacz et al. (2009) only the first of the two experiments described in this paper yielded sufficient data for fits to the extended DD model to be feasible. Comparison of extended DD parameters obtained by such data fits, averaged over the 20 subjects, to fits of the same data to the pure DD process reveals the following. (We identify the source of the parameters by the parenthetical notes ext. and pure respectively.)

Mean values of the threshold-to-drift ratios  $\theta$  (ext.) are approximately half  $\theta$  (pure), mean values of SNR  $\eta$  (ext.) are approximately triple  $\eta$  (pure), while mean values of the nondecision time  $T_0$  (ext.) are very close to  $T_0$  (pure). In all cases the extended and pure parameter values are strongly correlated across subjects ( $\theta$ :  $r = 0.61$ ,  $p = 0.004$ ,  $\eta$ :  $r = 0.92$ ,  $p < 10^{-5}$  and  $T_0$ :  $r = 0.67$ ,  $p = 0.001$  (Bogacz et al., 2009, Fig. 4)), and the resulting optimal threshold-to-drift ratios are also strongly correlated ( $r = 0.71$ ,  $p < 10^{-5}$ ). This leads to the

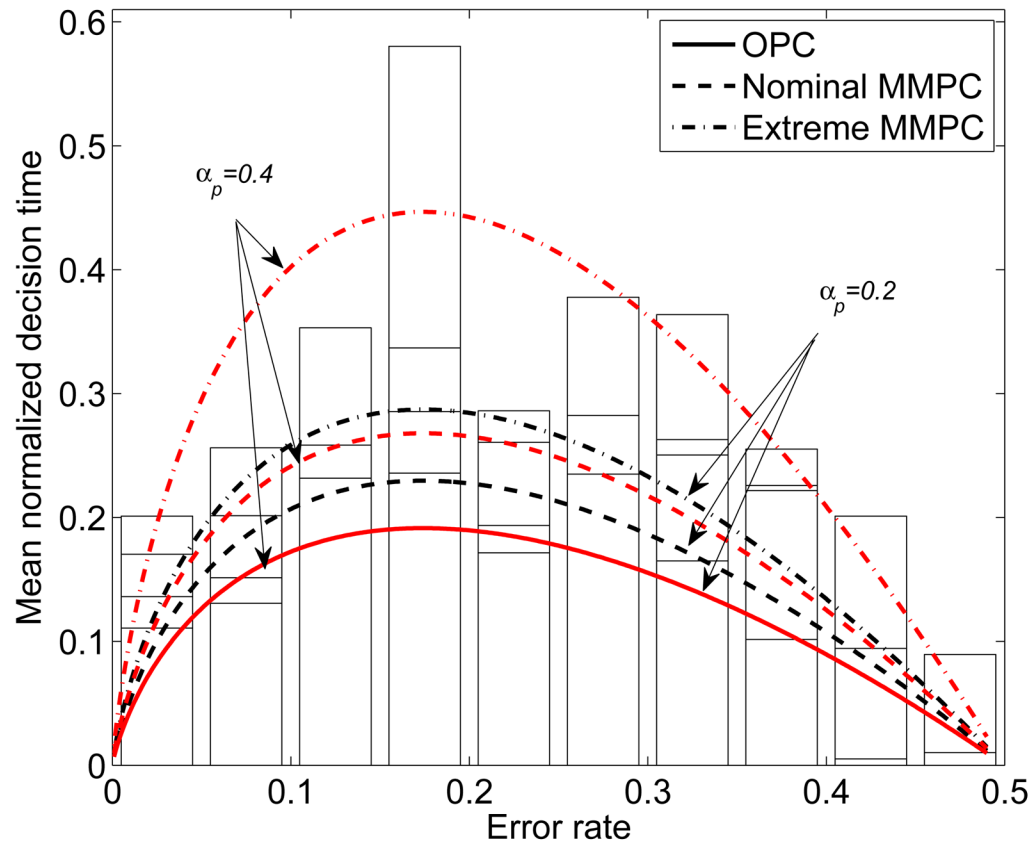
fact that overall correlations between subjects' thresholds and the optimal thresholds obtained from the pure and extended DD fits do not substantially differ ( $r_1 = 0.44$  ( $p < 10^{-5}$ ) for pure, and  $r_1 = 0.62$  ( $p < 10^{-5}$ ) for extended.) Moreover, comparisons of the mean thresholds (averaged over these 20 subjects) with the optimal thresholds computed analytically for the pure, and numerically for the extended DD processes, for all four delay conditions, reveal very similar patterns. See Bogacz et al. (2009, Fig. 5).

