Task Switching as a Two-Stage Decision Process

N. Sinha, J.T.G. Brown, and R.H.S. Carpenter
The Physiological Laboratory, University of Cambridge, Cambridge, United Kingdom

Submitted 8 November 2005; accepted in final form 31 January 2006


INTRODUCTION

A decision is the selection of one response from a number of alternatives. An extremely common example is the decision where to look, which we all make two or three times every second of our waking lives, resulting in the rapid, conjunct eye movements known as saccades. The reaction time or latency of these responses is surprising, being much longer than can be explained by conduction times in the shortest, collicular pathways linking vision to eye movement (Guitton 1992; Sparks 1986). Even more striking is the random variation of saccadic latency from trial to trial, and the study of this variability has provided some insight into the decision mechanisms that underlie the initiation of saccades.

The magnitude of this delay can be explained as a process of procrastination (Carpenter 1981): although the colliculus can translate a given visual location into an appropriate command to the brain stem to generate a foveating saccade, it lacks direct access to the information needed to decide between several competing targets. This demands the involvement of higher—essentially cortical—levels, which generate the procrastination by tonically suppressing the oversimplistic operations of the lower levels.

Therefore reaction time is largely decision time, and quantitative analysis of its variability shows that although reaction times are unpredictable on individual trials, they obey a simple stochastic law: the recurrences of reaction times, whether for saccades or other responses, have a Gaussian distribution. This can most easily be demonstrated by plotting cumulative latency distributions on a probit scale as a function of reciprocal latency (a reciprobit plot) when they typically yield a straight line (Fig. 1).

The LATER (Linear Approach to Threshold with Ergodic Rate) model provides an economical explanation of this behavior (Fig. 1) (Carpenter and Williams 1995). A decision signal $S$ rises linearly at a rate $r$ from a starting level $S_0$ until it reaches a threshold level $S_T$ when a response is initiated. In different trials, $r$ varies as a Gaussian, with mean $\mu$ and variance $\sigma^2$, giving rise to the observed random distribution of reaction times. The latency $\lambda$ on any trial is given by $k/r$, where $k = (S_T - S_0)$ so that the reciprocal of $\lambda$ is a Gaussian of mean $\mu/k$ and SD $\sigma/k$. Because the choice of units is arbitrary, only two parameters, for instance $\mu$ and $\sigma$, are needed to specify the behavior fully because we can define our units such that $k = 1$.

The values of the parameters can be determined directly from a reciprobit plot: the median latency, where the distribution intersects the $P = 0.5$ axis, determines $\mu$, and the slope of the line is equal to $1/\sigma$. In addition the intercept $I$ of the line with the $t = \infty$ axis is given by the dimensionless quantity $\mu/\sigma$, because this is unaffected by changes in $k$, alterations in $S_0$ or $S_T$ will cause the line to “swivel” around $I$, whereas changes in $\mu$ will shift it sideways in a self-parallel fashion (Fig. 1).

This empirically derived model is strengthened by its theoretical interpretation as an ideal Bayesian decision maker. $S$ is identified with the perceived probability of the existence of a situation demanding a response, on a logarithmic scale. $S_0$ then corresponds to prior probability, $\mu$ with the rate of arrival of information, and $S_T$ with a criterion at which a response is initiated—equivalent to a statistical significance level in conventional statistics. LATER can thus be regarded as collecting fragments of information until there is enough belief in the hypothesis to justify a response. In real life there would be many competing hypotheses, each presumably with its own LATER unit, racing against each other to reach the threshold.

Previous studies have confirmed this functional interpretation. LATER makes specific, quantitative predictions about the effects of altering the three parameters $S_T$, $S_0$, and $\mu$, both on the mean reaction times, but also—more demandingly—on the shape of their statistical distribution (Fig. 1). Modifying prior probability produces the predicted changes to the median latency and generates the expected “swivel” of the reciprobit plot about a fixed intercept on the $t = \infty$ axis (Carpenter and Williams 1995), and analogous results have been found for changing the criterion level (Reddi and Carpenter 2000). Manipulation of the information rate produced results consistent with a change in $\mu$ (Reddi et al. 2003). The notion of competing LATER units racing against one another can be verified by asynchronous presentation of targets (Leach and Carpenter 1986).

