Neural Networks 21 (2008) 1247–1260

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

# Integrated Bayesian models of learning and decision making for saccadic eye movements☆

Kay H. Brodersen[a,b,*], Will D. Penny[a], Lee M. Harrison[a], Jean Daunizeau[a], Christian C. Ruff[c], Emrah Duzel[c], Karl J. Friston[a], Klaas E. Stephan[a,d]

[a] Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, UK
[b] Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK
[c] Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK
[d] Branco-Weiss-Laboratory, Institute for Empirical Research in Economics, University of Zurich, Switzerland

ARTICLE INFO

*Article history:*
Received 7 September 2007
Received in revised form
29 August 2008
Accepted 31 August 2008

*Keywords:*
Saccades
Decision making
Reaction time
Bayesian learning
Model comparison

ABSTRACT

The neurophysiology of eye movements has been studied extensively, and several computational models have been proposed for decision-making processes that underlie the generation of eye movements towards a visual stimulus in a situation of uncertainty. One class of models, known as linear rise-to-threshold models, provides an economical, yet broadly applicable, explanation for the observed variability in the latency between the onset of a peripheral visual target and the saccade towards it. So far, however, these models do not account for the dynamics of learning across a sequence of stimuli, and they do not apply to situations in which subjects are exposed to events with conditional probabilities. In this methodological paper, we extend the class of linear rise-to-threshold models to address these limitations. Specifically, we reformulate previous models in terms of a generative, hierarchical model, by combining two separate sub-models that account for the interplay between learning of target locations across trials and the decision-making process within trials. We derive a maximum-likelihood scheme for parameter estimation as well as model comparison on the basis of log likelihood ratios. The utility of the integrated model is demonstrated by applying it to empirical saccade data acquired from three healthy subjects. Model comparison is used (i) to show that eye movements do not only reflect marginal but also conditional probabilities of target locations, and (ii) to reveal subject-specific learning profiles over trials. These individual learning profiles are sufficiently distinct that test samples can be successfully mapped onto the correct subject by a naïve Bayes classifier. Altogether, our approach extends the class of linear rise-to-threshold models of saccadic decision making, overcomes some of their previous limitations, and enables statistical inference both about learning of target locations across trials and the decision-making process within trials.



## 1. Introduction

In order to survive in a competitive, dynamic environment, animals must be able to integrate past experience with sensory evidence to infer the current state of the world and execute a behavioural response. Marked progress in our understanding of the neural basis of decision making has been achieved by focusing on sensory-driven decisions, such as the simple question of where to look next. Studying decision making in sensorimotor systems like the oculomotor system has the advantage that one can exploit a large body of neuroanatomical and neurophysiological knowledge that has been accumulated over the past decades. It seems conceivable that studying the neuronal mechanisms of visual-saccadic decision making could provide us with a blueprint of how the brain implements other sensorimotor decisions, or even deliver "a model for understanding decision making in general" (Glimcher, 2003).

The decision processes that underlie rapid eye movements towards a target have been studied in a variety of experimental paradigms. One seminal series of studies is based on the *random dot-motion* task designed by Newsome and colleagues (Newsome & Pare, 1988). In an initial fixed-duration version of this task, monkeys were trained to discriminate the motion direction of a set of moving dots with varying degrees of coherence, and indicate the perceived motion by a leftward or rightward

☆ This research has been funded by the Wellcome Trust (VS/06/UCL/A18), the German Academic Exchange Service (DAAD, D/06/49008), and the Stiftung Familie Klee (Frankfurt/Main).

* Corresponding author at: Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, UK.
*E-mail address:* kay.brodersen@gmx.net (K.H. Brodersen).

0893-6080/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2008.08.007

saccade (Newsome, 1997; Newsome, Britten, & Movshon, 1989; Newsome, Britten, Salzman, & Movshon, 1990; Salzman, Britten, & Newsome, 1990). Subsequently, Shadlen, Britten, Newsome, and Movshon (1996) suggested a computational explanation of the neuronal mechanisms producing the resulting saccade and provided experimental verification of its key assumptions (Gold & Shadlen, 2000; Kim & Shadlen, 1999; Shadlen et al., 1996; Shadlen & Newsome, 2001). In particular, they identified a gradual rise of spiking activity in the lateral intraparietal (LIP) area integrating motion direction-specific signals from the middle temporal (MT) area (Shadlen & Newsome, 1996, 2001).

Based on a reaction time version of the same task (Roitman & Shadlen, 2002), Shadlen and colleagues advanced the hypothesis that rising activity before a saccade, which had also been observed in the frontal eye fields (FEF), represented the ratio of the log likelihoods that the two possible eye movements would be executed (Gold & Shadlen, 2000, 2001). Based on their decision-theoretic analysis, they suggested that log likelihood ratios might be used as "a natural currency for trading off sensory information, prior probability and expected value to form a perceptual decision" (Gold & Shadlen, 2001).

Another key series of studies was carried out by Hanes, Schall, and colleagues, who investigated an *oddball* task (as well as the *countermanding* paradigm; Hanes and Carpenter (1999)) to study how neural signals in the FEFs would finally trigger the initiation of saccades (Hanes & Schall, 1996; Hanes, Thompson, & Schall, 1995; Schall & Thompson, 1999; Thompson, Bichot, & Schall, 1997; Thompson, Hanes, Bichot, & Schall, 1996). In their oddball task, monkeys were trained to indicate, by an eye movement, the location of the oddball within a circular arrangement of visual stimuli around a central fixation dot. They showed that FEF activity was consistent with psychophysical models about oddball reaction time tasks (Luce, 1986; Ratcliff, 1978; Sternberg, 1969a, 1969b). Specifically, their findings supported the notion that the saccadic decision would be made as soon as gradually increasing neural activity in the FEFs had crossed a biophysical threshold (Hanes, Patterson, & Schall, 1998; Schall & Thompson, 1999).

Motivated by the question of why saccadic latencies displayed large variance in all of the above tasks, an even simpler reaction time paradigm was investigated by Carpenter and colleagues (Carpenter & Williams, 1995; Reddi & Carpenter, 2000). In their *saccade-to-target* reaction time task, human subjects were asked to shift their gaze from a central fixation stimulus to an eccentric target as soon as it appeared on the screen. The critical manipulation was to vary the uncertainty about where the target would appear (Basso & Wurtz, 1997, 1998). It was found that saccade latencies became shorter with increasing prior probability of the corresponding target location. Specifically, response speed was found to be proportional to the log prior probability of target location (Basso & Wurtz, 1997, 1998; Carpenter & Williams, 1995).

The behavioural and electrophysiological findings from all three paradigms described above are consistent with the notion of a saccade being elicited once some gradually rising neuronal activity crosses a biophysical threshold. This idea has been formalized in terms of various mechanisms known as *rise-to-threshold* accumulator models. These models aim to provide a computational abstraction of a biophysically conceivable mechanism that explains saccade latencies and their variability across trials (for reviews see Glimcher (2001, 2003), Gold and Shadlen (2001), Platt (2002), Ratcliff and Smith (2004), Schall (2001, 2003), Smith and Ratcliff (2004) and Usher and McClelland (2001)).

