The neural basis of the speed–accuracy tradeoff

Rafal Bogacz1, Eric-Jan Wagenmakers2, Birte U. Forstmann2 and Sander Nieuwenhuis3

1Department of Computer Science, University of Bristol, Bristol BS8 1UB, UK
2Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands
3Institute of Psychology and Leiden Institute for Brain and Cognition, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands

In many situations, decision makers need to negotiate between the competing demands of response speed and response accuracy, a dilemma generally known as the speed–accuracy tradeoff (SAT). Despite the ubiquity of SAT, the question of how neural decision circuits implement SAT has received little attention up until a year ago. We review recent studies that show SAT is modulated in association and pre-motor areas rather than in sensory or primary motor areas. Furthermore, the studies suggest that emphasis on response speed increases the baseline firing rate of cortical integrator neurons. We also review current theories on how and where in the brain the SAT is controlled, and we end by proposing research directions that could distinguish between these theories.

Many decisions are based on information that accumulates over time. Is the approaching animal a predator or a prey? Should I take the next exit on a busy highway? What restaurant do I want to dine at tonight? For such decisions, one faces a dilemma that is known as the speed–accuracy tradeoff (SAT; Refs. [1–3]). On the one hand, the probability of making a correct decision increases as information accumulates; on the other hand, it takes time to accumulate information, and slow decisions that are almost certainly correct may be of little value when you are being mauled by a predator, when you have missed your exit, and when you discover that all restaurants in town have just closed. These examples highlight the importance of striking a reasonable balance between the competing demands of speed and accuracy.

In the neurosciences this process has recently been studied in the context of perceptual tasks that are relatively simple but which retain the SAT dilemma present in more complex decisions [4–6]. In the moving-dots paradigm [7], for example, animals and humans have to decide whether a cloud of dots, visually presented in the middle of a computer screen, appears to move to the right or to the left. Due to noise in the stimulus, the dominant direction of motion is not immediately obvious. When put under time pressure, animals and humans are able to make such decisions relatively quickly, but at the cost of making more errors. When required to be accurate, they can reduce the probability of making an error by accumulating information for a longer period. What are the neural mechanisms that underlie these adaptive changes in performance?

For almost half a century SAT has been studied almost exclusively using abstract mathematical models. Although these models provide a good account of behavioral data, they are silent about the structures and mechanisms by which SAT is implemented in the brain. Only very recently have researchers begun to study the neural basis of SAT, using experimental methods and neurocomputational models. This review focuses on three key aspects of SAT that this recent work has helped bring to the fore: (i) what processing stage does SAT affect? (ii) does speed emphasis cause an increase in baseline activity or a decrease in response threshold? and (iii) how is SAT controlled in the cortico–basal ganglia circuit? Before we can adequately address these questions, however, we first need to review briefly the neural mechanisms of decision making and the mathematical models of SAT.

The neural basis of information accumulation

For simple perceptual decisions (e.g. does the stimulus move left or right?), neural correlates of information accumulation have been found in multiple cortical and subcortical brain areas. These areas include regions associated with the response modality – for example the superior colliculus and the lateral intraparietal area (LIP) in saccadic response tasks and the pre-motor cortex in motor response tasks – but also regions associated with high-level decision making such as the dorsolateral prefrontal cortex (DLPFC) [8–10]. In these areas, pools of neurons are selectively tuned to different perceptual choices such that their level of activation is indicative of the subject’s response [8]. During the interval between stimulus onset and response, these neurons gradually increase their firing rate, suggesting that they integrate the inputs from sensory neurons over time [8–10]. A response is initiated when the firing rate of the corresponding integrator neurons exceeds a certain critical threshold [11].