We believe that this is because the substantial variances in SNR and threshold-to-drift ratio for the extended DD model are compensated by higher noise variance  $\sigma^2$  in the pure DD fits (i.e., all the sources of variance are lumped in a lower SNR  $\eta(\text{pure})$ ). To further gauge the effects of drift and initial condition variance we compared mean decision times and error rates for the extended DD fits with the corresponding quantities evaluated for a pure DD process with  $\eta$  and  $\theta$  equal to the mean values of the extended DD parameters. We found that  $\langle DT \rangle$ s for the two processes were very similar ( $r = 0.94$ ,  $p < 10^{-5}$ ), but that  $p(\text{err})$ s were significantly lower for the pure DD process (paired t-test  $p < 10^{-5}$ ), presumably due to the unrealistically high SNRs that result when the other sources of variance are removed from the extended DD model.



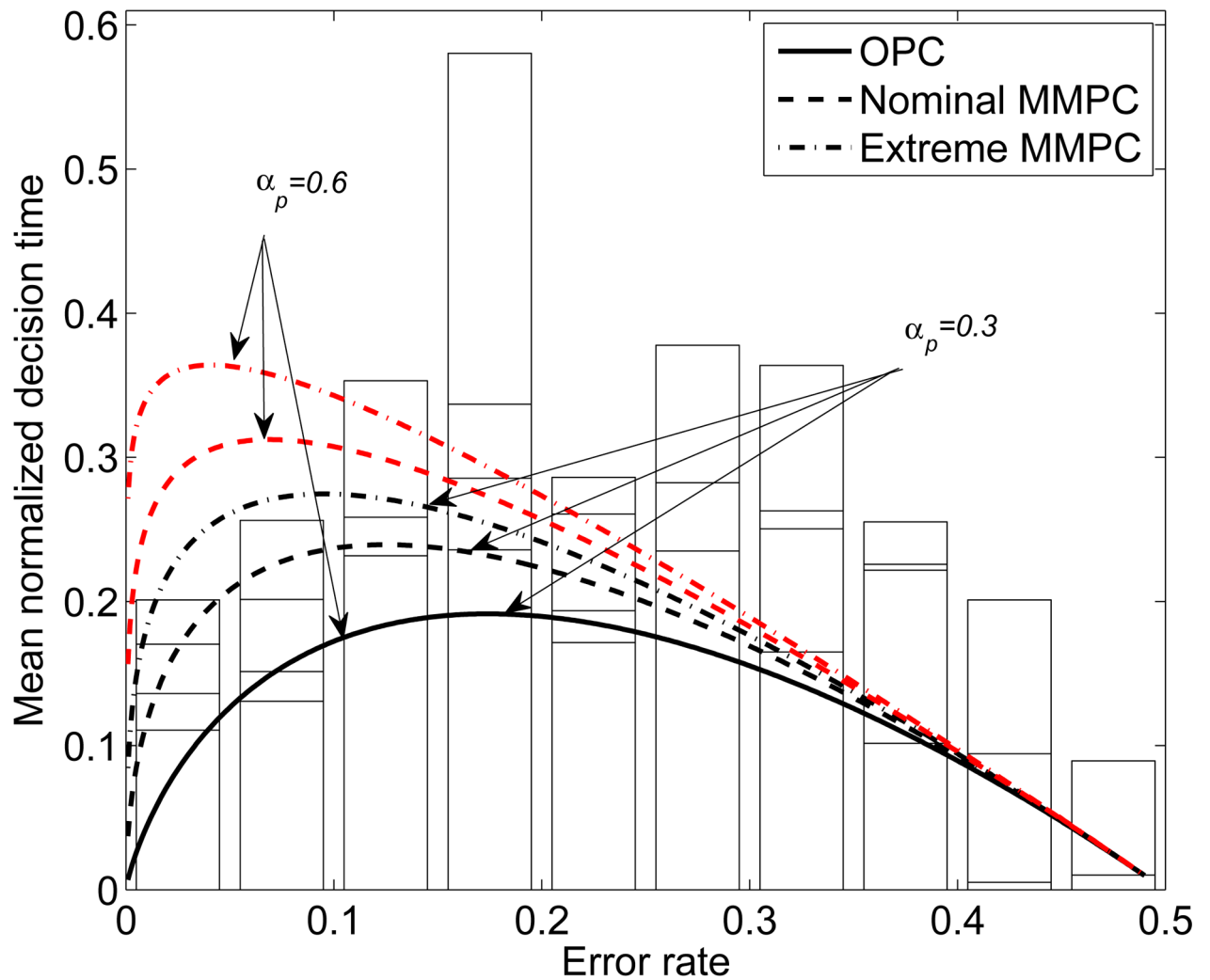
**Fig. 1.**

(a) Optimal performance curve of Eq. (7) relating mean normalized decision time to error-rate under varying task conditions. Triangles and circles mark the performance under the specified task conditions. Moving leftwards, the resulting RRs increase with SNR from 0.51 to 0.60, 0.84 and 0.97 for  $D_{tot} = 1$ , and from 0.26 to 0.33, 