Review

Cite this article: Schall JD, Palmeri TJ, Logan GD. 2017 Models of inhibitory control. *Phil. Trans. R. Soc. B* **372**: 20160193.
http://dx.doi.org/10.1098/rstb.2016.0193

Author for correspondence:
Jeffrey D. Schall
e-mail: jeffrey.d.schall@vanderbilt.edu

Models of inhibitory control

Jeffrey D. Schall, Thomas J. Palmeri and Gordon D. Logan

Center for Integrative and Cognitive Neuroscience, Vanderbilt Vision Research Center, Department of Psychology, Vanderbilt University, PMB 407817, Nashville, TN 37240-7817, USA

JDS, 0000-0002-5248-943X; TJP, 0000-0001-7617-9797

We survey models of response inhibition having different degrees of mathematical, computational and neurobiological specificity and generality. The independent race model accounts for performance of the stop-signal or countermanding task in terms of a race between GO and STOP processes with stochastic finishing times. This model affords insights into neurophysiological mechanisms that are reviewed by other authors in this volume. The formal link between the abstract GO and STOP processes and instantiating neural processes is articulated through interactive race models consisting of stochastic accumulator GO and STOP units. This class of model provides quantitative accounts of countermanding performance and replicates the dynamics of neural activity producing that performance. The interactive race can be instantiated in a network of biophysically plausible spiking excitatory and inhibitory units. Other models seek to account for interactions between units in frontal cortex, basal ganglia and superior colliculus. The strengths, weaknesses and relationships of the different models will be considered. We will conclude with a brief survey of alternative modelling approaches and a summary of problems to be addressed including accounting for differences across effectors, species, individuals, task conditions and clinical deficits.

This article is part of the themed issue 'Movement suppression: brain mechanisms for stopping and stillness'.

1. Introduction

The collaboration between mathematical psychology and neuroscience is yielding amazing insights into the mind and brain that were unimaginable 20 years ago. Mathematical psychology provides precise, explicit descriptions of mental processes that are linked tightly to behaviour, making strong predictions about behaviour that stand up to rigorous empirical tests. Accurate prediction of response time (RT) distributions for correct and error responses is now commonplace, and it is the standard by which models are judged. Neuroscience has opened the black box and shown us how the neural processes underlying behaviour unfold in real time. Analyses of spike trains, local field potentials and EEG reveal the time-course of encoding, transforming and acting. Information from anatomical connectivity and effects of lesion or inactivation reveal microcircuits and networks contributing to the production and withholding of movements. In recent years, we have seen a proliferation of theories merging insights from mathematical psychology and neuroscience through the mapping of the computations articulated by the mathematical models with the activity of individual or systems of neurons that implement the computation. Such mapping is tested rigorously by requiring quantitatively accurate fits or accounts of both behavioural and neural data. This mapping specifies linking propositions that connect the mathematical description to neurons, groups of neurons or brain regions [1,2]. The linking propositions identify the points of contact between theory and neural data, and specify the aspects of the data that are relevant to the theory.

In the domain of response inhibition, the converging constraints have offered particularly clear and compelling mapping between abstract formal computational concepts and detailed neural implementation. Response inhibition has been investigated with the greatest theoretical leverage through the stop-signal
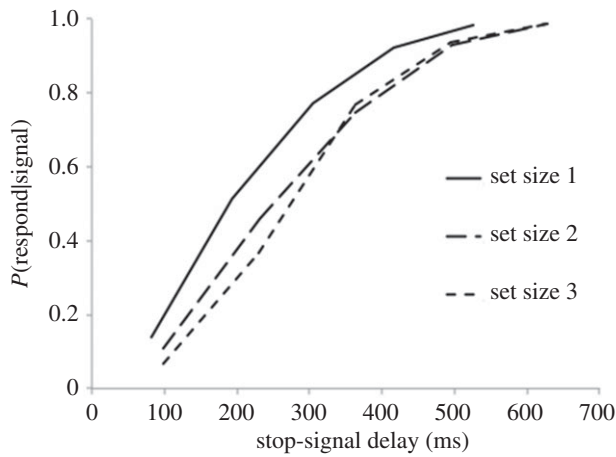
Figure 1. Inhibition functions from a memory-search experiment in which response time varied with the number of items in the memory set. The probability of responding given a stop signal increases as stop-signal delay increases and decreases as RT in the Go task increases as set size increases.

Figure 2. Probability ($RT < t$) of RT on no-stop-signal trials and on non-cancelled (signal respond) error trials with stop-signal delay (SSD) equal to 231, 364 and 496 ms. RT on non-cancelled trials was systematically shorter than RT on no-stop-signal trials and increased with SSD, beginning at common minima and ending with shorter maxima.
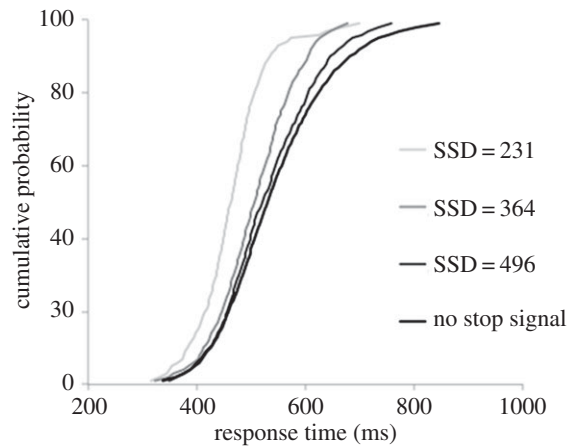
or countermanding task [3]. Performance can be explained in nuanced detail by an abstract theory based on the mathematics of race finish times.

For the stop-signal task, a theory of response inhibition should provide a quantitative account of the probability of inhibiting a response and explain how it varies with the time available to interrupt response preparation (stop-signal delay, or SSD). The theory must provide a quantitative account of the probability of successfully inhibiting a response (known as *signal inhibit* or *cancelled* trials) and the distributions of RT on trials with no stop signal and when inhibition fails (known as *signal respond* or *non-cancelled* trials).

2. Response inhibition in the stop-signal paradigm

The ability to inhibit our responses voluntarily is a paradigm case of cognitive control. It reveals itself in many behavioural paradigms, but it is revealed most directly in the stop-signal paradigm [4,5]. In this paradigm, subjects perform a 'go' task, in which they make a speeded response to an imperative stimulus. On some trials, a 'stop signal' is presented that tells subjects to inhibit their response to the Go signal. Whether or not they are able to is the main datum of interest. Many studies show that the ability to inhibit responses is probabilistic, and the probability of inhibition depends primarily on SSD (figure 1). SSD controls the amount of time available to detect the stop signal and countermand the Go response before it is executed; response inhibition is more likely when more time is available. Non-cancelled trial RT is also an important datum. It is usually faster than RT on trials with no stop signal, as if it comes from the faster tail of the Go RT distribution (figure 2).