In the context of saccadic decision making with a fixed set of potential target locations, rise-to-threshold models assume that subjects maintain a set of hypotheses each of which corresponds to one such location (Carpenter & Williams, 1995; Gold & Shadlen, 2002; McMillen & Holmes, 2006; Shadlen & Gold, 2004). As the stimulus appears, a measure of evidence for each of these hypotheses is continuously refined, implemented as a competition between alternative decision signals in the brain. At any given point in post-stimulus time, these decision signals might, for example, represent the posterior probabilities of the target hypotheses, as derived from the subject's prior (Basso & Wurtz, 1997, 1998; Platt & Glimcher, 1999) and the sensory evidence (i.e., the likelihood of the data) collected up to that point in time (Carpenter, 2004; Carpenter & Williams, 1995). As soon as one such signal reaches a preset threshold, a saccade is elicited towards the corresponding target. Depending on the way in which information is assumed to be accumulated over time, two specific types of rise-to-threshold model are often distinguished: random-walk models and linear rise-to-threshold models.

*Random-walk* or *diffusion* models are fundamentally based on a sequential probability ratio test that is being carried out continually (Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004; Ratcliff, Zandt, & McKoon, 1999; Wald, 1945). In these models, each new incoming piece of sensory evidence either increases or decreases a single decision variable until it has drifted beyond a threshold associated with the saccadic movement towards a particular target. The decision variable represents the relative evidence for the two alternatives (Ratcliff & Rouder, 1998). However, in the case of a simple saccade-to-target task in a high-contrast setting with highly salient targets, it has been questioned whether a random-walk process for target detection provides a sufficient explanation for the large variability in latencies (Carpenter, 2004; Carpenter & Reddi, 2001; Reddi, 2001).

In *linear* rise-to-threshold models, randomness is introduced as trial-by-trial changes in the otherwise constant rate of rise of the decision signal. This notion has been formalized by Carpenter in a model termed 'LATER' (linear approach to threshold with ergodic rate; Carpenter and Williams (1995), Leach and Carpenter (2001), Reddi, Asrress, and Carpenter (2003)). Like other rise-to-threshold models, LATER proposes that a saccade towards a target is elicited as soon as a neural decision signal has reached a particular threshold. But unlike other rise-to-threshold models (e. g., Grice (1968) and Nazir and Jacobs (1991)), it assumes a fixed threshold and a linear increase whose rate is subject to variation *across* trials, yet fixed *within* a given trial (for a debate on the relationship between the two approaches see Carpenter and Reddi (2001), Ratcliff (2001), Usher and McClelland (2001)). The neurophysiological recordings by Schall and colleagues (Hanes & Schall, 1996; Schall & Thompson, 1999) are consistent with these key assumptions of the LATER model: they had observed that the threshold for saccade release seemed to be constant, whereas the slope of the rise in activity varied considerably across trials (see Fig. 2a).

In their experiments on the saccade-to-target task, Carpenter and colleagues found that the observed saccadic latency was a function of the log probability of the corresponding target location: the more likely the target location, the shorter the latency (Carpenter & Williams, 1995). LATER accounts for this relationship by assuming that the learned a priori target probabilities determine the baseline levels of the decision signals, but not their rates of rise (cf. biased choice theory by Luce (1963)). Carpenter and colleagues used LATER to produce remarkably accurate predictions of human latency distributions in the saccade-to-target task as well as variations of it (Asrress & Carpenter, 2001; Carpenter & Williams, 1995; Leach & Carpenter, 2001; Reddi et al., 2003; Reddi & Carpenter, 2000).

A strength of the LATER model is the straightforward interpretability of its parameters. LATER has thus been used to relate various features of observed latency distributions to the putative underlying neurophysiological process (Anderson, 2008; Asrress & Carpenter, 2001; Carpenter & McDonald, 2007; Kurata & Aizawa, 2004; Leach & Carpenter, 2001; Loon, Hooge, Berg, & den, 2002; Madelain, Champrenaut, & Chauvin, 2007; Reddi et al., 2003; Sinha, Brown, & Carpenter, 2006). Furthermore, various extensions have been proposed, such as arrangements of multiple LATER units in parallel (Carpenter, 2004; Robinson, 1973), mixture models (Nakahara, Nakamura, & Hikosaka, 2006), or the assumption that both the rate of rise and the baseline level of the decision signal are trial-by-trial random variables (Nakahara et al., 2006).

However, the simplicity of this model limits its applicability in three ways. First, linear rise-to-threshold models like LATER have only been applied to saccade-to-target situations in which no learning took place: in previous studies, prior probabilities of target location were always fixed in a given experimental session, and subjects were initially given extensive training until their performance levelled off. During learning, by contrast, the baseline levels of the decision signals are expected to *change* across trials. Even though the notion of variable baseline levels has been discussed before (Glimcher, 2001; Nakahara et al., 2006), no specific model has been put forward how they might evolve dynamically depending on the history of previous trials. Second, LATER only accounts for simple marginal probabilities, where the probability distribution of target locations is described by a single vector of probabilities. It does not account for higher-order contingencies, that is, situations in which the target location probability depends on the target location during the previous trial. Third, within the class of linear rise-to-threshold models, no generative model has been proposed so far that would allow for statistical inference about parameter estimates and for model comparison (e.g., with regard to the type of learning that occurs across trials).

In this study, we propose a more general linear rise-to-threshold model for visual-saccadic decision making that overcomes the restrictions outlined above. First, we explicitly model how subjects' priors are systematically altered by the sequence of stimuli observed so far. This approach makes it possible to investigate how learning dynamically shapes decision making about saccades. Second, our model is able to account for different forms of learning which can be evaluated by model comparison. In particular, this allows us to investigate whether subjects' behaviour is not only driven by *marginal* but also by *conditional* probabilities. Third, based on computational considerations, we propose a specific parameterization of the model. This enables parameter estimation within a maximum likelihood scheme and the subsequent construction of a classifier that can be used to distinguish subjects with different learning profiles.

## 2. Methods

### 2.1. Task

For the present study, subjects were engaged in a sequential reaction time task (SRTT) during which they had to elicit saccades towards a given target in quick succession. The predictability of the target location was modified between blocks to induce varying forms of learning. The degree to which subjects learned the underlying contingency of a particular block was measured by the latencies of their saccades, that is, the time between stimulus onset and the beginning of the saccade towards the stimulus.

The specific setup adopted in this study was based on the saccade-to-target task proposed by Carpenter and Williams (1995). Subjects placed their heads on a chinrest in front of a computer screen in a dark, soundproof booth. At the beginning of a trial, they focused on a red fixation dot (hue 0°, luminance 0.5) at the centre of a black screen. After a random waiting period between 500 and 1500 ms, a second red dot, the target, appeared on the screen, either located at 15° to the left or to the right. Since the original fixation dot remained visible, this design represented an overlap task rather than a gap task (alternative types of waiting-period probability distribution are examined in Oswal, Ogden, and Carpenter (2007)). Subjects were asked to foveate the target as quickly as possible, but not at the cost of errors. After another 700 ms, both dots disappeared, and the screen remained blank for an inter-trial interval of 500 ms.