The Physiological Laboratory, University of Cambridge, Cambridge, United Kingdom

Address for reprint requests and other correspondence: R.H.S. Carpenter, The Physiological Laboratory, University of Cambridge, Cambridge CB2 3EG, UK (E-mail: rhsc1@cam.ac.uk).
and in countermanding paradigms (Hanes and Carpenter 1999). Finally, electrophysiological recordings provide support for the LATER model. Visuomotor neurons in FEF show an approximately linear increase in activity, before saccadic onset, whose rate varies randomly from trial to trial (Hanes and Schall 1996), whereas the time of initiation of the subsequent saccade has a fixed relation to the time when the activity reaches a fixed criterion. Saccade-related cells in the superior colliculus and the parietal cortex show rise-to-threshold behavior, and in the colliculus the starting level depends on prior probability (Basso and Wurtz 1997, 1998); with suitably designed experiments it is possible to demonstrate a functional distinction between this kind of motor-related behavior and activity representing an earlier, “perceptual” stage of processing that could be regarded as being related more to detection than to decision (Thompson et al. 1996).

However, although psychophysical and physiological investigations over the last 10 years have done much to improve our understanding of relatively simple neural decision making, how we make more complex decisions remains a mystery. Are there similar mechanisms, influenced by many more factors, or is the process qualitatively different? One example of a more complex decision that is still relatively controllable is the scenario known to psychologists as task switching.

A task is a situation in which behavior is governed by an instruction. Task switching is thus the result of a decision to follow one instruction instead of another. Studies typically compare reaction times in a sequence of trials in which the task is constant with one in which the task changes because of an instruction; the main results of this extensive body of research have recently been reviewed (Monsell 2003; Schneider and Logan 2005). Much of this work has used complex tasks, the mechanisms of which are themselves poorly understood, for example switching between color naming and word reading (Stroop 1935) or switching between classifying the digit member of a pair of characters as even or odd, and the letter member as consonant or vowel (Rogers and Monsell 1995). Given that complex tasks of this kind may themselves involve working and long-term memory, and have relatively ill-defined timing, it is necessarily difficult to isolate the decision to switch. In contrast, saccades have precise timing and their neural mechanisms are relatively well described both functionally and physically.
anatomically; as has recently been persuasively argued (Schall 2003; Schall et al. 2002; Shadlen and Gold 2004) the analysis of saccadic tasks provides a uniquely effective approach to the understanding of these more complex operations.

We propose that a key logical step has been overlooked in uniting the neurally plausible models of primates decision making with the “attentional level” models (Norman and Shallice 1980) of task switching, namely that recognizing an instruction is itself a decision. For switching between simple tasks, we envisage a model consisting of two LATER units in series, the first detecting the instruction and the second implementing it. More specifically, when the decision signal of the instruction detector reaches threshold it triggers the selection of a corresponding subset of the second-stage units (Fig. 1). This provides a neurally plausible model of saccadic task switching that generates several quantitative predictions. Furthermore, if the instructions are presented by means of stimulus elements identical to those of the task itself, decisions about instructions can be studied on a comparable basis to decisions in the task, so that similarities and differences can be highlighted.

Here, we investigate whether our two-stage model can simulate the distribution of both conventional and task-switching reaction times. We also vary the prior probability of switching, to check whether the first unit shows the behavior ing errors such as blinks, head movements, or other artifacts were withheld from further analysis.

**Data collection**

We recorded horizontal eye movements using a binocular infrared oculometer (Ober Consulting, Poznan, Poland) (Ober et al. 2003), consisting of dual differential infrared reflectance transducers comparing reflectance from the medial sclera and pupil of each eye. It had a bandwidth of 250 Hz, symmetrically linear to 7% within ±30°, and a noise level equivalent to a 10-min arc. Its output voltage was sampled at 100 Hz, synchronized to the display frames. Saccades were recorded in real time and stored in 10-ms bins by the PC-based saccadic analysis program SPIC (Carpenter 1994), which also controlled the presentation of stimuli. After experimental runs, all records were checked manually, and those with latencies <50 ms or containing errors such as blinks, head movements, or other artifacts were withheld from further analysis.