These effects have been observed in several species, including rats, monkeys and humans, in several subject populations, including children, adolescents, young adults and the elderly. These effects have been observed in several psychiatric disorders, including attention deficit hyperactivity disorder and schizophrenia, and in several neurological disorders, including stroke and Parkinson's disease. They have been observed in different stimulus and response modalities, in different experimental conditions, and with different strategies. The patterns

are the same qualitatively, but they differ quantitatively, and the quantitative differences reveal important changes or deficits in cognitive control.

3. Independent race model

Two facts led to the independent race model of stop-signal performance: (i) the probability of response inhibition depends on the time available to detect the stop signal before the Go response is executed and (ii) non-cancelled RTs are faster than RTs on no-stop-signal trials. These characteristics are found in the large majority but not all studies of countermanding [6,7]. These two facts suggested that response inhibition depends on the outcome of a race between a GO process, initiated by the Go stimulus, and a STOP process, initiated by the stop signal. If the STOP process finishes before the GO process, the response is inhibited, producing a cancelled trial. If the GO process finishes before the STOP process, the response is not inhibited, producing a non-cancelled trial. The model assumes that the finishing times for the STOP and GO processes are independent random variables, and demonstrates that the fundamental results in the stop-signal paradigm follow from these assumptions (figure 3).

The independent race model provides a measure of the latency of the STOP process, called stop-signal reaction time (SSRT). This is a powerful virtue because the STOP process is not directly observable. If the STOP process finishes before the GO process, there is no response whose latency can be measured. If the STOP process finishes after the GO process, we know SSRT must have been longer than non-cancelled RT, but we do not know how much longer. The independent race model provides several converging methods for estimating SSRT from the observed data. These measures of SSRT have been important in documenting differences in the ability to inhibit responses across lifespan development, between clinical and control groups, and between neurological patients and controls. They have also been important in understanding the neurophysiology of response inhibition. Neural processes that cause response inhibition must modulate before SSRT; neural processes that are consequences of response inhibition modulate after SSRT.

3
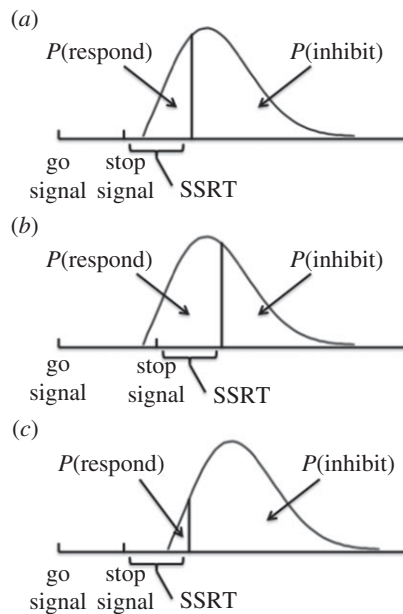
Figure 3. Predictions of the independent race model, assuming SSRT is constant. Onset of go signal followed by onset of stop signal. Vertical line in the distribution represents the finishing time of the STOP process. Probability of responding in spite of a stop signal is the area under the distribution to left of line. Probability of inhibiting in response to a stop signal is the area under the distribution to right of line. (a) Stop signal occurs before earliest RT, resulting in larger ratio of P(inhibit) to P(respond). (b) Here, stop-signal delay is increased, so probability of responding is increased. (c) Here, Go response time is increased, so probability of responding is decreased.

The independent race model has been used in virtually every stop-signal experiment. It provides important measures of cognitive control, like SSRT, and it provides a benchmark against which other models can be evaluated. Its prevalence results from its generality. It is formulated in terms of generic finishing time distributions for the STOP and GO processes. It makes no commitment to the underlying computational or neural processes that generate these finishing times. It expresses relationships that must hold for any and all distributions, regardless of the process that generates them. This is important because the independent race model provides an important check for the models we consider here that address the computations performed by the underlying neural processes: these models must predict the empirical relationships predicted by the independent race model.

The independent race model captures the essence of computation but not the details of implementation. It formulates the constraints that any model of response inhibition must follow, but it does not provide the structure necessary to explain recent developments in stop-signal research. For example, many studies have shown that Go RT is slower when stop trials occur more frequently, as if the GO process changes to balance the competing demands of stopping and going. Many other studies have shown that Go RT is slower on trials following stop signals than on trials before them, suggesting that a stop trial results in some kind of strategic adjustment to the GO process. To explain how these adjustments occur, we need a more detailed model of the GO process that tells us which parts can support this strategic adjustment. The independent race model provides no model of the underlying process. It can describe these effects, but it cannot explain them.

4. Neurophysiological principles

Hanes & Schall [8] showed that macaque monkey performance of a saccade stop-signal task corresponded in detail to that of human performance of manual stop-signal tasks. This correspondence was subsequently established for human stopping of saccades [9,10] and pursuit [11–13]. The probability they would inhibit their eye movements depended on SSD and their non-cancelled RTs were faster than their no-stop-signal RTs. Hanes et al. [14] recorded from frontal eye fields in monkeys performing the saccadic stop-signal task, isolating neurons involved in gaze-shifting and gaze-holding that represent a larger circuit of such neurons that extends from cortex through basal ganglia and superior colliculus to brainstem [15]. They found that these neurons modulated on stop-signal trials, just before SSRT when the monkey stopped successfully. These results have been replicated in the superior colliculus [16] and in FEF of monkeys doing related tasks [17,18] as well as in the basal ganglia of rats [19,20]. These rich neurophysiology datasets offer the opportunity to develop a model of response inhibition that explains how the computational constraints of the independent race model are implemented.

5. Interactive race model

The juxtaposition of the independent race model and observations of pronounced neural modulation at the moment of SSRT suggested a paradox: how can interacting circuits of mutually inhibitory gaze-holding and gaze-shifting neurons instantiate STOP and GO processes with independent finishing times? An interactive race model has been formulated [21] and extended [22–25] to address this paradox, which may also explain other kinds of response inhibition such as antisaccades [26]. The STOP and GO processes were instantiated as stochastic accumulators within various model architectures with lateral or feed-forward inhibition (figure 4). The GO accumulator begins after an afferent delay, D_{go}, accumulating activation until it reaches a threshold, whereupon a response occurs. The STOP accumulator begins after an afferent delay, D_{STOP}, inhibiting the Go response in proportion to its activation. If the STOP accumulator becomes active soon enough (if $SSD + D_{STOP} < RT_{go}$), it prevents the GO accumulator from reaching threshold and the response is inhibited. If the STOP process becomes active too late (if $SSD + D_{STOP} > RT_{go}$), the GO accumulator reaches threshold and the response is not inhibited.