Based on this design, Carpenter and Williams (1995) investigated the effects of fixed state probabilities for the two target locations on saccadic reaction times. For example, prior to the actual experiment, subjects were trained extensively on a sequence of trials during which the target appeared on the left-hand side with a probability of 70%, and on the right-hand side with a probability of 30%.

In our study, we extended this experimental design in two ways (see Fig. 1). First, each block contained a comparatively small number of trials, and subjects were not trained on a particular setting before the beginning of data acquisition. In this way, data were acquired while learning was in progress. Experimental pilots showed that 150 trials allowed for the subjects' performance to stabilize sufficiently. Second, in addition to modifying target probabilities across blocks, the probability *structure* underlying the sequence of target locations was varied.

In a *state-oriented block*, as in previous experiments, the sequence of target locations was generated according to fixed state probabilities. They were specified as a vector $(p, 1-p)$, $0 \leq p \leq 1$, where $p$ and $1-p$ denote the marginal probabilities of leftward and rightward targets, respectively.

In a *transition-oriented block*, the probability distribution of the target location of the current trial was conditional on the target location of the previous trial. Thus, given a sequence of past trials, the probability distribution of the target on the next trial depended on the last item of the sequence, and only on this one. A sequence with this property is known as a first-order Markov chain, and the change from one trial to the next as a transition. The probability that the next target location is $j$, given that the current target location is $i$, is given by $p_{i,j}$. Thus, the sequence of target locations was specified by the transition matrix of its underlying Markov chain, $P = [p_{i,j}]_{1 \leq i,j \leq 2}$, where $p_{i,1}$ and $p_{i,2}$ denote the probabilities of leftward and rightward targets, respectively, given that the target of the previous trial appeared at location $i$. The first target in the sequence was drawn from a uniform initial distribution $\pi = (0.5, 0.5)$; that is, the sequence of target locations was initialized randomly, either with a leftward or with a rightward target. The example in Table 1 shows a short sequence of trials generated from the transition matrix

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}.$$

For each trial, the table shows the probability distribution vector from which the current target location is drawn. For the first trial, it is (0.5, 0.5). In all subsequent trials, it is either the top or the bottom row of $P$, depending on whether the previous target location was 'left' (top row) or 'right' (bottom row). The example illustrates that, given a transition matrix with high diagonal probabilities (a 'stable' transition matrix), the target tends to stay where it was on the previous trial, and only occasionally switches to the other side.

Finally, in a *uniform block*, target locations occurred on the left-hand side and the right-hand side with equal chance, rendering the

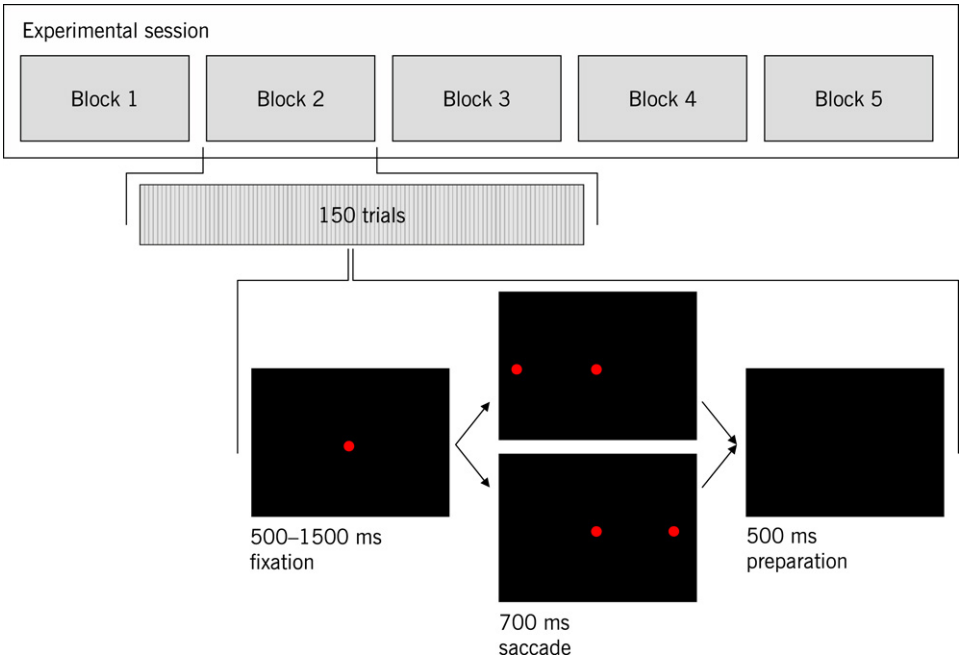


**Fig. 1.** Experimental design. A complete session consists of 5 blocks, each of which contains 150 trials generated from the same block-specific transition matrix. All matrices shown in the main text were used to generate samples in each session. A trial consists of three consecutive stages: a fixation screen (showing a central red fixation dot); a target screen (showing both the fixation dot and an additional leftward or rightward target dot); an inter-trial interval (showing a black screen until the beginning of the next trial).

**Table 1**

Example of a sequence of target locations generated from the transition matrix $\begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$

| Trial | 1 | 2 | 3 | 4 | 5 | 6 | … |
|---|---|---|---|---|---|---|---|
| Target probabilities | (0.5, 0.5) | (0.7, 0.3) | (0.7, 0.3) | (0.7, 0.3) | (0.3, 0.7) | (0.3, 0.7) | (0.7, 0.3) |
| Target location drawn | Left (1) | Left (1) | Left (1) | Right (2) | Right (2) | Left (1) | … |

On trial 1, the target location is always drawn from a uniform distribution (0.5, 0.5). On all subsequent trials, its probability distribution depends on the target location of the previous trial.

sequence of targets maximally unpredictable. This block structure served as a control condition in which no statistical learning across trials should take place.

In order to avoid drowsiness, which subjects in pilot experiments had displayed after 30 min of constant testing, a single experimental session was chosen to contain only 5 blocks. A break of 3 min between any two blocks was introduced to reduce the potential confound of learning effects carrying over from one block to another.

In order to allow for a unified formalism, all blocks were specified in terms of a transition matrix $P$. The blocks for each session were chosen according to the scheme in Box I.

Each session contained all five block types. Their order was randomized in each session, and the two alternative matrices underlying the state-oriented blocks were counterbalanced across subjects. In order to distinguish transition-oriented learning from simpler state-oriented learning, all transition-oriented blocks were designed in such a way that the states of the implied Markov chain, 1 and 2, had a uniform steady state distribution $\pi^* = (0.5, 0.5)$ (see Papoulis (1991)). Hence, in transition-oriented blocks, targets would, on average, appear equally often on either side, and no state-oriented learning should take place.

Experimental data were collected from three healthy male right-handed authors of this article with normal vision aged between 23 and 40 years (KHB, KES, WDP; see Table 2). Eye movements were recorded at a sampling frequency of 120 Hz using an ASL 504 infrared remote optics eye tracker. Targets were presented on a 27 cm × 37 cm CRT screen at a viewing distance of 67 cm. Data acquisition and analysis were implemented using MATLAB, Cogent 2000, and ILAB (Gitelman, 2002).