Mathematical models of SAT

The above neurophysiological mechanisms of decision making are strikingly consistent with models that have been produced in the field of mathematical psychology. These formal models seek to capture the error rate and the distribution of response times in various two-choice tasks [12–18]. Many of these models include abstract units or
accumulators that are assumed to gradually accrue noisy sensory input until a threshold of activation is reached [19,20]. More recent models proposed in the field of computational neuroscience explicitly model individual integrator neurons [21–23] or assume that each accumulator corresponds to a population of integrator neurons associated with a particular choice alternative [24,25]. Importantly, all these models assume that SAT is controlled by the distance between the initial activity of the integrators (i.e. the baseline) and the response threshold. If this difference is small, decisions are fast but error-prone; if the difference is large, decisions are accurate but slow (Figure 1). Therefore, these models predict that neural changes associated with SAT should be visible in brain areas involved in decision making (i.e. areas containing integrator neurons), rather than in areas specialized in stimulus encoding and motor execution.

Mathematical models can account for SAT in two ways: either by changing the baseline of accumulators or by changing their threshold (Figure 2). Although many modelling studies have assumed for simplicity that the baseline stays constant and that SAT is controlled by changing the threshold, in almost all mathematical models the crucial factor controlling SAT is the baseline–threshold distance. In most models an increase in baseline and a corresponding decrease in threshold produce exactly the same changes in simulated behavior. Thus, behavioral data alone do not allow one to distinguish whether SAT is controlled by changing the baseline, threshold, or both. To address this question, one needs to analyze neural activity.

fMRI studies of SAT: advantages and limitations

To date, the most direct evidence concerning the neural basis of SAT comes from three recent BOLD-fMRI studies in humans [4–6]. BOLD is an fMRI technique that reveals the local changes in blood oxygenation that are closely coupled with local increases in neural activation [26]. Compared to cell recordings in animals, human fMRI has distinct advantages as a method for studying SAT. First, unlike animals, human subjects can simply be instructed to be fast or accurate. Furthermore, fMRI permits whole-brain coverage at a spatial resolution sufficient to delineate regional changes in activation. Whole-brain coverage is important for studying phenomena, like SAT, that are likely to be dependent on the interplay between various brain areas.

Two general limitations of fMRI are its low spatial and temporal resolution. In contrast to neurophysiological recordings, fMRI does not allow one to monitor the activity of specific integrator neurons as evidence accumulates over time. In addition to these general limitations, fMRI studies of SAT are also confronted with two specific challenges. First, as pointed out by van Veen et al. [6], the amplitude of decision-related BOLD responses is not only proportional to the baseline–threshold distance but also to the duration of the decision process [26], and this duration is likely to be longer in the accuracy condition. Therefore it is hard to attribute changes in decision-related BOLD signals unequivocally to changes in baseline–threshold distance. Second, it is difficult to examine response thresholds with

Figure 1. An accumulator model account of SAT. The figure shows a simulation of a choice between two alternatives. The model includes two accumulators, whose activity is shown by blue lines. The inputs to both accumulators are noisy, but the input to the accumulator shown in dark blue has a higher mean, because this accumulator represents the correct response. Lowering the threshold (horizontal lines) leads to faster responses at the expense of an increase in error rate. In this example, the green threshold leads to a correct but relatively slow response, whereas the red threshold leads to an incorrect but relatively fast response.

Figure 2. Schematic illustration of changes in the activity of neural integrators associated with SAT. Horizontal axes indicate time, while vertical axes indicate firing rate. The blue lines illustrate the average activity of a neural integrator selective for the chosen alternative, and the dashed lines indicate baseline and threshold. (a) Accuracy emphasis is associated with a large baseline–threshold distance. (b,c) Speed emphasis can be caused either by increasing the baseline (panel b) or by lowering the threshold (panel c); in formal models, these changes are often mathematically equivalent.
fMRI because the amplitude of decision-related BOLD signals is measured relative to a baseline that, as we will review below, itself is also changing with differences in SAT; the careful measurement of threshold levels requires an absolute measure. These two specific problems with interpreting decision-related BOLD responses do not apply to the interpretation of sustained BOLD changes brought about by the SAT instruction. Indeed, the latter measurements have proved to be highly informative about the neural basis of SAT.