0.45 and 0.49 for  $D_{tot} = 2$ . Diamonds mark suboptimal performance points resulting from thresholds 25% above and below the optimal threshold for SNR=1 and  $D_{tot} = 2$ ; both yield  $\approx 1.3\%$  degradation in the RR. (b) Thick curve shows the optimal performance curve of Eq. (7), and histograms show data collected from 80 human subjects, sorted according to total rewards accrued over multiple blocks of trials. Refer to the text for experimental conditions. White bars: all subjects; lightest bars: lowest 10% excluded; medium bars: lowest 50% excluded; darkest bars: lowest 70% excluded. Vertical line segments indicate standard errors. From Holmes et al. (2005, Fig. 1).

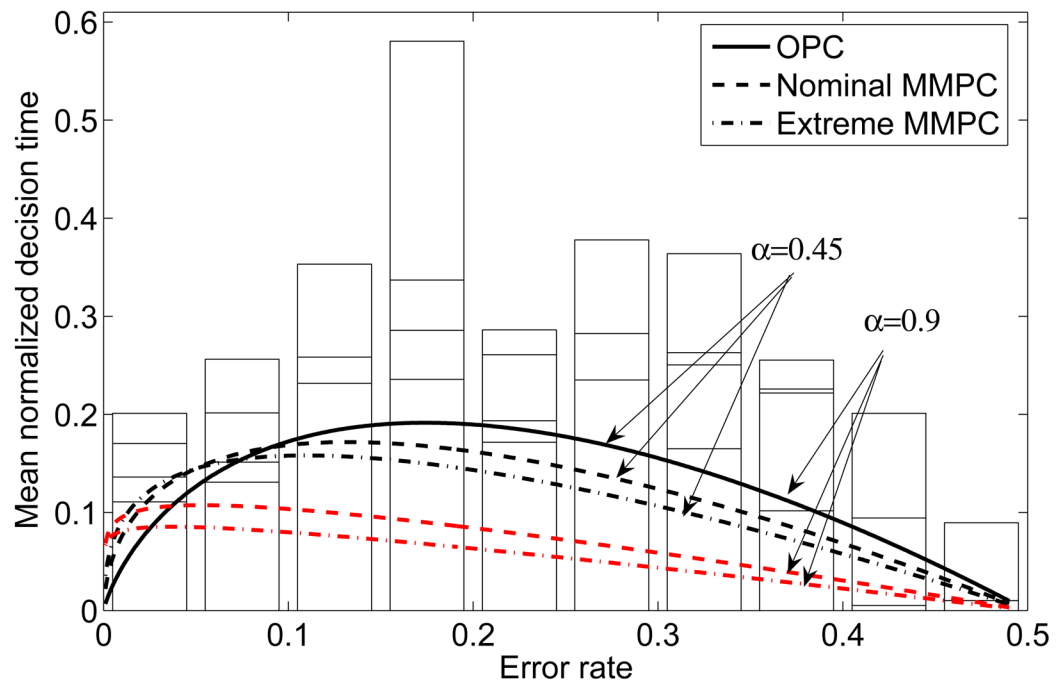


**Fig. 2.**

Maximin performance curves (MMPCs) and performance bands for two presumed levels of uncertainty  $\alpha_p = 0.2, 0.4$  in delays. Nominal MMPCs for  $D_{tot} = \tilde{D}_{tot}$  (dashed), and extreme MMPCs for  $D_{tot} = \tilde{D}_{tot}(1 - \alpha_p)$  (dash-dotted) and for  $D_{tot} = \tilde{D}_{tot}(1 + \alpha_p)$  (solid) are shown; the latter coincide with the OPC for all  $\alpha_p$  (cf. Eq. (12) with  $\gamma = (1 + \alpha_p)^{-1}$  and Eq. (7)). Performance bands consistent with the presumed level of uncertainty are bounded by the extreme MMPC and the OPC.