Before extracting latencies from eye recordings, blinks were filtered by searching for invalid pupil size values. Pupil coordinates within a time window of 25 ms around the beginning and the end of a blink were removed. Saccades were then detected using a standard algorithm by Fischer, Biscaldi, and Otto (1993): in the raw recorded eye coordinates we looked for an initial pupil velocity of 250°/s and searched the consecutive 100 ms time window for a saccade of at least 10° that resulted in a fixation of at least 100 ms. Any latencies below 10 ms or above 800 ms were interpreted as artifacts and removed, as were blocks with an overall recognition rate below 80%. Altogether, 15% of the recorded blocks were rejected, as were 4% of the trials from accepted blocks. For the remaining trials, we computed the latency between target onset and the beginning of the first detected saccade.

In order to reduce the variance of latencies, each subject took part in many sessions with an overall number of more than 20 000 trials.

### 2.2. Modelling

Various models have been proposed over the past two decades to explain the variability in the latencies between the appearance of a target and the initiation of an eye movement towards it. In one class of models, a decision signal is assumed to rise at a linear rate until reaching a fixed threshold. The release of a saccade is then modelled as the final outcome of this linear rise-to-threshold

| | | | | |
|---|---|---|---|---|
| $\begin{bmatrix}0.5 & 0.5\\ 0.5 & 0.5\end{bmatrix}$ | $\begin{bmatrix}0.7 & 0.3\\ 0.7 & 0.3\end{bmatrix}$ or $\begin{bmatrix}0.3 & 0.7\\ 0.3 & 0.7\end{bmatrix}$ | $\begin{bmatrix}0.9 & 0.1\\ 0.9 & 0.1\end{bmatrix}$ or $\begin{bmatrix}0.1 & 0.9\\ 0.1 & 0.9\end{bmatrix}$ | $\begin{bmatrix}0.7 & 0.3\\ 0.3 & 0.7\end{bmatrix}$ | $\begin{bmatrix}0.9 & 0.1\\ 0.1 & 0.9\end{bmatrix}$ |
| 'uniform' | 'weak state orientation' | 'strong state orientation' | 'unstable transition orientation' | 'stable transition orientation' |

**Box I.**

**Table 2**
Number of blocks (B) and trials (T) with successfully extracted saccade latencies, per subject and type of transition matrix

| Subject | Uniform | | Weak state orientation | | Strong state orientation | | Unstable transition orientation | | Stable transition orientation | | All | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | T | B | T | B | T | B | T | B | T | B | T |
| S-1 | 7 | 948 | 7 | 985 | 4 | 557 | 5 | 691 | 5 | 685 | 28 | 3866 |
| S-2 | 12 | 1727 | 12 | 1736 | 11 | 1554 | 11 | 1558 | 9 | 1285 | 55 | 7860 |
| S-3 | 11 | 1599 | 11 | 1594 | 12 | 1779 | 12 | 1755 | 11 | 1625 | 57 | 8352 |

mechanism (Carpenter & Williams, 1995; Leach & Carpenter, 2001; Reddi et al., 2003). This type of model can be extended in two ways: (i) within an individual trial, the linear rise to threshold can be parameterized and turned into a generative model; (ii) across trials, the dynamics of alternative forms of learning can be integrated into the model.

The two levels can be formalized as hierarchically related *intra-trial* and *inter-trial* sub-models, respectively. They are described separately in the following sections. Put together, they predict saccade latencies on the basis of the sequence of target locations observed so far, as well as three model parameters.

### 2.2.1. Intra-trial modelling

We propose a generative intra-trial model that extends previous models of the relation between prior expectations about target location and saccadic onset times. It describes a computational abstraction of putative neurophysiological processes between target onset and the release of a saccade.

Each trial has two potential outcomes: a target appears either on the left-hand side or on the right-hand side of the screen. Within each trial, we model a subject's belief in these two outcomes at any given point of time as distinct decision signals $S$ that evolve linearly over time $t$ until one of them hits a threshold $S_T$. At the beginning of a trial, both decision signals have specific initial values, with the signal of the 'winning' hypothesis defined to start at $S_0$. The initial level $S_0$ reflects the subject's prior as provided by the inter-trial model described in Section 2.2.2. As the stimulus appears at $t = t_0$, the two signals increase or decrease, respectively, representing the changing belief in the two hypotheses. As soon as one of them hits threshold, a saccade is released towards the corresponding target location at time $\tau$. The rate at which the decision signals rise or fall varies over trials, accounting for the large variance of the resulting latency distribution (see Fig. 2).

The decision process can be parameterized by modelling subjects as Bayesian observers who collect evidence about the true state of the world and, combined with their prior expectations, accept one of two competing hypotheses about it (Knill & Pouget, 2004). In this scheme, the initial level of the 'winning' decision signal, $S_0$, is associated with the subject's prior, and the rate of rise is associated with the likelihood of the true hypothesis.

Evidence from several studies provides support for this parameterization. For instance, based on psychophysical experiments in humans, Carpenter and Williams (1995) plotted latencies on a reciprobit scale, in which normally distributed data approach a straight line. They found that a change in the marginal probabilities of the two target locations led to a reciprobit swivel. This change in slope is consistent with a change in the threshold height but not with a change in the mean rate of rise, which would cause a reciprobit shift (Sinha et al., 2006). Further support for the notion that priors determine the initial level of the decision signal rather than its rate of rise comes from neurophysiological experiments using a similar task to ours in monkeys (Basso & Wurtz, 1997, 1998; Ratcliff, Cherian, & Segraves, 2003). These studies identified neurons in the superior colliculus whose firing rates just before target onset reflected the target probability but not target salience. Altogether, these human and primate studies provide a robust foundation for the assumption that the subjective prior systematically influences the initial level of the decision signal, $S_0$, before it starts to rise until hitting threshold.

Let the two possible states of the world be denoted by $h_i$, $i \in \{1, 2\}$, corresponding to the target location being $x_k = i$, respectively, within the current trial $k$. The sensory evidence for the hypotheses $h_i$ is provided by time-continuous visual input. Assuming this supportive evidence to be processed in small, discrete timesteps, in a lossless fashion without any form of temporal filter (Ludwig, Gilchrist, McSorley, & Baddeley, 2005), the evidence at time $t$ is referred to as $e_t$, and the accumulated evidence for one or another hypothesis up to time $t$ is denoted by $e_{1..t}$. Writing $p(e_t)$ for the probability density of the piece of evidence at time $t$, we make two simple assumptions. First, it is assumed that $p(e_2|e_1, h_i) = p(e_2|h_i)$, that is, $e_1$ and $e_2$ are conditionally independent. This means that $p(e_t|e_{1..t-1}, h_i) = p(e_t|h_i)$. Second, since sensory stimuli $e := e_1 = \ldots = e_t$ are equal throughout the duration of the trial, the likelihood term $p(e|h_i)$ is taken to be constant. It follows that

$$\begin{aligned} p(e_{1..t}|h_i) &= p(e_t|e_{1..t-1}, h_i) \times p(e_{t-1}|e_{1..t-2}, h_i) \times \cdots \times p(e_1|h_i) \\ &= p(e_t|h_i) \times p(e_{t-1}|h_i) \times \cdots \times p(e_1|h_i) \\ &= p(e|h_i)^t. \end{aligned} \tag{1}$$