**What processing stage does SAT affect?**

The mathematical models discussed above suggest that SAT affects only the decision stage, and not stimulus encoding or motor execution. What is the neurobiological support for this claim? Three recent fMRI studies have significantly advanced our understanding of the brain areas involved in SAT [4–6; Box 1]. Each of the three studies manipulated SAT using visual cues that emphasized either response speed or accuracy. Despite differences in task, design, and analysis, all three studies report that sustained activity in the pre-supplementary motor area (pre-SMA) is increased when response speed is emphasized. This is in line with studies that show the pre-SMA is involved in the internal generation of action plans and in the anticipation of a motor response [27]. Furthermore, the three studies [4–6] found that speed cues increased activity in the striatum, the input nuclei of the basal ganglia. In general, the fMRI studies revealed increased baseline activity associated with speed emphasis in several decision-related associative and premotor areas. In contrast, none of the studies found SAT-related changes in sensory cortical areas or the primary motor cortex. In line with mathematical models, fMRI studies thus suggest that the modulation of SAT occurs in

---

**Box 1. Methodology used in fMRI studies on SAT**

Three recent fMRI experiments manipulated SAT using cues that instructed participants to respond either quickly or accurately [4–6]. The main challenge confronted by these studies was to distinguish direct effects of SAT instructions from indirect effects due to concomitant changes in the decision process. That is, the studies needed to disentangle the sustained change in neural activity associated with the SAT instruction from the changes in decision-related neural responses that are superimposed on this sustained change.

To address this challenge Forstmann et al. [4] presented a SAT cue (speed, accuracy, or neutral) at the start of each trial of a moving-dots task. Each cue was visible for 4.8 sec before the motion stimulus was presented. This relatively long period of cue presentation was introduced due to the poor temporal resolution of fMRI, and allowed the analysis to focus solely on the percent signal change brought about by the cue.

Ivanoff et al. [5] presented a cue that emphasized speed or accuracy, followed by a block of seven trials with a moving-dots task. All dots initially moved randomly. On some trials, the dots never started to move coherently. These trials were used to examine ‘baseline’ neural activity, because neural integrator responses were absent in this condition. In the remaining trials, a fraction of the dots started to move coherently 0, 4.5 or 9 sec after stimulus onset. These trials were used to examine the size of the transient BOLD responses associated with detecting motion-coherence under both levels of SAT. Thus, sustained SAT-related changes and corresponding decision-related changes in neural activity were distinguished by examining the fMRI response to motion displays with or without information.

In the study of van Veen et al. [6] each SAT cue was followed by a short mini-block of trials with a Simon interference task. The authors tried to account for both sustained and decision-related changes in BOLD signal during the mini-block, by using a hybrid general linear model with a sustained baseline regressor that accounted for the baseline activity throughout each mini-block, and a number of transient regressors (convolved with the hemodynamic response function) to account for decision making under both levels of SAT.

A second important methodological step involved correlating, across subjects, the obtained SAT-related fMRI effects with behavioral indices or model parameters of SAT. Forstmann et al. [4] calculated individual measures of response caution by fitting the linear ballistic accumulator model [19] to the behavioral data, and deriving the value of the baseline-threshold distance parameter. They then examined whether SAT-related changes in the value of this parameter were associated with changes in activation in the pre-SMA and striatum (Figure 3b). Ivanoff et al. [5] took a similar approach, but used a different estimate of baseline-threshold distance – the criterion metric of signal-detection theory. Finally, van Veen et al. [6] correlated sustained SAT-effects on BOLD signal intensities with a behavioral index of SAT – the difference in error rates between speed and accuracy conditions.

---

Figure 3. Changes in pre-SMA activity due to SAT manipulations. (a) The pre-SMA exhibited increased baseline activity under speed emphasis in the fMRI study by Forstmann et al. [4]. (b) Increases in pre-SMA baseline activity in the speed versus the accuracy condition showed a significant negative cross-subject correlation with an individual measure of response caution: the baseline-threshold distance estimated from the linear ballistic accumulator model, a mathematical model of decision making [19]. Thus, those participants who showed a relatively large decrease in response caution also showed a relatively large increase in baseline activity in the pre-SMA.
decision-related rather than in early sensory or primary motor areas.