The models are tested through computer simulations of particular sets of stochastic differential equations that govern the STOP and GO accumulators. The simulations were fit to behavioural data from two monkeys, who also provided neural data from the same test sessions. Simulations that could not fit well the inhibition function and distributions of RT were excluded. Two mechanisms whereby the STOP unit interrupts the GO process have been explored. The first involved inhibition from the STOP unit onto the GO unit, a natural extension of the lateral inhibition well known to be crucial for motor control. The second involved interrupting the drive on the GO unit with leak attenuating the GO unit activation [3,27]. The premise is that Go responses are driven by input from perceptual systems, and responses can be countermanded by blocking the input to the motor system. The input can be blocked in several ways. One possibility is
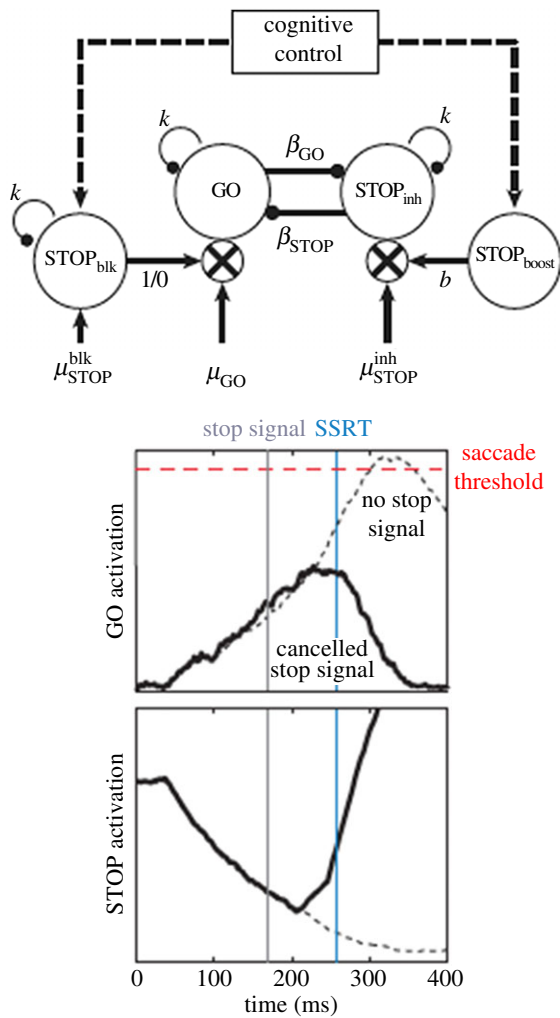
Figure 4. Interactive race model. The GO unit receives input after an afferent delay (D_{GO}) and the STOP unit receives input after stop-signal delay (SSD) plus an afferent delay (D_{STOP}). GO and STOP units inhibit each other. Inhibition from STOP to GO is much greater than inhibition from GO to STOP. A Go response occurs if GO activation reaches threshold. The Go response is inhibited if inhibition from the STOP unit prevents it from reaching threshold. The top panel diagrams the architecture of the circuit: excitatory connections have arrowheads, and inhibitory connections have circle terminators. Activation of a GO unit, driven by an input (μ_{GO}), accumulates with leak (k) to specify whether and when a saccade will be initiated. A variety of alternative mechanisms can interrupt GO unit activation. All of these mechanisms instantiate delayed potent inhibition, allowing a network of interacting units to produce behaviour that can be described as the outcome of a race between stochastically independent processes. Inhibition (β_{STOP}) from a STOP unit driven by an input (μ_{STOP}) can reduce activation of the GO unit, and this inhibition can be boosted by a cognitive control signal that potentiates the activation of the $STOP_{inhibit}$ unit. Alternatively, activation of a STOP unit can interrupt or block the input to the GO unit. The inputs to the $STOP_{block}$ (μ_{STOP}^{block}) and $STOP_{inhibition}$ ($\mu_{STOP}^{inhibition}$) units are distinguished because they assumed different values. These alternative architectures draw attention to the flexibility and adaptability of countermanding behaviour afforded through cognitive control. The bottom panel illustrates the activation of the GO unit (upper) and the STOP unit (lower) for trials with no stop signal (dashed lines) and trials with a stop signal that successfully cancelled the saccade (solid lines). Saccades are produced when inhibition of the STOP unit is released and the activation of a GO unit reaches a threshold (red dashed line). In response to the stop signal (solid grey line), the STOP unit becomes active, interrupting the accumulation of GO unit activation. This interruption occurs immediately before the stop-signal reaction time (SSRT) (blue line), a measure of STOP process duration derived from the independent race model.

deleting the goal to act. In production system models, action depends on two conditions: a goal and an appropriate stimulus. The action can be countermanded by removing the goal, by removing the stimulus, or by removing both. Another way to countermand responses is to suppress the input from perceptual systems. In stochastic accumulator models, this involves setting the drift rate to zero (or less). A third possibility is to break the connection between perceptual and motor systems. The mapping of Go stimuli onto Go responses is often arbitrary (e.g. 'press the left key if an X appears') and must be maintained somewhere in the cognitive system [28]. Disabling the mapping rules would prevent the growth of activation in the motor system. For example, in our model of visual search [29,30] the connection between perceptual and motor activity is controlled by a gate that prevents noise from accumulating in stochastic accumulators. Responses could be countermanded by raising the gate to a much higher level [31]. Both model architectures fit the performance data, providing accurate quantitative accounts of the inhibition function, no-stop-signal RTs, and non-cancelled RTs across SSDs (see fig. 6 in [21]).

A further critical test of the models was based on the trajectories of activation of the GO and STOP accumulators that were simulated using the parameters that produced the best fits to the performance data. To assess the match between simulated activation dynamics and recorded neural activity, specific aspects of the recorded activity were assessed. The activation of individual neurons has many idiosyncrasies, but all patterns show some general characteristics. The key feature was the distribution of cancel times, which are the times at which neural activity modulates on trials on which subjects stop successfully, relative to SSRT. This was measured in the simulated data in the same way that it was measured in the neural data, by determining the point at which activation on successful stop trials first differed significantly from activation on latency-matched no-stop-signal trials. In the neural data, this point ranges from 50 ms before to 50 ms after SSRT, with a mean 5–10 ms before SSRT. The model predicted distributions with the same range (see fig. 7 in [21]). Note, these are genuine predictions. They were generated with a fixed set of parameters that provided the best fit to the behavioural data, without any further adjustment to optimize the fit to neural data. Thus, these interactive race models are computationally explicit, explaining core features of the necessary computational and underlying neural processes.