Subjects are modelled as permanently testing a decision rule which determines whether they continue their observation—or accept one of the hypotheses. From a Bayesian learning perspective it is intuitive to consider, as a decision variable, the subjective posterior probability of each hypothesis, given the supporting evidence up to time $t$. In an iterative form, its dynamics can be written as

$$\begin{aligned} \mathbb{P}(h_i|e_{1..t}) &= \mathbb{P}(h_i|e_t, e_{1..t-1}) \\ &= \frac{p(e_t|h_i, e_{1..t-1}) \times \mathbb{P}(h_i|e_{1..t-1})}{p(e_t|e_{1..t-1})} \\ &= \frac{p(e_t|h_i) \times \mathbb{P}(h_i|e_{1..t-1})}{p(e_t|e_{1..t-1})}, \end{aligned} \tag{2}$$

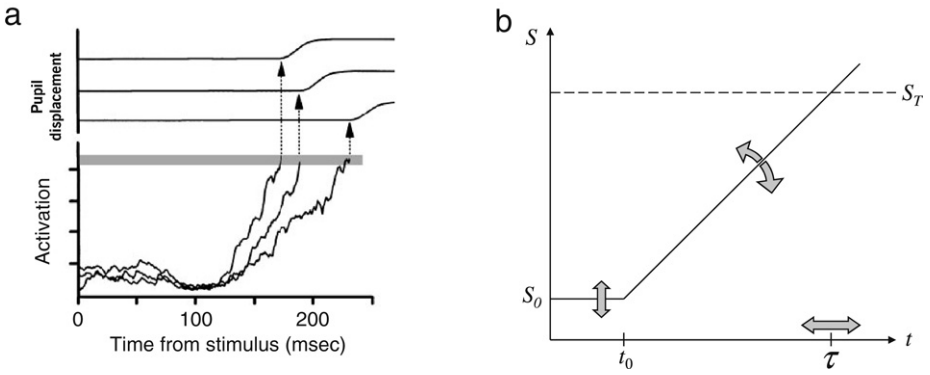


**Fig. 2.** Neuronal responses of a decision process and translation into a computational model. (a) Neuronal responses prior to a saccade from three trials. In their experiment, Schall and Thompson (1999) trained rhesus monkeys to stare at a central fixation stimulus and, as soon as eight secondary targets appeared, to elicit a saccade towards the oddball. The targets were arranged radially around the central fixation stimulus, and the location of the oddball was random. The diagram shows the recorded activity of single movement-related neurons in the saccadic movement maps of the frontal eye fields (FEF). Trials were grouped into those with slow, medium and fast saccades. The three plots show the averaged activity within these groups of trials, in each trial taking the activity from that neuronal response field corresponding to the correct target location. The activity patterns show that there is a fairly constant biophysical threshold at which a saccade is irrevocably elicited (grey bar) whereas the rate at which the signals rise varies between the groups of saccades. (Reprinted, with permission, from the Annual Review of Neuroscience, Volume 22 (c) 1999 by Annual Reviews, www.annualreviews.org) (b) Translation into a computational model of the decision process for a single trial. The rising activation in the FEFs is modelled as a linearly rising decision signal $S$. It starts off at an initial level $S_0$ and rises at a variable rate until reaching threshold $S_T$ at time $\tau$.

illustrating how the prior probability $\mathbb{P}(h_i|e_{1..t-1})$ is turned into a posterior probability $\mathbb{P}(h_i|e_{1..t})$ as new evidence $e_t$ is processed. Using Bayes' theorem and Eq. (1), we obtain the closed form

$$\begin{aligned}\mathbb{P}(h_i|e_{1..t}) &= \frac{p(e_{1..t}|h_i) \times \mathbb{P}(h_i)}{\sum_{i=1}^{2} p(e_{1..t}|h_i) \times \mathbb{P}(h_i)} \\ &= \frac{p(e|h_i)^t \times \mathbb{P}(h_i)}{\sum_{i=1}^{2} p(e|h_i)^t \times \mathbb{P}(h_i)},\end{aligned} \tag{3}$$

in which the assumption of discretized time is no longer necessary. However, this quantity does not rise linearly over time. Therefore, as an alternative decision variable that can be constructed in the case of two possible target locations, we consider the log posterior ratio. Using (2), its iterative form can be written as

$$\ln \frac{\mathbb{P}(h_1|e_{1..t})}{\mathbb{P}(h_2|e_{1..t})} = \ln \frac{\mathbb{P}(h_1|e_{1..t-1})}{\mathbb{P}(h_2|e_{1..t-1})} + \ln \frac{p(e_t|h_1)}{p(e_t|h_2)}. \tag{4}$$

Using (3), the closed form is

$$\ln \frac{\mathbb{P}(h_1|e_{1..t})}{\mathbb{P}(h_2|e_{1..t})} = \ln \frac{\mathbb{P}(h_1)}{\mathbb{P}(h_2)} + t \times \ln \frac{p(e|h_1)}{p(e|h_2)}, \tag{5}$$

which can be written in an analogous fashion for its counterpart $\ln \frac{\mathbb{P}(h_2|e_{1..t})}{\mathbb{P}(h_1|e_{1..t})}$ by interchanging $h_1$ and $h_2$. These log-odds are attractive candidates for computational models of neuronal processes of decision making because they (i) allow for optimal decision making and (ii) rise linearly over time, as shown in Fig. 2a. It is assumed that a hypothesis is accepted when its decision variable reaches a fixed threshold. This yields a decision rule that is evaluated at each point of time $t$:

$$\text{accept} \begin{Bmatrix} x_k = 1 \\ x_k = 2 \end{Bmatrix} \text{when} \begin{Bmatrix} \ln \frac{\mathbb{P}(h_1|e_{1..t})}{\mathbb{P}(h_2|e_{1..t})} \\ \ln \frac{\mathbb{P}(h_2|e_{1..t})}{\mathbb{P}(h_1|e_{1..t})} \end{Bmatrix} > \vartheta, \tag{6}$$

and otherwise neither is accepted. Note that equivalently optimal decision rules, although framed somewhat differently, have been proposed by previous authors. For example, in the case of a forced-choice task, Gold and Shadlen (2001) consider a decision rule that is based on the likelihood ratio of the two hypotheses: accept $x_k = 1$ when $\frac{p(e_{1..t}|h_1)}{p(e_{1..t}|h_2)} > \frac{\mathbb{P}(h_2)}{\mathbb{P}(h_1)}$, and accept $x_k = 2$ otherwise. This rule can be turned into a decision rule for our task by multiplying the right-hand side criterion by an additional factor $c$ that introduces the necessary 'temporal gap' in which neither hypothesis is accepted. Equating $\vartheta \equiv \ln c$, taking logarithms, and using Eq. (1), this modified rule can be rewritten as

$$\begin{aligned}&\frac{p(e_{1..t}|h_1)}{p(e_{1..t}|h_2)} > \frac{\mathbb{P}(h_2)}{\mathbb{P}(h_1)} \times e^{\vartheta} \\ &\ln \frac{\mathbb{P}(h_1)}{\mathbb{P}(h_2)} + \ln \frac{p(e|h_1)^t}{p(e|h_2)^t} > \vartheta \\ &\ln \frac{\mathbb{P}(h_1)}{\mathbb{P}(h_2)} + t \times \ln \frac{p(e|h_1)}{p(e|h_2)} > \vartheta,\end{aligned} \tag{7}$$

which is precisely the same rule as in (6). This means that the two approaches are decision-equivalent.