**Does speed emphasis increase baseline or decrease threshold?**

In mathematical models of decision making, an increase of the baseline and reduction of the threshold are often assumed to be equivalent (Figure 2). However, as discussed below, human brain-imaging studies and neurophysiological recordings may provide an answer to the question of whether in the brain SAT is accomplished by changing the baseline or the threshold of integrator neurons. Because the changing-baseline and the changing-threshold hypotheses are not mutually exclusive, we will discuss the relevant evidence for both these scenarios.

The recent fMRI studies of SAT provide strong evidence for the changing-baseline hypothesis. Contrasts between the speed and accuracy conditions revealed many brain areas with increased sustained BOLD activity during the speed condition and very few or no areas showing the opposite pattern. For example, Ivanoff et al. [5] and van Veen et al. [6] found increased activity with speed emphasis in several frontal and parietal areas, including the DLPFC [6]. Similarly, Forstmann et al. [4] found increased activity in the pre-SMA (Figure 3) and striatum during the interval between a cue instructing the participants to be fast and stimulus onset. Under the assumption that the observed BOLD signal in these areas is produced by the activity of integrator neurons, the data from these three fMRI studies suggest that speed instructions increase the baseline activity of these neurons.

We are not aware of any neurophysiological studies in which animals were trained and instructed to trade accuracy for response speed. However, several studies have yielded evidence that is consistent with the changing-baseline hypothesis. Preparatory activity in the monkey parietal reach area was found to predict trial-to-trial fluctuation in reach reaction time [28]. This is consistent with the changing-baseline hypothesis under the assumption that a substantial portion of such response-time fluctuations are due to trial-to-trial changes in boundary separation [29,30], meaning that subjects spontaneously modify SAT between trials (but note that alternative explanations of the fluctuations in baseline are possible; Refs. [15,31]). Another study examined the effect of instructed movement speed on monkeys’ preparation to reach manually for a target [32]. Prior to the actual movement, neural activity in premotor and motor cortex showed an increase in baseline activity following the fast-speed instruction compared to the slow-speed instruction.

Unlike for the changing-baseline hypothesis, there is no strong evidence for the changing-threshold hypothesis. One fMRI study reported results consistent with a fixed threshold across different levels of SAT [5]. This study found that the increased baseline activity with speed emphasis in the pre-SMA was associated with a roughly equal decrease in decision-related transient BOLD responses. Furthermore, participants who showed a larger increase in baseline activity also showed a larger decrease in decision-related activity in the pre-SMA. However, a similar inverse relationship was not found for another decision-related area, the lateral PFC, which led Ivanoff et al. [5] to suggest that in the lateral PFC baseline changes were accompanied by threshold changes.

As mentioned above, results from studies that have examined trial-to-trial variability in response times may provide additional, yet indirect, clues regarding the neural basis of SAT. These studies have shown that different response–time bins are associated with a fixed activity level at the time of the response in the lateral intraparietal area (LIP) [11], frontal eye field [33] and superior colliculus [34]. This suggests that integrator neurons in these areas accumulate evidence until they reach a fixed threshold. Studies on event-related potentials in humans have yielded very similar results: at the time of response onset, voltage potentials reflecting the accumulation of evidence in the motor cortex are at the same level for different response–time bins [35,36].

Taken together, the results from brain-imaging and neurophysiological studies suggest that SAT is implemented by changes in the baseline activity of integrator neurons. More evidence is necessary to establish whether the threshold also changes between emphasis on speed and accuracy.

**How is SAT controlled in the cortico–basal ganglia circuit?**

Mathematical models of decision making are silent about the neural structures and processes that are ultimately responsible for how people control SAT. This section reviews four recent theories on how SAT is controlled in the cortico–basal ganglia circuit (Figure 4) involved in action selection (Box 2). We refer to these theories as

---

**Figure 4.** Schematic representation of the cortico–basal-ganglia–thalamic circuit. STN corresponds to the subthalamic nucleus, and Output corresponds to the substantia nigra pars reticulata and the external segment of globus pallidus in primates, or its rodent homolog the entopeduncular nucleus. Green arrows denote excitatory connections; red lines with filled circles denote inhibitory connections. Blue arrows indicate the areas where the input controlling SAT could be provided, according to the different theories described in the text.
Box 2. Mechanism of action selection in the basal ganglia