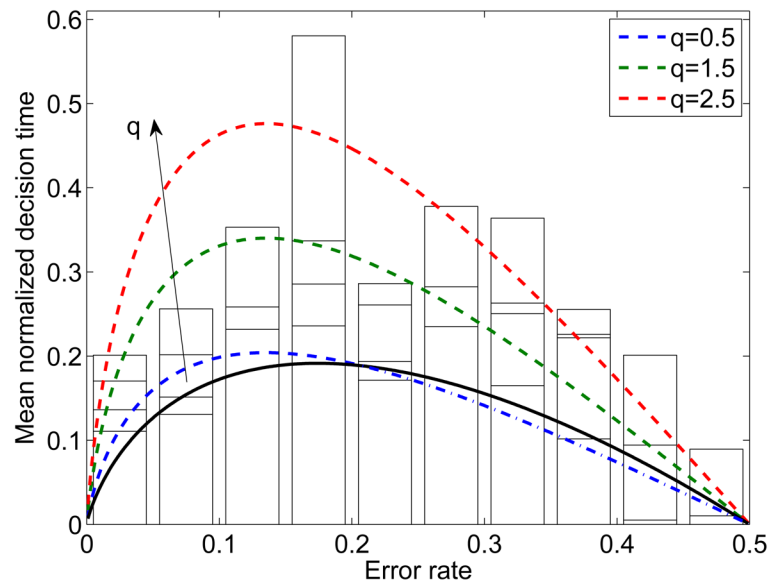


**Fig. 3.** MMPCs for two presumed levels  $\alpha_p = 0.3, 0.6$  of uncertainties in SNR arising from uncertain noise variance. Nominal MMPC for  $\eta = \tilde{\eta}$  (dashed), and extreme MMPC for  $\eta = \tilde{\eta}(1 + \alpha_p)$  (dash-dotted) and for  $\eta = \tilde{\eta}(1 - \alpha_p)$  (solid); the latter coincide with the OPC for all  $\alpha_p$ . Performance bands for SNR within the range consistent with the presumed uncertainty are bounded by the extreme MMPC and the OPC.

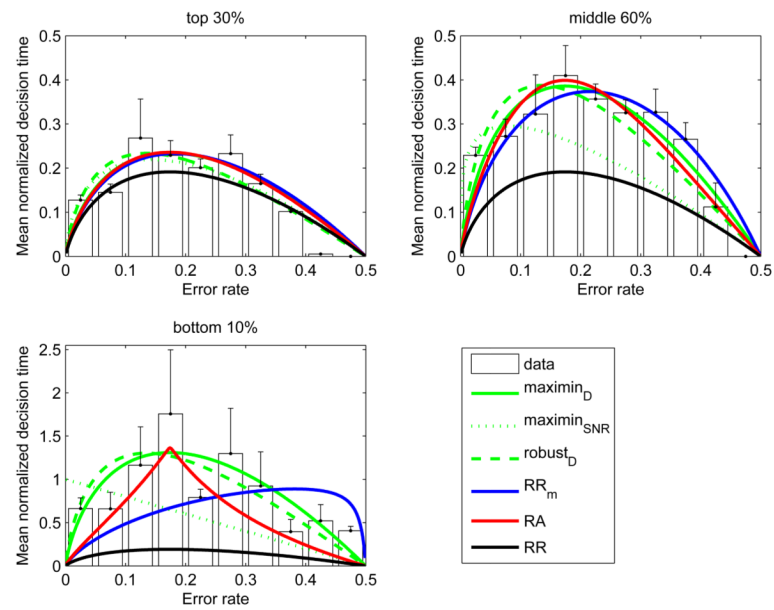


**Fig. 4.**

MMPCs for two presumed levels  $\alpha_p = 0.45, 0.9$  of uncertainty in SNR arising from uncertain drift rate. Nominal MMPC for  $\eta = \tilde{\eta}$  (dashed), and extreme MMPCs for  $\eta = \tilde{\eta}(1 + \alpha_p)$  (dash-dotted) and for  $\eta = \tilde{\eta}(1 - \alpha_p)$  (solid); the latter coincides with the OPC for all  $\alpha_p$ . Performance bands for SNR within the range consistent with the presumed uncertainty are bounded by the extreme MMPC and the OPC.