Both the decision variable for the true hypothesis in (7) and its counterpart for the alternative hypothesis start at specific initial levels that represent the subject's prior, and then rise or fall, respectively, over time. This corresponds to the notion of the accumulation of supportive evidence for the two rival hypotheses. A saccade to the true target location is released at time $\tau$ when

$$\begin{aligned}&\ln \frac{\mathbb{P}(h_{x_k}|e_{1..\tau})}{\mathbb{P}(h_{\bar{x}_k}|e_{1..\tau})} = \vartheta \\ &\ln \frac{\mathbb{P}(h_{x_k})}{\mathbb{P}(h_{\bar{x}_k})} + \tau \times \ln \frac{p(e|h_{x_k})}{p(e|h_{\bar{x}_k})} = \vartheta \\ &\tau = \frac{\vartheta - \ln \frac{\mathbb{P}(h_{x_k})}{\mathbb{P}(h_{\bar{x}_k})}}{\ln \frac{p(e|h_{x_k})}{p(e|h_{\bar{x}_k})}},\end{aligned} \tag{8}$$

where $x_k$ and $\bar{x}_k \in \{1, 2\}$ denote the true and the false target location of the current trial $k$, respectively.

The likelihood term $p(e|h_i)$ can be thought of as a descriptor of a subject's visual discrimination efficiency or processing capacity. The larger it is the more quickly will an observed sensory stimulus make the subject increase their posterior belief in the corresponding hypothesis, and the shorter the resulting saccade latency. One way of parameterizing this quantity is to assume arbitrary 'evidence units'. With

$$p(e|h_{x_k}) = 1 + \rho \quad \text{and} \quad p(e|h_{\bar{x}_k}) = 1 \tag{9}$$

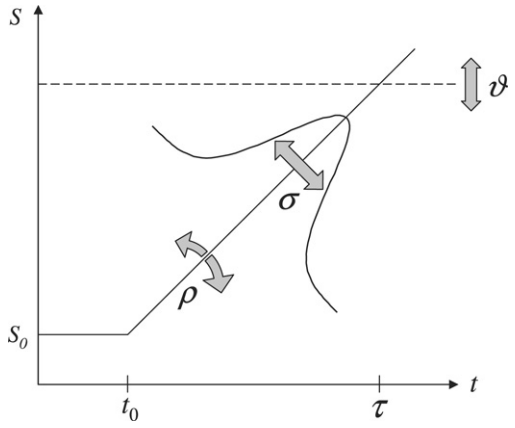


**Fig. 3.** Intra-trial model parameterization. The proposed intra-trial model has three free parameters, represented by grey arrows. (i) $\rho$ and (ii) $\sigma$ determine the mean and the standard deviation of the normally distributed slope of the decision signal that corresponds to the true target location of the current trial. The larger $\rho$, the shorter the predicted saccade latency $\tau$. The larger $\sigma$, the larger the variability of the distribution of $\tau$. (iii) $\vartheta$ specifies the threshold the decision signal has to reach in order to evoke a saccade. The larger $\vartheta$, the longer the latency and the less the influence of the initial value $S_0$.

the supportive evidence per unit time for the true hypothesis is larger than the evidence for the false hypothesis, by an amount determined by a second model parameter $\rho > 0$.

In addition to the parameters $\vartheta$ and $\rho$, we must account for the fact that, across trials, the rate of the decision signal varies (see Fig. 2a). Previous experiments based on the same paradigm as in this study have found reciprocal latencies to conform to a Gaussian distribution (Carpenter & Williams, 1995). Hence, in trial $k$, the rate of the assumed decision signal, $r_k = \Delta S_k \frac{1}{\tau_k}$, will have a Gaussian distribution as well, with $\Delta S_k = S_T - S_{0,k}$ denoting the difference between the threshold $S_T$ and the initial level $S_{0,k}$ of the decision variable (see Fig. 2b). This introduces a third model parameter $\sigma$ that describes the standard deviation of $r$. The parameterization of the intra-trial model is summarized in Fig. 3.

### 2.2.2. Inter-trial modelling

A central assumption of our model is that, at the beginning of each trial, the starting point of the decision signal associated with the true hypothesis corresponds to $\ln \frac{\mathbb{P}(h_{x_k})}{\mathbb{P}(h_{\bar{x}_k})}$, i.e. the log ratio of the prior probabilities of the correct and the incorrect hypothesis, see Eq. (5). The evolution of the prior probabilities $\mathbb{P}(h_i)$ during a sequence of trials should reflect the learning of certain statistical properties of the target locations. In order to investigate what type of learning might happen during a sequence of trials $x_1, x_2, \ldots, x_N$, we propose three different inter-trial models that formalize alternative learning hypotheses.

*2.2.2.1. The transition model.* In the first candidate inter-trial model, we assume that subjects act like ideal observers: while responding to the sequence of stimuli within a block, they continuously refine an estimate of the underlying Markov transition matrix (see Minka (2001)).

Let $P = [p_{i,j}]$ be a hidden homogeneous transition matrix with a uniform initial distribution $\pi = (0.5, 0.5)$ underlying the sequence of target locations $x_1, \ldots, x_N$. The states of the Markov chain, 1 and 2, encode leftward and rightward targets, respectively. Let $n_{i,j}^{(k-1)}$ denote the number of transitions $x_i \to x_j$, for $i, j \in \{1, 2\}$, that have occurred in the sequence of $k-1$ trials observed so far. A maximum likelihood estimate of $P$ could be obtained by maximizing the log likelihood function

$$\ln L(\hat{P}; x_1, \ldots, x_{k-1}) = \sum_{i,j} n_{i,j}^{(k-1)} \ln \hat{p}_{i,j} \qquad (10)$$

subject to $\sum_j p_{i,j} = 1 \; \forall i \in \{1, 2\}$. Using Lagrange's method, the maximum likelihood estimates $\hat{p}_{i,j}^{\mathrm{ML}} = \frac{n_{i,j}}{\sum_{l=1}^{2} n_{i,l}}$ follow.

For example, having observed the first six trials in Table 1, an ideal observer will have counted two 'left→left' transitions, one 'left→right' transition, and so forth. From this follows an estimated transition matrix $\hat{P}^{\mathrm{ML}} = \begin{bmatrix} 0.67 & 0.33 \\ 0.5 & 0.5 \end{bmatrix}$. It is the matrix that makes the observed sequence of trials most likely.