What about the paradox? The interactive race model affords investigation of the dynamics of possible interactions between gaze-holding and gaze-shifting neurons that control eye movements. How can it fit performance data just as well by the independent race model? In short, the interaction of the STOP with the GO unit must be late and potent—late to preserve the independence of the GO and STOP processes through SSRT and potent because it must be late. In other words, D_{STOP} must be almost as long as SSRT, and the interruption of the GO unit by the STOP unit must be strong whether through powerful lateral inhibition or rapid leak after interruption of the GO unit input.

6. Interactive race with spiking neuron networks

The lateral inhibition interactive race model was subsequently implemented in a recurrent network model consisting of
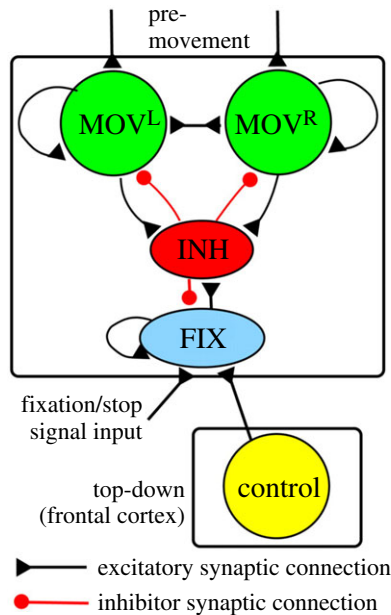
Figure 5. Spiking neuron model. Two pools of units producing different movements (MOV^L and MOV^R) are excited by input from a Go signal. Another pool of units preventing movement (FIX) is excited by input from the fixation point, which is the stop signal. A pool of inhibitory units (INH) mutually inhibits the MOV and FIX units. A control unit excites the FIXATION unit. A Go response occurs when the control unit releases excitation on the FIXATION unit. The Go response is inhibited if the stop signal excites the FIXATION unit before a MOVE unit makes its response.

spiking movement (GO) neurons and fixation (STOP) neurons (figure 5) [32]. The model assumes hundreds of units representing populations of movement neurons, fixation neurons and inhibitory interneurons, and a control unit that turns the fixation neurons on and off. Each population produces Poisson spike trains that depend on the ratio of parameters representing NMDA and AMPA inputs. The model addresses fixation activity at the beginning of a trial and the transition from fixation to movement as well as the rise in movement activation to threshold. The model produces the transition from fixation to movement, and ultimately RT, by turning off the control unit that excites fixation units, thereby releasing tonic inhibition on the movement units and allowing their activity to rise to threshold.

This model fixed the number of units and many of the parameters across all conditions and manipulated just three parameters to maximize goodness of fit—the mean and standard deviation of a Gaussian distribution for the time at which the control unit turned off and the time at which the stop signal turned the fixation units back on (analogous to D_{STOP} described above). Unusual for this neural level of modelling, parameters were found to fit the inhibition function and RT distributions for no-stop-signal and non-cancelled trials. As observed in the simpler interactive race models, Lo *et al.* [32] found that D_{STOP} had to be relatively long to produce appropriate non-cancelled RT distributions; again, inhibition of stop on GO had to be late and potent. This complex network model predicted the modulation of movement and fixation neurons and the cancel time distributions qualitatively as well as the original interactive race model, although these predictions were not assessed quantitatively.

The additional complexity of this model provided additional insights. For example, variation in baseline activation of fixation and movement units affected the probability of successful inhibition. Successful inhibition was less likely when movement units were more active during the baseline period and was more likely when fixation units were more active. They confirmed this prediction in the data from the original frontal eye field [14] and superior colliculus [16] countermanding studies, finding lower baseline firing rates in movement neurons prior to successful inhibition. In retrospect, differences in baseline activation could be implemented in the simpler interactive race model architectures and would probably produce similar results.

Although this spiking network interactive race model fit the performance data, replicated the neural trajectories, and made a novel prediction, the model has two limitations to consider. First, RT depends on turning off a control unit that tonically excites fixation units, which releases inhibition on movement units and allows their activity to rise to the trigger threshold. The variability in RT depends almost entirely on the variability in the time at which the control signal is turned off, which is determined by a Gaussian distribution whose mean and standard deviation were free parameters that were adjusted to optimize goodness of fit. The model provides no explanation for the mechanism producing this Gaussian distribution. Therefore, the control unit resembles a homunculus endowed with the ability to intervene at the right time to produce the right effect. Future work should specify more completely the computational and neural mechanism of this control element [33].

Second, the spiking network model required many parameters (e.g. AMPA and NMDA ratios for each interaction between units) in addition to the three parameters that were varied to optimize goodness of fit. These parameters were fixed for the fit, but they were tweaked to produce firing rates in the desired range for movement and fixation cells before they were fixed. From the perspective of mathematical psychology, where fitting large amounts of data with a small number of parameters is desirable, this is inelegant. From the perspective of neuroscience, where accounting for the details of implementation is the goal, this is necessary. We look forward to further research in which additional constraints, such as neuropharmacological effects on countermanding performance, might be integrated in spiking network models. Another constraint not applied in this spiking network model is the actual anatomical details of the brain regions engaged in response inhibition.