**Protocols**

In the main experiments, we used two kinds of trial: Repeat trials and Switch trials. In all experiments, an individual trial began with presentation of the black central fixation dot during a foreperiod whose duration was random and distributed uniformly in the interval 0.5–1.5 s. In Repeat trials, this was followed by presentation of a pair of dots, red and blue, with the red appearing randomly on the right or left with equal probability. However, in a Switch trial, at the same time as the pair of targets an additional red or blue central dot appeared. This told the volunteer which color to saccade to on that and subsequent trials, until instructed differently (Fig. 2). The central dot in Switch trials was alternately red and blue so that an instruction was not presented twice in succession. In a block of trials, Switch trials occurred randomly with a fixed (prior) probability, which for any particular block was set to one of 10, 20, 50, 70, or 90%. Blocks of trials for each probability were interleaved and each began with a Switch trial. Volunteers undertook a set of practice trials before the main experimental blocks.

In addition, as controls we ran blocks of the same kind but with different instructions: we told the volunteer to saccade to only one color throughout the block, regardless of the color of the central dot in Switch trials. This served to give baseline latency values for both Repeat and Switch trial displays in the absence of task switching because the visual stimuli were identical with experimental runs: only the instructions differed.

In the second part of this study, we varied the time delay in a Switch trial between presenting the central instructional dot and the appearance of the actual targets. Six time delays, or instruction lead times (ILTs), were chosen in addition to the 0-ms case: 50, 75, 100, 125, and 200 ms.
150, and 250 ms. Each block of trials implemented three of these ILTs, using a 20% probability of a Switch (Fig. 2).

Warning tones indicated wrong responses, or no response at all, and a further sound denoted the successful completion of a trial. Volunteers were told before each experiment to aim for accuracy of response rather than rapidity. An experimental run consisted of block of 300 trials, with rests as necessary between blocks.

Statistical analysis

Data sets pertaining to specific experimental conditions (color, direction, Switch or Repeat trials) were displayed and analyzed separately within SPIC. We used the Kolmogorov–Smirnov (K-S) one-sample test to determine the agreement between observed distributions and those predicted by the LATER model, and K-S two-sample tests to compare observed distributions with one another (Kolmogorov 1941). Best-fit estimates of the LATER parameters (median latency and slope) were obtained by iterative minimization using the K-S statistic as criterion. To compare two alternative distributional models we calculated log-likelihood measures for each subject and model and summed them across all subjects.

Simulations. Using SPIC, we ran Monte Carlo simulations of the responses predicted for each subject by a model consisting of two LATER units in series, the input to the implementation unit being applied only when the detection unit had reached its threshold. On each simulated trial the values of the rate of rise \( r \) for each unit were independently chosen as random Gaussian variates using a standard congruence procedure with a cycle length of over 107 (Abramowitz and Stegun 1965). The parameters \( (\mu_2, \sigma_2) \) of the implementation units were previously selected as those that minimized the K-S one-sample statistic for the observed distributions in Control trials. For the detection unit, we used the reciprocal of the median task switch cost (the average difference between the latency for Switch and Control trials) to provide \( \mu_1 \). We ran 1,000 trials in each simulation and compared the resultant distributions with the observed data with the K-S two-sample statistic. Values for \( \sigma_1 \) were iteratively estimated by minimization of this statistic.

Results

Distributions of response times

The necessarily small proportion of switch trials means that the collection of enough data to produce distributions from which statistically significant conclusions can be drawn is extremely time consuming. Across all subjects, we recorded a total of 43,700 saccades. There were 1,000 Control trials and 9,970 Switch trials. From these raw data we first determined the distribution of the latencies of saccades during Control trials and during successful Switch trials. In Fig. 3 (left) we show examples of reciprobit plots in which the subject’s cumulative histogram is plotted on reciprobit axes. The most striking difference between the distribution of reaction times for Switch and Control conditions is an increased median with a larger SD, the difference in medians being the task switch cost. As can be seen in Fig. 3 (right), across subjects it varies between about 100 and 250 ms and shows no obvious relationship to the latency in Control or Repeat trials.