In this form, however, an individual matrix element $\hat{p}_{i,j}^{\mathrm{ML}}$ remains undefined as long as $n_{i,1} + n_{i,2} = 0$. Instead, we assume an initial uniform prior of $p_{i,j}^{(0)} = \frac{1}{2}$ for all $i, j \in \{1, 2\}$, which can be thought of as an imaginary 'prior observation count' of 1 for each transition event (Minka, 2001). The posterior predictive distribution $\hat{p}_{i,j}^{(k)}$ based on $k$ trials then allows subsequent generative models to yield predictions for all trials. Specifically, at the beginning of trial $k$, subjects are assumed to have constructed the estimate

$$\hat{P}^{(k-1)} = [\hat{p}_{i,j}^{(k-1)}]_{i,j \in \{1,2\}} \quad \text{with} \quad \hat{p}_{i,j}^{(k-1)} := \frac{n_{i,j}^{(k-1)} + 1}{\sum_{l=1}^{2} (n_{i,l}^{(k-1)} + 1)}, \qquad (11)$$

such that $\hat{p}_{i,j}^{(k-1)} > 0 \; \forall i, j, k$ and $\hat{p}_{i,j}^{(0)} = \frac{1}{2} \; \forall i, j$. Using this alternative formulation, the estimated transition matrix in the above example (Table 1) would be $\hat{P}^{(6)} = \begin{bmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix}$.

Obtaining maximum likelihood estimates of the transition matrix elements in this way can equivalently be viewed as a Bayesian update scheme. By counting how often each type of transition has occurred so far, an ideal observer can estimate the joint probabilities $\mathbb{P}(x_{k-1} = i, x_k = j)$, from which estimates of the conditional probabilities $\hat{p}_{i,j} \equiv \mathbb{P}(x_k = j | x_{k-1} = i)$ can be derived (see Harrison, Duggins, and Friston (2006), for an example).

The initial uniform matrix at the beginning of an experimental block corresponds to maximal uncertainty. As more and more trials are observed, the posterior of the transition matrix is refined and gradually approaches the true matrix.

*2.2.2.2. The state model.* The inter-trial model outlined so far is *transition*-oriented in that it assumes an observer who estimates a transition matrix underlying the sequence of stimuli. Alternatively, a *state*-oriented observer can be imagined who simply estimates a state probabilities vector $\hat{p} = (p_1, p_2)$ by counting the frequencies $n_i \; \forall i \in \{1, 2\}$ of the two target locations while not paying attention to the transitions between them. At the beginning of trial $k \in \{1, \ldots, N\}$, in analogy to the Bayesian update scheme outlined above, this estimate is

$$\hat{p}^{(k-1)} = [\hat{p}_i^{(k-1)}]_{i \in \{1,2\}} \quad \text{with} \quad \hat{p}_i^{(k-1)} := \frac{n_i^{(k-1)} + 1}{\sum_{l=1}^{2} (n_l^{(k-1)} + 1)}, \qquad (12)$$

such that $\hat{p}_i^{(k-1)} > 0 \; \forall i$ for all $k \in \{1, \ldots, N\}$. Again, the initial prior is assumed to be uniform, $\hat{p}_1^{(0)} = \hat{p}_2^{(0)} = \frac{1}{2}$.

Having observed the first six trials of the above example (Table 1), a 'state' observer will have counted four 'left' trials and two 'right' trials. Using (12), an estimated state probabilities vector $(0.63, 0.38)^{\mathrm{T}}$ follows. Using the same notation as in the transition model, this estimate can be written as $\hat{P}^{(6)} = \begin{bmatrix} 0.63 & 0.38 \\ 0.63 & 0.38 \end{bmatrix}$.

*2.2.2.3. The uniform model.* The two alternative inter-trial models proposed so far account for different forms of learning, but they do not question whether learning occurs at all. Therefore, a third candidate model is proposed in which subjects are assumed to be entirely ignorant of the history of stimuli. In this *uniform* model, subjects maintain a uniform prior belief of $\hat{p}_1 = \hat{p}_2 = 0.5$ in either target location throughout the experiment. Using the same notation as in the transition model, this prior can be written as a constant estimate $\hat{P}^{(k-1)} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$ $\forall k = 1, \ldots, N$.

Fig. 4 illustrates exemplary predictions generated by the alternative inter-trial models operating on alternative block structures. The individual diagrams show the extent to which the models are able to adapt to the transition matrix underlying the observed sequence of target locations. Crucially, the rate of convergence is highest when the model structure most closely matches the block structure. In particular, convergence takes longer when the true block structure is more complicated than the assumed one. For example, in the case of a uniform block, all three models eventually settle around 0.5/0.5 predictions, but the 'state' model and the 'transition' model take longer to converge. Thus, we will be able to make use of measured reaction times from *all* blocks to find out which model explains a particular subject's behaviour best (Section 3.4).

*2.2.3. Model construction*

The two sub-models outlined above can now be integrated into an overall generative model with three free parameters.

In our paradigm, reciprocal latencies have previously been observed to follow a normal distribution (Carpenter & Williams, 1995). Thus, the rate of the rising decision signal can be modelled as a random variable

$$R_k \sim \mathcal{N}(\mu_r, \sigma^2), \tag{13}$$

where $\mu_r$ and $\sigma$ describe the mean and the standard deviation of the rate across trials. Using $\frac{1}{\tau} = \frac{r}{\Delta S}$, the reciprocal latency $y_k := \frac{1}{\tau_k}$ can then be modelled as a random variable

$$Y_k \sim \mathcal{N}\left(\frac{1}{\Delta S_k}\mu_r, \frac{1}{(\Delta S_k)^2}\sigma^2\right). \tag{14}$$

The difference $\Delta S_k$ between the initial level $S_{0,k}$ of the decision signal and its threshold can be expressed in terms of a model parameter $\vartheta$ and the log prior ratio as derived in Section 2.2.1:

$$\Delta S_k = S_T - S_{0,k} = \vartheta - \ln\frac{\mathbb{P}(h_{x_k})}{\mathbb{P}(h_{\bar{x}_k})}. \tag{15}$$

The mean rate of the decision signal was derived in (5) and (9):

$$\mu_r = \ln\frac{p(e|h_{x_k})}{p(e|h_{\bar{x}_k})} = \ln(1+\rho). \tag{16}$$

The variability of the rate across trials is described by the model parameter $\sigma$. The prior of the two hypotheses is given by the corresponding entry in the transition matrix, as estimated within the inter-trial model:

$$\mathbb{P}(h_{x_k}) = \hat{p}_{x_{k-1},x_k} \quad \text{and} \quad \mathbb{P}(h_{\bar{x}_k}) = \hat{p}_{x_{k-1},\bar{x}_k} = 1 - \hat{p}_{x_{k-1},x_k}. \tag{17}$$

The overall probability distribution for reciprocal latencies, conditional on the model parameters, is then given by

$$Y_k|\rho, \vartheta, \sigma \sim \mathcal{N}\left(\frac{\ln(1+\rho)}{\vartheta - \ln\frac{\hat{p}_{x_{k-1},x_k}}{1-\hat{p}_{x_{k-1},x_k}}}, \left[\frac{\sigma}{\vartheta - \ln\frac{\hat{p}_{x_{k-1},x_k}}{1-\hat{p}_{x_{k-1},x_k}}}\right]^2\right). \tag{18}$$

Note that the overall model does not propose a single parameterized distribution $Y$, but a sequence of distributions $(Y_k)_{1\le k\le N}$. This is because of its dependence on the output of the inter-trial model, $\hat{p}_{x_{k-1},x_k}$, which in turn depends on the sequence of target locations $(x_1, \ldots, x_{k-1})$ observed so far.