7. A brain circuitry model

A more general level of modelling incorporates connectivity between distinct brain structures such as the cerebral cortex and basal ganglia [34,35]. This brain network model describes interactions between units in frontal cortex (dorsolateral prefrontal cortex, right inferior frontal gyrus, frontal eye fields), basal ganglia (striatum, globus pallidus external segment, substantia nigra pars compacta, substantia nigra pars reticulata, subthalamic nucleus), and superior colliculus (figure 6). It addresses the stop-signal task and an anti-saccade task in which a peripheral target is presented, and subjects must inhibit their natural tendency to look directly at it and shift their gaze to a position opposite to it. The model explains stop-signal performance by assuming that the stop signal activates right inferior frontal gyrus, which activates subthalamic nucleus, which activates substantia nigra pars reticulata, which then inhibits superior colliculus. If the superior
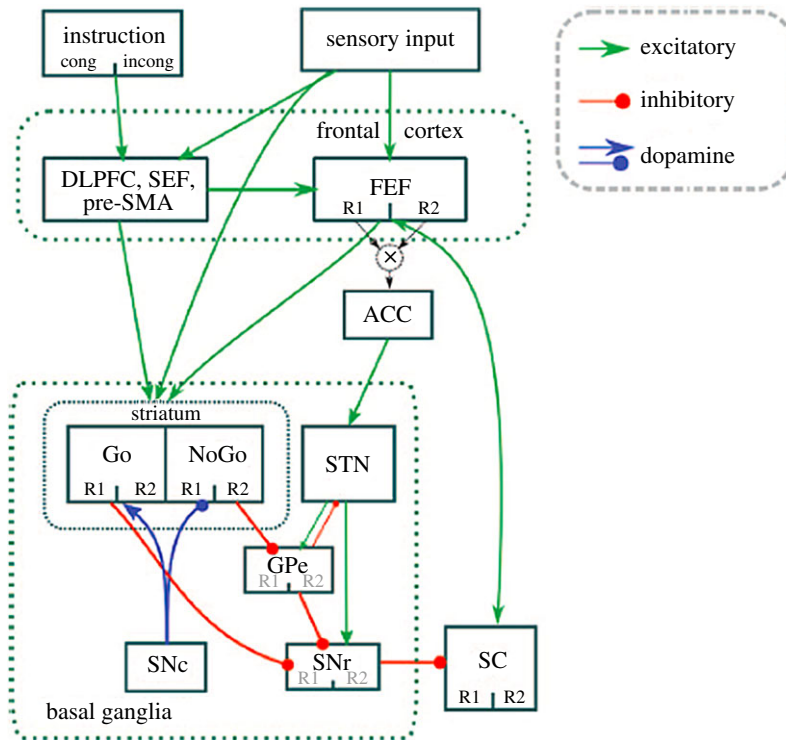
Figure 6. Cortico-basal ganglia network model. The model is designed to produce either of two responses (R1, R2). A sensory input layer excites the frontal eye fields (FEF), striatum, plus dorsolateral prefrontal cortex (DLPFC), supplementary eye fields (SEF) and pre-supplementary motor area (pre-SMA). Through excitatory projections from FEF to the superior colliculus (SC) output layer saccades are produced when activation reaches a threshold. During fixation, the SC is inhibited by tonically active substantia nigra pars reticulata (SNr) units. SC units become excited through disinhibition by the direct path from striatal GO unit activation and inhibition of SNr units. Responses are selectively suppressed through the indirect path leading from striatal NoGo units through the external segment of the globus pallidus (GPe) to the SNr. Coactivation of mutually incompatible R1 and R2 response units in FEF is registered by anterior cingulate cortex (ACC), which is supposed to activate the subthalamic nucleus (STN), delaying production of a response until the conflict is resolved through excitatory projections to SNr. Innervation of striatum by dopamine (DA) from substantia nigra pars compacta (SNc) allows the system to learn which actions to gate or to suppress through amplification or attenuation of GO relative to NoGo activity in proportion to reward value. The instruction layer encodes task rules to guide production of responses that are congruent (cong) like prosaccades or incongruent (incong) like antisaccades. The DLPFC is supposed to integrate the task cue with the sensory input to guide a correct response by activating the appropriate pool of units in FEF and striatum.

colliculus is inhibited before its activation reaches threshold, the response is inhibited, producing a cancelled trial. If superior colliculus reaches threshold before it is inhibited, the response is executed, producing a non-cancelled trial.

This model was parametrized to simulate performance on the stop-signal task, but it was not fit to the data unlike the previous two models. The model simulated inhibition functions and RT distributions on no-stop-signal and non-cancelled trials but did not compare the simulated functions quantitatively to observed data. The model also simulated activation for units in striatum, substantia nigra pars reticulata, subthalamic nucleus and dorsolateral prefrontal cortex; however, the times of changes in activation trajectory were not compared quantitatively with observed neural data.

The model provides an explanation of the computations occurring within and between units, but it does not achieve quantitative fits of performance data. In scope and performance, it is difficult to compare this model to the more basic interactive race models. Every model of the stop-signal task predicts inhibition functions and RT distributions for no-stop-signal and non-cancelled trials, so the model prediction of the shapes of these functions is not crucially diagnostic. This brain network model aims to account for the integrated action of the many brain structures on which stop-signal performance depends. However, the linking propositions connecting model units and their interactions to brain structures and their interactions are based on incomplete, uncertain information and

are not evaluated rigorously. We acknowledge the complexity of accomplishing quantitative comparisons of critical features of the data (e.g. cancel times, diversity of neuron types), but we know of no other path available for rigorous model comparison. Moreover, a given brain network model must be seen as one instance of a family of plausible alternatives. Distinguishing between complex, mimicking models is difficult or even impossible. Nevertheless, such brain network models are useful for organizing large sets of observations across anatomy, physiology and neuroimaging.

8. Unresolved questions

The integration of mathematical psychology and neuroscience has only just begun [31,36–38]. The models have moved our understanding of response inhibition forward significantly, but much remains to be done. We sketch out some remaining questions.

(a) Alternative modelling approaches

Yu and co-workers have proposed a model assuming that participants decide whether and when to go based on continually updated Bayesian information about both the Go and Stop stimuli, repeatedly assessing the relative value of stopping versus going [39]. This model replicates key characteristics of stop-signal task performance. Because the repetitive evaluation

is designed to optimize outcomes based on benefits and costs, it can adapt performance based on changing reward contingencies or motivational factors. The authors also show how the independent race model can be seen as an approximation to optimal decision-making. Through this link the parameters of the race model, such as the stopping latency, can be changed with task parameters and individual experiences or ability. Accounting for adaptive countermanding performance is an important problem to address in the interactive race framework.

(b) Similarities and differences of effectors

At the biomechanical and neural level, significant differences are well known to distinguish the nimble jerks of the eyes from the ponderous rotations and displacements of head or limb. For example, although co-contraction is a viable means of inhibiting movements of the limbs, it is not possible for the eyes. Also, saccadic eye movements are produced by a brainstem circuit that features powerful inhibition from omnipause neurons, but these have no counterpart in the spinal control of the limbs. The independent and interactive race models are formulated at a level that generalizes across effectors. They can account for the probability and timing of countermanding of eye, head or hand movements. However, countermanding performance with the head or limbs exhibits features seen rarely with the eyes, such as partial responses and mid-flight corrections [40,41]. Meanwhile, when countermanding head movements, the timing of muscle contractions provides a trial-by-trial measure of SSRT [42]. Also, the patterns of muscle contractions can resolve whether body movements are cancelled through inactivation of agonist muscle groups or co-activation of antagonist groups [43]. Of course, muscle co-activation is not possible for the eyes [44]. Comprehensive models of response inhibition will need to diversify to explain the particularities of the different effectors and their coordination [45].

(c) Choice in going and stopping

Another pressing problem highlighted in recent research is choice, both in response to the Go stimulus and to the Stop stimulus. The original interactive race model considered only one accumulator for the Go task. The spiking network and brain network models included two accumulators, one for each possible Go response but did not model activity in the competing accumulator. This was guided by the original saccadic stop-signal task in which a single target was presented, making choice errors impossible, but the stop-signal tasks most common in the literature include response choice.