Simulating a two-stage model

We performed Monte Carlo simulations to determine whether a model consisting of two LATER units in series could correctly predict the distributions resulting from the switch process. Of the four underlying parameters, \( \mu_1 \) [representing the mean rate of rise for the first (detection) unit] can be immediately estimated from the task switch cost. We estimated the two parameters describing the second (implementation) unit \( (\mu_2, \sigma_2) \) by K-S best-fit minimization in Control trials, in which the detection unit plays no part, and \( \sigma_1 \) can similarly be estimated by minimization for the Switch trials. The reciprobit plots in Fig. 3 (left) show distributions of simulated and actual saccade latency distribution for Switch and Control trials. Actual and simulated distributions were not significantly different.
different in any subjects (K-S, \( P > 0.1 \)). At bottom right the values of the parameters for each subject are plotted. It can be seen that the values for the implementation unit are relatively tightly clustered and very similar to those found in comparable saccadic reaction time tasks in previous studies (Carpenter and Williams 1995); \( \mu_2 \) and \( \sigma_2 \), the parameters for the detection unit are both larger, on average, and show considerably more variation between subjects.

**Altering prior probability of a switch**

We tested the predictive power of our model by systematically varying the prior probability of the appearance of the instruction (the central colored dot that signaled the switch trial). In different blocks of trials, the proportion of occasions on which the stimulus signaling task switching appeared was 10, 20, 50, 70, or 90%.

Interpreting the decision signal as a representation of log likelihood provides a particularly stringent and quantitative prediction. In this case, \( S_0 \) should represent the prior log probability \( P \); consequently, median latency, directly proportional to \( (S_T - S_0) \), should be reduced linearly in proportion to \( P \). In Fig. 4 we plot median latency for the Switch trial against log prior probability; as would be expected of a LATER unit, median latency—and thus \( S_T \)—does appear to vary linearly with log probability.

In addition to altering median latencies, the LATER model predicts that modifying the prior probability will produce a characteristic alteration in the distribution of reaction times (Fig. 1). More specifically, alteration of \( S_0 \) should cause both the median and slope of the reciprocit plot to alter, but the intercept \( I \) with the \( t = \infty \) axis should remain fixed, so that the distribution appears to “swivel” about this intercept (Fig. 1). Figure 4 shows two representative plots of switch trial distributions for different prior probabilities, together with best-fit lines constrained to a common intercept \( I \). For all subjects, the distributions were compatible with a common intercept (swivel), but not with the alternative, the parallel shift expected if the change in median were the result of a change in \( \mu \) (K-S one-sample test, with a criterion of \( P = 0.05 \)). Over all subjects, a log-likelihood analysis of these two alternative hypotheses yielded a support value of 9.8 log units for swivel rather than shift, equivalent to a conventional significance level of \( P < 10^{-4} \).

**Varying the preparation time**

A further prediction of our model is that increasing the instruction lead time (ILT, the interval between the presentation of the instruction and the stimulus) should produce a decrease in the switch trial latency that is equal to the ILT, until a critical value of the ILT is reached at which the reaction time for the Switch and Control trials is on average identical. For all subjects, latency in Switch trials did indeed fall in the expected manner. Figure 5 shows typical results for three subjects, together with predicted form of the relationship, the only free parameter being the critical time, chosen by minimizing sum of squares over each data set. For all five subjects tested, the...
change in the median latency of the Switch trials with a change in preparation time is not significantly different from the prediction for any value of the ILT; we also found that error rates decreased systematically with increased ILT (Fig. 5, bottom right). Finally, a point of some theoretical interest in relation to current ideas about task switching in general (see DISCUSSION) is the notion of residual cost: even when the preparation time is very long, Switch trials still exhibit latencies that are longer than controls. However, a comparison of mean latencies for Control and Switch trials with a 250-ms ILT showed no significant difference (paired $t$-test, $P = 0.99$), demonstrating that there is apparently no residual cost in this task.