*2.3. Parameter estimation*

Each candidate model constructed in the preceding section proposes a particular distribution of reciprocal saccade latencies parameterized by a vector $\theta = (\rho, \vartheta, \sigma)$. Fig. 3 indicates that the parameters are not perfectly independent in their effect on the resulting predictions. For example, an increase in predicted response speed can be obtained by either decreasing the threshold $\vartheta$ or by increasing the rate of the decision signal $\rho$. Note, however, that this also alters the dependence of reaction times on the subject's priors, so the parameters are not completely interchangeable. Nevertheless, to avoid numerical identifiability problems during parameter estimation, we reparameterized the model such that the rate $\rho$ is expressed as per unit threshold, i.e. $\rho \to \rho/\vartheta$. Moreover, since the dispersion parameter $\sigma$ must take a non-negative value, we used a log-transform. Thus, during numerical parameter estimation, the model was parameterized by $\theta' = (\rho/\vartheta, \vartheta, \ln\sigma)$.

The maximum likelihood principle identifies those model parameters of the distributions of $Y_1, \ldots, Y_N$ that are most likely to give rise to the data $y_1, \ldots, y_N$. An estimate $\hat{\theta}_{\mathrm{ML}} = (\hat{\rho}, \hat{\vartheta}, \hat{\sigma})$ can be found by maximizing

$$\ln\prod_{k=1}^{N} p_k(y_k|\theta) = -\frac{N}{2}\ln 2\pi - \sum_{k=1}^{N}\ln\frac{\sigma}{\vartheta - \ln\frac{\hat{p}_{x_{k-1},x_k}}{1-\hat{p}_{x_{k-1},x_k}}} - \sum_{k=1}^{N}\frac{\left(y_k - \frac{\ln(1+\rho)}{\vartheta - \ln\frac{\hat{p}_{x_{k-1},x_k}}{1-\hat{p}_{x_{k-1},x_k}}}\right)^2}{2\left(\frac{\sigma}{\vartheta - \ln\frac{\hat{p}_{x_{k-1},x_k}}{1-\hat{p}_{x_{k-1},x_k}}}\right)^2} \tag{19}$$

with respect to $\theta$. The implementation can be simplified by omitting the term $-\frac{N}{2}\ln 2\pi$ which does not depend on the free parameters. Multiplying the expression by $(-1)$, the maximum likelihood estimate $\hat{\theta}_{\mathrm{ML}}$ is then found as the solution to a minimization problem with respect to $\vartheta, \rho, \sigma$.

## 3. Results

*3.1. Data analysis*

In order to validate our paradigm, we replicated two key results by Carpenter and Williams (1995). First, latencies varied considerably over trials. Across our 20 078 trials, we found an overall mean of $\mu = 262$ ms and a standard deviation of $\sigma = 65$ ms. Second, reciprocal latencies $\tau_k = 1/y_k$, with $\mu = 4.02 \times 10^{-3}$ ms and $\sigma = 1.67 \times 10^{-3}$ ms, closely approximated a normal distribution (see Fig. 5).

Above and beyond these descriptive features, Carpenter and colleagues showed that latencies declined as the learned state probability of the corresponding target location increased (Basso & Wurtz, 1997, 1998; Carpenter & Williams, 1995). Specifically, they found a significant negative linear relation between log prior probabilities and latencies. In our study, we replicated these results as well, finding a very similar negative linear relation between log state probabilities and latencies in our state-oriented blocks ($p < 0.01$). We further extended this analysis to transition probabilities. Using a linear regression analysis, we found a significant relation ($p < 0.01$) between reciprocal latencies and log transition probabilities (see Fig. 6). Even though this analysis neglects the dynamics of learning completely, it already indicates that human observers are sensitive to transition probabilities in visual input statistics. Yet formal model comparison will provide much stronger evidence for this claim.

**Fig. 4.** Log prior ratios predicted by the alternative inter-trial models. Each diagram is based on the combination of a particular block structure (transition-oriented, state-oriented, or uniform) and a particular inter-trial model ('transition' model, 'state' model, or 'uniform' model). For each trial, the diagrams show the target location (black dots), with high and low markers indicating leftward and rightward targets. Furthermore, they show the log ratio between the prior probability of the true and the false target location (grey squares) as well as a prediction for this log ratio, generated by the respective inter-trial model (black crosses). Since the models are always initialized with uniform priors, the predicted log ratio for trial $k = 1$ is $\ln(0.5/0.5) = 0$. The upper left diagram, for example, shows how the 'transition' model gradually adapts to the transition-oriented block structure underlying the observed sequence of trials. By contrast, the central diagram in the left column shows that the 'state' model is incapable of learning the structure of a transition-oriented block.

**Fig. 5.** Histogram of latencies and reciprocals. The diagrams are based on 28 blocks containing 3 866 trials from subject S-1. While the latencies themselves have often been described as log normally distributed (Glimcher, 2003), their reciprocals can be approximated by a normal distribution (Carpenter & Williams, 1995).

### 3.2. Parameter estimation

In order to obtain maximum likelihood parameter estimates from Eq. (19), a gradient-descent scheme was run on the acquired data from all three subjects separately. The results are given in Table 3.

In order to visualize the dependence between conditional probabilities and latencies, and provide face validity for our parameter estimates, one of the subjects was engaged in an additional session consisting of 10 transition-oriented blocks with identical sequences initially generated from a $\begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$ transition matrix. We fitted the model to the data and predicted the priors $\hat{p}_{x_{k-1},x_k}$ using the 'transition' model. Fig. 7 shows how observed latencies develop over time and how this is reflected by model predictions.

In order to give a qualitative comparison between the predictive power of the three competing inter-trial models, Fig. 8 visualizes measured latencies versus model predictions. The individual diagrams show a marked separation between trials in which the target has just switched to the other side (crosses) and those in which it has not (dots). As expected from a subject that learns transition probabilities, switch trials lead to long latencies, i.e. short reciprocals; accordingly, crosses have low $x$-coordinates. Conversely, non-switch trials lead to short latencies, i.e. long reciprocals; accordingly, dots have high $x$-coordinates. Looking at the $y$-coordinates, the 'transition' model is the only model that predicts these two clusters of trials.

### 3.3. Statistical classification

The parameter estimates obtained for each subject are sufficiently distinct to demonstrate the use of a statistical classifier that maps an unseen test sample onto the correct class (Jain, Duin, &