The probability and latency of choice errors during countermanding and related tasks need to be modelled. An independent race model with two GO processes for correct and error responses plus a STOP process accounts for macaque and human performance of countermanding saccades when multiple potential targets are offered during visual search [46]. Moreover, this model also explained the incidence of corrective saccades including mid-flight corrections. Other recent work extended the independent race model to deal with choice in the Go task and found some evidence for selective influence [47]. In an interactive race model framework the alternative responses must be modelled as stochastic accumulators, and their interaction with the stochastic accumulator for the correct response must be specified. Race models, feed-forward inhibition models, and lateral inhibition models are viable alternatives [29,30,48]. Choice tasks provide the opportunity to manipulate several factors that affect the GO process concurrently, and these factors may influence different parameters of the GO process selectively. Selective influence provides important leverage in modelling. Some parameters should stay constant across conditions while others vary, and this adds important constraints in fitting data. Recent work has begun to test countermanding performance while systematically manipulating visual choice difficulty, finding that SSRT is not influenced by choice difficulty [49].

Choice is also possible in the STOP process. Several investigators have studied varieties of 'selective inhibition' in which some responses but not others must be inhibited when a stop signal occurs [50], or all responses must be inhibited when a stop signal occurs but not when another similar 'ignore' signal occurs [51]. Selective stopping will require interesting extensions for models of response inhibition, because selective stopping to one stimulus but not another often produces violations of the independence assumptions of the race model [51]. This is important because all of the models we have discussed assume that the STOP process and GO process are independent for much of their duration. Independence makes modelling simpler. Non-independent STOP and GO processes are much harder to characterize, but work is beginning on models of selective stopping.

(d) Motivation, memory and strategy

The models we have reviewed consider countermanding in rather impoverished contexts. However, most behaviour is embedded in circumstances shaped by motivation, memory and strategy. Perhaps, not surprisingly, recent studies have demonstrated that measures of stop-signal task performance can vary as a function of motivation signalled by reward magnitude [52–54]. Other studies have described contributions of learning, memory and task goals in countermanding performance [55,56]. We look forward to future work extending the race model framework to incorporate these additional findings.

(e) Individual differences

A major strength of the original independent race model was the measurement of SSRT, interpreted as a measure of response inhibition. Numerous studies have demonstrated significant changes of SSRT and other measures of countermanding performance in individuals diagnosed with various neurological [57] and psychopathological disorders [58]. Even healthy participants exhibit incidental or strategic variation of performance across individuals. The development of interactive race models provides a foundation on which to explore whether parameter modification alone is sufficient or process addition is necessary to account for such individual differences.

(f) Model mimicry and linking propositions

The models we discussed make common predictions for behaviour and physiology. Quantitative fits to behaviour—inhibition functions and RT distributions for no-stop-signal and non-cancelled trials—were equivalent for the various interactive race architectures. The brain network model of frontal cortex and basal ganglia also produced the same qualitative trends. On the one hand, one approach to resolving such model mimicry involves designing more complex experiments with more resolved manipulations and dissociations. On the

other hand, to the extent that all of these models are designed to account for behaviour that corresponds to the outcome of a race between GO and STOP processes, each must mimic the other.

Neural measures derived from the models exhibit mimicry too [59]. All of the interactive race models predict similar patterns and timing of modulation of activity in movement and fixation neurons. Mimicry in model predictions of neural measures may be broken by examining additional neural measures including levels of activation and modulation [29,30]. Crucially, the use of neural measures to resolve mimicry of models requires the validity of the mapping between the neural measures and the modelled process. Verification of such linking propositions [1,2] requires empirical investigation. For example, a recent study demonstrated that the duration of activation of FEF movement neurons paralleled the duration of the race model GO processes, establishing a necessary relationship for the mapping between model and neural processes [60]. Similarly, studies have also demonstrated other neurons in FEF and in other cortical areas such as the supplementary eye field that modulate not at all, too late or in the wrong direction during countermanding tasks, so they cannot instantiate the race model processes. Meanwhile, other studies have produced evidence that seems to contradict the identity mapping of the stochastic accumulation process onto movement neuron activity [61,62]. For example, the systematic delay of RT observed after cancelled stop signal trials was accomplished not by an elevation of threshold of FEF and SC movement neurons as predicted by accumulator models but instead by postponement of beginning of activation after target presentation. However, model simulations indicate that measures of dynamics need not agree with the values of the parameters that generated them, especially if the data and model predictions are noisy. For example, the model and measured threshold can be different among ensembles of accumulators [63]. Also, measured accumulation onsets need not correspond to non-decision times, nor measured rates of growth to drift rates, nor measured thresholds to model thresholds [59,64].

Despite these complicacies, we are optimistic that model-based cognitive neuroscience is the most secure bridge between behaviour and brain. Many insights have been provided by the recursive blending of rigorous models from mathematical psychology with detailed measures from neurophysiology and neuroimaging during tasks with well-grounded computational theories. Models that make the same predictions about performance data can be distinguished through neural measures [21,25,29,30], neural data can provide useful constraints on the nature of model mechanisms [29,30,65], and model mechanisms can provide a language for understanding the nature of neural dynamics [59].

References

1. Schall JD. 2004 On building a bridge between brain and behavior. *Annu. Rev. Psychol.* **55**, 23–50. (doi:10.1146/annurev.psych.55.090902.141907)

2. Teller DY. 1984 Linking propositions. *Vision Res.* **24**, 1233–1246. (doi:10.1016/0042-6989(84)90178-0)

3. Logan GD, Cowan WB. 1984 On the ability to inhibit thought and action: a theory of an act of control. *Psychol. Rev.* **91**, 295–327. (doi:10.1037/0033-295X.91.3.295)

4. Logan GD. 1994 On the ability to inhibit thought and action: a users' guide to the stop signal paradigm. In *Inhibitory processes in attention, memory, and language* (eds D Dagenbach, TH Carr), pp. 189–239. San Diego, CA: Academic Press.