**DISCUSSION**

In this study we have attempted to bring two approaches to the study of behavioral control that have previously been somewhat disparate: cognitive and relatively abstract on the one hand, and neurophysiological and relatively specific on the other. The neurophysiological approach has tended to use combined electrophysiology and psychophysical experimentation and has often targeted perceptual decisions (Shadlen and Gold 2004). Such experiments focus on simple tasks in which there is an intimate and accessible relationship between the sensory analysis and the motor intention. In contrast, psychologists have preferred to investigate more complex decisions in which this relationship is often much less apparent, and this work has led to the discovery of a variety of behavioral phenomena. However, it has offered little in the way of potential neural explanation. We propose that there is a need to expand the neurophysiological perspective and systematically tackle these observations—a good starting point seems be the decision to switch task.

The study of task switching has a long history (Jersild 1927) that has so far been exclusively from a psychological perspective. It has recently been cogently reviewed by Monsell (2003): essential points are that responses take longer to initiate on switch trials than on nonswitch trials, the difference being the task-switch cost (Rogers and Monsell 1995); that providing advance knowledge of the upcoming switch, allowing time to prepare, tends to reduce the average switch cost; and that even with very long preparation times, there is a residual cost: switch latencies are greater than nonswitch latencies. Rather little of this work has been specifically saccadic, of which most has been centered on a paradigm with switching between saccades made toward the stimulus (prosaccades) and in the opposite direction (antisaccades) (Barton et al. 2006; Cherkasova et al. 2002; Hallet and Adams 1980; Hodgson et al. 2004; Hunt and Klein 2002; Weber 1995), although this task is intrinsically unsatisfactory because of the inescapable asymmetry between these two types of responses.

**Task switching as decision**

The starting-point for the alternative approach is the realization that task switching implies a decision: identification of the instruction to switch. Tasks would then compete with one another just as stimuli do. It does not seem unreasonable to suggest that this decision process might be implemented in essentially the same way as other decision processes that have been previously studied (reviewed in the INTRODUCTION). More specifically, it is natural to imagine a model for task switching consisting of two LATER units in series, the first detecting the instruction, and then giving permission to a second unit that
behaviors are viewed as vectors and the decision procedure is
implemented by a model consisting of two LATER units in series.

Residual cost

Whatever its strengths, our model cannot provide a complete
account of the phenomena observed when we switch task, in
particular the perplexing phenomenon of residual cost. Rogers
and Monsell (1995) suggest that some part of switching cannot
occur until triggered by the task stimulus itself. An alternative
view (DeJong 2000) is that although subjects attempt switching
before stimulus onset, they succeed only in a proportion of
trials. However, it is important to recall that the task that we
investigated in this study does not in fact generate a residual
cost. This has been noted in previous saccadic studies: in a
situation in which the instruction—consisting of two vertical
or horizontal dots presented centrally, indicating a prosaccade
and antisaccade, respectively—were presented on every trial,
there was a mean increase in reaction time of 70 ms in mixed
blocks compared with controls (Weber 1995). This cost was
eliminated completely with an ILT of 300 ms. Hunt and Klein
(2002) similarly found that allowing sufficient preparation time
in a paradigm consisting of switching between pro/antisaccadic
switching could completely eliminate the switch cost.

Conceivably this lack of residual cost occurs because con-
tventional saccadic tasks are rather simpler than has tradition-
ally been the case in the general task switching literature, and
impose less of a burden on working memory. Hodgson et al.
(2004) found no cost in a relatively simple pro/antisaccadic
task, but in further experiments in which the subject had to
learn an arbitrary relation between the color of a central
stimulus and the direction of the saccade a substantial residual
cost was observed. If the switching process occurs by two
LATER units in series, then this suggests that the residual cost
is determined by a different mechanism that must cause a delay
to the input to one of these units; however, absence of a
residual cost in our study means that further speculation is
probably of little value.

Clarifying the boundary between tasks that do or do not
generate residual costs is clearly a priority, although perhaps it
is not surprising that highly complex tasks demand untidy
explanations. We suggest that studies on tasks designed to be
as simple as possible while still exemplifying the phenomenon
of interest, combined with models that make quantitative
predictions and are consistent with neurophysiological find-
ings, will together provide incisive tools for developing plau-
sible models for task switching and possibly many other
cognitive phenomena.

Acknowledgments

We thank our subjects for considerable patience during these long exper-
iments.

Grants

This work was supported by grants from the Caius Medical Association and
the Altounyan Fund.
REFERENCES