The basal ganglia are a set of subcortical nuclei involved in action selection [47]. In a default state the output nuclei of the basal ganglia send inhibition to the thalamus and other subcortical areas and thus block the execution of any movements [48]. In order to initiate a movement, the cortical integrators send input to the striatum, that in turn inhibits the output nuclei (Figure 4). The resulting reduction in the activity of the output nuclei releases the subcortical areas from inhibition, and thus allows the movement to be executed [48].

The above process is additionally modulated by the subthalamic nucleus. This sends excitatory projections to the output nuclei, and hence its activity can block or postpone action execution [49]. It has recently been proposed that the modulation provided by the subthalamic nucleus allows longer and more accurate decisions on difficult trials [50,51].

‘cortical’, ‘striatal’, ‘STN’, or ‘synaptic’, indicating the part of the cortico–basal ganglia circuit that modulates SAT (STN, subthalamic nucleus).

The cortical theory proposes that with speed emphasis, cortical integrators receive additional excitatory input, thereby increasing their baseline activity. The cortical theory was proposed independently in several articles, two of which were simulation studies that demonstrated that a common input to cortical integrators can control SAT efficiently [21,37]. In the studies by van Veen et al. [6] and Ivanoff et al. [5] the theory was motivated by the observation that speed emphasis increases the baseline activity in decision-related cortical areas. Van Veen et al. [6] reported evidence suggesting that the input to integrators controlling SAT is provided by the DLPFC.

The striatal theory suggests that, with speed emphasis, the striatum receives excitatory input from cortical (non-integrator) neurons, this increases striatal activity and thus lessens the inhibitory control of the basal ganglia over the brain, and thereby facilitates faster but possibly premature responses [4]. This theory was motivated by the increase in striatal activity with speed emphasis observed in Refs. [4,6]. Forstmann et al. speculated that the input to the striatum controlling SAT could be provided by pre-SMA [4]. Under the cortical and striatal theories, the DLPFC and pre-SMA, respectively, are involved in the control of SAT rather than in the integration of sensory evidence – the increased activity in these areas with speed emphasis should therefore be interpreted as an increased control signal that modulates cortical and striatal activity.

The STN theory proposes that with accuracy emphasis the subthalamic nucleus (STN) receives additional excitatory input from frontal areas [38]. Increased STN activity produces slower and more accurate choices (Box 2). This theory was motivated by the observation that when participants attempt to stop initiated movements, this increases activity in the STN and in frontal areas that project to the STN, such as the pre-SMA and the inferior frontal cortex [39,40]. Frank et al. speculated that the input to STN controlling SAT could be provided by the anterior cingulate cortex [38], an area thought to be involved in monitoring response conflict [41]. When response conflict increases, the anterior cingulate cortex promotes, via the STN, a more careful mode of responding.

The synaptic theory suggests that speed emphasis strengthens the synaptic connections between cortical integrators and striatal neurons [42]. This reduces the level of activity of cortical integrators required to trigger a decision, and hence reduces the threshold. The synaptic theory was proposed in a different context than the previous theories, namely for tasks in which subjects need to discover the balance between speed and accuracy that optimizes a performance criterion such as reward rate [43]. The synaptic theory was motivated by reinforcement–learning theories that postulate that cortico–striatal synapses are modified during learning processes aiming at reward maximization [44]. However, Furman and Wang point out that the modification of synaptic weights is a gradual and relatively slow process, and hence the mechanisms suggested by the other three theories would enable more rapid, and thus more adaptive control of SAT [21].