5. Verbruggen F, Logan GD. 2009 Models of response inhibition in the stop-signal and stop-change paradigms. *Neurosci. Biobehav. Rev.* **33**, 647–661. (doi:10.1016/j.neubiorev.2008.08.014)

6. Ozyurt J, Colonius H, Arndt PA. 2003 Countermanding saccades: evidence against independent processing of go and stop signals. *Percept. Psychophys.* **65**, 420–428. (doi:10.3758/BF03194573)

7. Gulberti A, Arndt PA, Colonius H. 2014 Stopping eyes and hands: evidence for non-independence of stop and go processes and for a separation of central and peripheral inhibition. *Front. Hum. Neurosci.* **8**, 61. (doi:10.3389/fnhum.2014.00061)

8. Hanes DP, Schall JD. 1995 Countermanding saccades in macaque. *Vis. Neurosci.* **12**, 929–937. (doi:10.1017/S0952523800009482)

9. Hanes DP, Carpenter RH. 1999 Countermanding saccades in humans. *Vision Res.* **39**, 2777–2791. (doi:10.1016/S0042-6989(99)00011-5)

10. Logan GD, Irwin DE. 2000 Don't look! Don't touch! Inhibitory control of eye and hand movements. *Psychon. Bull. Rev.* **7**, 107–112. (doi:10.3758/BF03210728)

11. Kornylo K, Dill N, Saenz M, Krauzlis RJ. 2003 Cancelling of pursuit and saccadic eye movements in humans and monkeys. *J. Neurophysiol.* **89**, 2984–2999. (doi:10.1152/jn.00859.2002)

12. Missal M, Heinen SJ. 2017 Stopping smooth pursuit. *Phil. Trans. R. Soc. B* **372**, 20160200. (doi:10.1098/rstb.2016.0200)

13. Krauzlis RJ, Goffart L, Hafed ZM. 2017 Neuronal control of fixation and fixational eye movements. *Phil. Trans. R. Soc. B* **372**, 20160205. (doi:10.1098/rstb.2016.0205)

14. Hanes DP, Patterson WF, Schall JD. 1998 Role of frontal eye field in countermanding saccades: visual, movement and fixation activity. *J. Neurophysiol.* **79**, 817–834.

15. Pouget P, Murthy A, Stuphorn V. 2017 Cortical control and performance monitoring of interrupting and redirecting movements. *Phil. Trans. R. Soc. B* **372**, 20160201. (doi:10.1098/rstb.2016.0201)

16. Paré M, Hanes DP. 2003 Controlled movement processing: superior colliculus activity associated with countermanded saccades. *J. Neurosci.* **23**, 6480–6489.

17. Murthy A, Ray S, Shorter SM, Schall JD, Thompson KG. 2009 Neural control of visual search by frontal eye field: effects of unexpected target displacement on visual selection and saccade preparation. *J. Neurophysiol.* **101**, 2485–2506. (doi:10.1152/jn.90824.2008)

18. Costello MG, Zhu D, Salinas E, Stanford TR. 2013 Perceptual modulation of motor—but not visual—responses in the frontal eye field during an urgent-decision task. *J. Neurosci.* **33**, 16 394–16 408. (doi:10.1523/JNEUROSCI.1899-13.2013)

19. Mallet N, Schmidt R, Leventhal D, Chen F, Amer N, Boraud T, Berke JD. 2016 Arkypallidal cells send a stop signal to striatum. *Neuron* **89**, 308–316. (doi:10.1016/j.neuron.2015.12.017)

20. Schmidt R, Berke JD. 2017 A Pause-then-Cancel model of stopping: evidence from basal ganglia neurophysiology. *Phil. Trans. R. Soc. B* **372**, 20160202. (doi:10.1098/rstb.2016.0202)

21. Boucher L, Palmeri TJ, Logan GD, Schall JD. 2007 Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol. Rev.* **114**, 376–397. (doi:10.1037/0033-295X.114.2.376)

22. Wong-Lin K, Eckhoff P, Holmes P, Cohen JD. 2010 Optimal performance in a countermanding saccade task. *Brain Res.* **1318**, 178–187. (doi:10.1016/j.brainres.2009.12.018)

23. Ramakrishnan A, Sureshbabu R, Murthy A. 2012 Understanding how the brain changes its mind: microstimulation in the macaque frontal eye fields reveals how saccade plans are changed. *J. Neurosci.* **32**, 4457–4472. (doi:10.1523/JNEUROSCI.3668-11.2012)

24. Salinas E, Stanford TS. 2013 The countermanding task revisited: fast stimulus detection is a key determinant of psychophysical performance. *J. Neurosci.* **33**, 5668–5685. (doi:10.1523/JNEUROSCI.3977-12.2013)

25. Logan GD, Yamaguchi M, Schall JD, Palmeri TJ. 2015 Inhibitory control in mind and brain 2.0: blocked-input models of saccadic countermanding. *Psychol. Rev.* **122**, 115–147. (doi:10.1037/a0038893)

26. Cutsuridis V. 2017 Behavioural and computational varieties of response inhibition in eye movements. *Phil. Trans. R. Soc. B* **372**, 20160196. (doi:10.1098/rstb.2016.0196)

27. Band GPH, van Boxtel GJM. 1999 Inhibitory motor control in stop paradigms: review and reinterpretation of neural mechanisms. *Acta Psychol.* **101**, 179–211. (doi:10.1016/S0001-6918(99) 00005-0)

28. Logan GD. 1979 On the use of a concurrent memory load to measure attention and automaticity. *J. Exp. Psychol. Hum. Percept. Performance* **5**, 189–207. (doi:10.1037/0096-1523.5.2.189)

29. Purcell BA, Heitz RP, Cohen JY, Schall JD, Logan GD, Palmeri TJ. 2010 Neurally constrained modeling of perceptual decision making. *Psychol. Rev.* **117**, 1113–1143. (doi:10.1037/a0020311)

30. Purcell BA, Schall JD, Logan GD, Palmeri TJ. 2012 From salience to saccades: multiple-alternative gated stochastic accumulator model of visual search. *J. Neurosci.* **32**, 3433–3446. (doi:10.1523/JNEUROSCI.4622-11.2012)

31. Palmeri TJ, Schall JD, Logan GD. 2015 Neurocognitive modelling of perceptual decisions. In *Oxford handbook of computational and mathematical psychology* (eds JR Busemeyer, J Townsend, ZJ Wang, A Eidels). Oxford, UK: Oxford University Press.

32. Lo CC, Boucher L, Paré M, Schall JD, Wang XJ. 2009 Proactive inhibitory control and attractor dynamics in countermanding action: a spiking neural circuit model. *J. Neurosci.* **29**, 9059–9071. (doi:10.1523/JNEUROSCI.6164-08.2009)

33. Logan GD, Schall JD, Palmeri TJ. 2015 Neural models of stopping and going. In *An introduction to model-based cognitive neuroscience* (eds B Forstmann, EJ Wagenmakers). Berlin, Germany: Springer Neuroscience.

