



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.

© 2008 Elsevier Ltd. All rights reserved.

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).

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