The four theories rely on a similar mechanism: speed emphasis causing a change in activity in one of the elements of the cortico–basal ganglia circuit. Consequently, all theories make similar predictions. In particular, because of the strong functional projections from the basal ganglia back to the cortex via the thalamus [45] (Figure 4), all four theories predict that speed emphasis increases the baseline of cortical integrators. For example, a purported reduction of STN activity with speed emphasis will decrease the inhibitory activity of the basal ganglia output nuclei, increase the activity in the thalamus, and consequently increase the activity of the cortical integrators. An increase in striatal activity – predicted by the striatal and synaptic theories – will increase the activity of cortical integrators by the same route (thus the striatal and synaptic theories predict both an increase in the baseline and a decrease in the threshold with speed emphasis). Analogously, all four theories predict an increase in striatal activity with speed emphasis. Because all four theories are consistent with the increased activity in the striatum and cortical areas (that presumably contain integrator neurons) observed in the recent fMRI studies, we believe that the present data do not provide a pivotal test between these theories.

The four theories do, however, make different predictions that could be tested in future experiments. First, the STN theory predicts that, before stimulus onset, speed emphasis decreases STN activity, whereas the other three theories predict that speed emphasis increases STN activity – due to increased input from cortical integrators. Second, the cortical, striatal, and STN theories claim that SAT is controlled by external input, but differ in the brain areas hypothesized to receive this input. Thus, these theories could be distinguished by single-unit recordings in cortical integrator areas, the striatum, and the STN, and by comparing the onset latencies of cue-induced changes in these three areas.

Proponents of the cortical and striatal theories have speculated about the origin of the input controlling SAT and have pointed to some areas that show increased activity with speed emphasis, most notably the DLPFC and the pre-SMA. However, the increased activity in the DLPFC and the pre-SMA is also consistent with the possibility that these areas include integrator neurons whose baseline is increased. It is generally difficult, on the basis of existing fMRI data, to determine whether an
area controls SAT, is being controlled, or includes both neurons involved in integration of information and neurons responsible for the control of SAT. Several possible experiments could address this issue. First, neurons exclusively involved in controlling SAT could be identified with single-unit recordings: these should increase their firing rate following speed or accuracy instructions and maintain this activity level until movement initiation. In contrast, integrator neurons are expected to increase their firing rates gradually after stimulus onset. Second, the interactions between different areas involved in SAT could be investigated with functional connectivity methods such as dynamic causal modeling – an fMRI analysis method that was developed to determine which areas provide and which receive the input [46].

Conclusion and outlook

In this article we have reviewed recent developments in neuroscience that shed light on how the brain implements SAT, allowing decision-makers to negotiate the competing demands for response speed and response accuracy. The existing experimental data show that the modulation of SAT occurs in association areas and the pre-SMA rather than in early sensory and primary motor areas. Furthermore, the data strongly suggest that emphasis on speed is associated with an increase in baseline activity of cortical integrator neurons. Current observations are consistent with at least four different theories of how SAT is controlled and implemented in neural decision circuits. More diagnostic information is likely to result from experiments suggested in the previous section, involving single-cell recordings and causal network analysis of fMRI data.

In their 2007 article on the neural basis of decision making, Gold and Shadlen [8] pointed to an important question for future research: ‘How and where in the brain is the decision rule (e.g. a bound crossing) implemented?’ (p. 563). In only one or two years since then, several studies have already been published that address exactly this question. We now know more about how the brain controls and implements SAT. However, the main contribution of the SAT studies to date has been to raise questions that are more precise, such as those discussed in this article (Box 3). This holds out hope that future work will be able to address the pivotal issues that the current work has identified.

Box 3. Outstanding questions

- Does emphasis on speed change the threshold of cortical integrator neurons?
- Which brain area provides the input controlling SAT? And to which area in the cortico–basal-ganglia circuit is this control input provided?
- The recent fMRI studies have almost exclusively focused on how the brain achieves greater response speed. Are there separate neural mechanisms responsible for achieving greater accuracy?
- Are individual differences in SAT mediated by structural differences in brain areas, such as the STN or the pre-SMA, or their connectivity?
- Is SAT driven by a single underlying neural system, or are there more? (Ref. [52])
- What drives the changes in SAT that are observed as people grow older? (Ref. [53]).

Acknowledgement

This research was supported by VENI and VIDI grants from the Netherlands Organization for Scientific Research.

References