34. Frank MJ. 2006 Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Netw.* **19**, 1120–1136. (doi:10.1016/j.neunet.2006.03.006)

35. Wiecki TV, Frank MJ. 2013 A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychol. Rev.* **120**, 329–355. (doi:10.1037/a0031542)

36. Forstmann BU, Wagenmakers EJ, Eichele T, Brown S, Serences JT. 2011 Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? *Trends Cogn. Sci.* **15**, 272–279. (doi:10.1016/j.tics.2011.04.002)

37. Turner BM, Forstmann BU, Wagenmakers E-J, Brown SD, Sederberg PB, Steyvers M. 2013 A Bayesian framework for simultaneously modeling neural and behavioral data. *Neuroimage* **72**, 193–206. (doi:10.1016/j.neuroimage.2013.01.048)

38. Turner BM, Forstmann BU, Love BC, Palmeri TJ, Van Maanen L. In press. Approaches to analysis in model-based cognitive neuroscience. *J. Math. Psychol.* (doi:10.1016/j.jmp.2016.01.001)

39. Shenoy P, Yu AJ. 2011 Rational decision-making in inhibitory control. *Front. Hum. Neurosci.* **5**, 48. (doi:10.3389/fnhum.2011.00048)

40. Corneil BD, Elsley JK. 2005 Countermanding eye-head gaze shifts in humans: marching orders are delivered to the head first. *J. Neurophysiol.* **94**, 883–895. (doi:10.1152/jn.01171.2004)

41. Corneil BD, Cheng JC, Goonetilleke SC. 2013 Dynamic and opposing adjustment of movement cancellation and generation in an oculomotor countermanding task. *J. Neurosci.* **33**, 9975–9984. (doi:10.1523/JNEUROSCI.2543-12.2013)

42. Goonetilleke SC, Wong JP, Corneil BD. 2012 Validation of a within-trial measure of the oculomotor stop process. *J. Neurophysiol.* **108**, 760–770. (doi:10.1152/jn.00174.2012)

43. Scangos KW, Stuphorn V. 2010 Medial frontal cortex motivates but does not control movement initiation in the countermanding task. *J. Neurosci.* **30**, 1968–1982. (doi:10.1523/JNEUROSCI.4509-09.2010)

44. Godlove DC, Schall JD. 2016 Microsaccade production during saccade cancelation in a stop-signal task. *Vision Res.* **118**, 5–16. (doi:10.1016/j.visres.2014.10.025)

45. Boucher L, Stuphorn V, Logan GD, Schall JD, Palmeri TJ. 2007 Stopping eye and hand movements: are the processes independent? *Percept. Psychophys.* **69**, 785–801. (doi:10.3758/BF03193779)

46. Camalier CR, Gotler A, Murthy A, Thompson KG, Logan GD, Palmeri TJ, Schall JD. 2007 Dynamics of saccade target selection: race model analysis of double step and search step saccade production in human and macaque. *Vision Res.* **47**, 2187–2211. (doi:10.1016/j.visres.2007.04.021)

47. Logan GD, Van Zandt T, Verbruggen F, Wagenmakers E-J. 2014 On the ability to inhibit thought and action: general and special theories of an act of control. *Psychol. Rev.* **121**, 66–95. (doi:10.1037/a0035230)

48. Ratcliff R, Smith PL. 2004 A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* **111**, 333–367. (doi:10.1037/0033-295X.111.2.333)

49. Middlebrooks PG, Schall JD. 2014 Response inhibition during perceptual decision making in humans and macaques. *Atten. Percept. Psychophys.* **76**, 353–366. (doi:10.3758/s13414-013-0599-6)

50. Aron AR, Verbruggen F. 2008 Dissociating a selective from a global mechanism for stopping. *Psychol. Sci.* **19**, 1146–1153. (doi:10.1111/j.1467-9280.2008.02216.x)

51. Bissett PG, Logan GD. 2014 Selective stopping? Maybe not. *J. Exp. Psychol. Gen.* **143**, 455–472. (doi:10.1037/a0032122)

52. Leotti LA, Wager TD. 2010 Motivational influences on response inhibition measures. *J. Exp. Psychol. Hum. Percept. Perform.* **36**, 430–447. (doi:10.1037/a0016802)

53. Padmala S, Pessoa L. 2010 Interactions between cognition and motivation during response inhibition. *Neuropsychologia* **48**, 558–565. (doi:10.1016/j.neuropsychologia.2009.10.017)

54. Boehler CN, Schevernels H, Hopf J-M, Stoppel CM, Krebs RM. 2014 Reward prospect rapidly speeds up response inhibition via reactive control. *Cogn. Affect. Behav. Neurosci.* **14**, 593–5609. (doi:10.3758/s13415-014-0251-5)

55. Verbruggen F, Logan GD. 2008 Automatic and controlled response inhibition: associative learning in the go/no-go and stop-signal paradigms. *J. Exp. Psychol. Gen.* **137**, 649–672. (doi:10.1037/a0013170)

56. Verbruggen F, Logan GD. 2008 Long-term aftereffects of response inhibition: memory retrieval, task goals, and cognitive control. *J. Exp. Psychol. Hum. Percept. Perform.* **34**, 1229–1235. (doi:10.1037/0096-1523.34.5.1229)

57. Bissett PG, Logan GD, van Wouwe NC, Tolleson CM, Phibbs FT, Claassen DO, Wylie SA. 2015 Generalized motor inhibitory deficit in Parkinson's disease patients who freeze. *J. Neural Trans.* (*Vienna*) **122**, 1693–1701. (doi:10.1007/s00702-015-1454-9)

58. Thakkar KN, Schall JD, Boucher L, Logan GD, Park S. 2011 Response inhibition and response monitoring in a saccadic countermanding task in schizophrenia. *Biol. Psychiatry* **69**, 55–62. (doi:10.1016/j.biopsych.2010.08.016)

59. Purcell BA, Palmeri TJ. In press. Relating accumulator model parameters and neural dynamics. *J. Math. Psychol.* (doi:10.1016/j.jmp.2016.07.001)

60. Nelson MJ, Murthy A, Schall JD. 2016 Neural control of visual search by frontal eye field: chronometry of neural events and race model processes. *J. Neurophysiol.* **115**, 1954–1969. (doi:10.1152/jn.01023.2014)

61. Heitz R, Schall JD. 2012 Neural mechanisms of speed-accuracy tradeoff. *Neuron* **76**, 616–628. (doi:10.1016/j.neuron.2012.08.030)

62. Pouget P, Logan GD, Palmeri TJ, Boucher L, Paré M, Schall JD. 2011 Neural basis of adaptive response time adjustment during saccade countermanding. *J. Neurosci.* **31**, 12 604–12 612. (doi:10.1523/JNEUROSCI.1868-11.2011)

63. Zandbelt B, Purcell BA, Palmeri TJ, Logan GD, Schall JD. 2014 Response times from ensembles of accumulators. *Proc. Natl Acad. Sci. USA* **111**, 2848–2853. (doi:10.1073/pnas.1310577111)

64. Purcell BA. 2013 Neural mechanisms of perceptual decision-making. Doctoral Dissertation, Vanderbilt University.

65. Hanes DP, Schall JD. 1996 Neural control of voluntary movement initiation. *Science* **274**, 427–430. (doi:10.1126/science.274.5286.427)