**Fig. 5.**

Robust satisficing performance curves (RSPCs, dashed) for uncertainties in delays given three normalized required performance levels  $q = (R_r^{-1} - D) / D_{tot}$ . Here the nominal delay ratio  $\tilde{D} / \tilde{D}_{tot} = 1$  is assumed, corresponding to the case of zero penalty delay. The OPC (solid) bounds the RSPCs below, but extensions of RSPCs under the OPC are also relevant (cf. dash-dotted line for  $q = 0.5$ ) and see Appendix B.7).

**Fig. 6.**

Comparisons of performance curves with mean normalized decision times from experimental blocks with  $D_{pen} = 0$  for three groups of subjects sorted by total rewards acquired. Error bars show standard error of mean normalized decision times. Note the difference in scale of the vertical axes in upper and lower panels. Here and throughout this section,  $\text{maximin}_{\text{SNR}}$  refers to the MMPC due to uncertainty in noise variance.



**Table 1**

List of abbreviations.

Abbreviation	Meaning
2AFC	2 alternative forced choice
COBRA	Competition between reward and accuracy
DD	Drift-diffusion
<i>DT</i>	Decision time
MMPC	Maximin performance curve
OPC	Optimal performance curve
<i>RA</i>	Reward accuracy rate
<i>RR</i>	Reward rate
<i>RR<sub>m</sub></i>	Modified reward rate
RSI	Response-to-stimulus interval
RSPC	Robust-satisficing performance curve
<i>RT</i>	Reaction time
SAT	Speed-accuracy tradeoff
SNR	Signal-to-noise ratio
SPRT	Sequential probability ratio test

**Table 2**

Numbers of experimental conditions falling into each error rate bin for the three panels of Fig. 6.

Error rate range	Top 30%	Middle 60%	Bottom 10%
0 – 5%	37	82	9
5 – 10%	17	32	7
10 – 15%	12	12	3
15 – 20%	12	23	6
20 – 25%	19	35	2
25 – 30%	15	39	5
30 – 35%	7	13	4
35 – 40%	1	6	2
40 – 45%	1	5	2
45 – 50%	4	3	2

**Table 3**

Values of performance curve parameters ( $q$  or  $\alpha$ ) estimated using the maximum likelihood method from the data from the three groups of subjects sorted by total earning. For the MMPC with uncertainty in SNR and bottom 10%, the higher the value of  $\alpha$  the higher the likelihood of the data. See Table 1 for abbreviations.

Performance curve	Top 30%	Middle 60%	Bottom 10%
MMPC for $D$	0.22	1.02	5.84
MMPC for SNR	0.19	0.54	$\infty$
RSPC for $D$	0.72	1.85	8.57
OPC for $RR_m$	0.14	0.49	0.98
OPC for $RA$	0.15	0.55	1.096

**Table 4**

Likelihood ratios of the data given the MMPC with uncertainty in  $D$  compared to other performance curves (rows), for different groups of subjects (columns). The last column is the product of the first three and thus expresses likelihood ratios for all data. The higher the likelihood ratio, the less likely is it that the data could be generated by the corresponding model, in comparison to the MMPC with uncertainty in  $D$ . To account for the fact that all curves have one free parameter except the OPC, the bottom row includes in brackets the ratio of the likelihood of the data given the two curves divided by  $\exp(1)$ , as suggested by Akaike (1981); see Appendix C. See Table 1 for abbreviations.

Performance curve	Top 30%	Middle 60%	Bottom 10%	Product
MMPC for SNR	1.13	$> 10^8$	308.38	$> 10^{10}$
RSPC for $D$	1.99	4.04	1.72	13.82
OPC for $RR_m$	1.28	38.79	208.47	$> 10^4$
OPC for $RA$	1.11	8.11	4.83	43.48
OPC for $RR$	47.34 [17.42]	$> 10^{21}$	$> 10^6$	$> 10^{28}$