Action inhibition revisited: Stopping is not faster than going

Yue Du[1*], Alexander D. Forrence[1, 2], Delaney M. Metcalf[1, 3], Adrian M. Haith[1]

[1]Department of Neurology, Johns Hopkins University, Baltimore, MD, USA
[2]Department of Psychology, Yale University, New Haven, CT, USA
[3]School of Medicine, University of Georgia, Augusta, GA, USA

*Corresponding author: ydu27@jhmi.edu

Abstract

We often need to swiftly abort a prepared response at the last moment before it is initiated. Our ability to abort a planned response is thought to be a fundamental facet of action control, which is distinguished by being more rapid than initiating an action, and has been suggested to be enabled by specialized neural mechanisms. This narrative has, however, largely been established based on experiments in which there is much greater urgency to abort an action than there is to generate an action. Here, we demonstrate that, under conditions of matched urgency, the speed at which participants are able to abort an action is comparable to the speed at which they can initiate an action. Our results challenge the prevailing view that reactive stopping behaviors have a privileged status over action initiation. Instead, action initiation may be systematically delayed to allow time to abort an action if needed. We propose that action cancellation and action initiation may reflect two opposing states of a single process supporting a decision about whether to act or not.

Keywords: Action inhibition, Stop-signal task, Race model, Action initiation

Introduction

The ability to cancel or inhibit a voluntary action moment before it is initiated is widely considered to be a fundamental component of action control (Logan and Cowan, 1984). For instance, a pedestrian about to step into the street must swiftly abort this act if a fast car suddenly approaches. In the laboratory, reactive action inhibition of this kind has primarily been studied using the *stop-signal task*, in which people are asked to respond to an imperative “go” stimulus as quickly as possible with a button press (or other action) but must cancel this response in the event of a “stop” signal. Using such tasks, it has been found that people are able to cancel an intended response even if the stop signal is received after the “go” cue is presented (Logan and Cowan, 1984; Verbruggen et al., 2019). The time required to successfully cancel the initiation of a response, therefore, seems to be shorter than the time needed to initiate a response (Slater-Hammel, 1960). This observation has prompted the suggestion that reactive action inhibition may be a privileged component of action control, potentially supported by a specialized mechanism through the prefrontal-cortex – basal-ganglia hyperdirect pathway that acts as an “emergency brake” to prevent unwanted responses from being initiated (Aron et al., 2014; Dunovan et al., 2015; Hannah and Aron, 2021; Wiecki and Frank, 2013).

The idea that stopping an action is inherently faster than generating one depends critically on the assumption that the reaction time for a movement represents the minimum possible time at which it could be initiated. Recent findings have, however, shown that this assumption is not true. Instead, movements seem to be initiated at a delay of up to 100 ms after they are prepared and ready to execute (Carlsen et al., 2004; Haith et al., 2016; Valls-Solé et al., 1995). Moreover, the size of the delay between preparation and initiation can be influenced by the urgency of the task (Haith et al., 2016). In stop-signal tasks, the delay in action initiation may be further exacerbated, with reaction times averaging up to around 450 ms (e.g., Leunissen et al., 2017) – much longer than the typical reactions times of 200-250 ms to respond to a go cue (Luce, 1991; Welford, 1980). By contrast, the need to cancel a response has high urgency because it has to be done before the response is initiated and the time needed to abort a response ranges from 180 ms to 270 ms (He et al., 2021; Leunissen et al., 2017; Logan and Cowan, 1984; Matzke et al., 2021). In the stop-signal paradigm, therefore, there is a marked asymmetry in urgency between the requirements to initiate a response and the potential requirement to cancel one. This asymmetry could account for action inhibition seeming to be much faster than action initiation.

Here, we performed an experiment to more fairly compare the speed at which participants could initiate an action or cancel an intended action by measuring them under separate conditions in which the urgency was matched as closely as possible. We adopted a “timed-response” approach in which participants were trained to always respond at a prescribed time in each trial (Haith et al., 2016). By occasionally and unexpectedly switching the required behavior, either from requiring a response at the prescribed time to requiring no response (Coxon et al., 2006; Leunissen et al., 2017; Slater-Hammel, 1960), or from requiring no response to requiring a response, we were able to establish the time course over which participants were able to inhibit or initiate responses and directly compare them.

Results

Participants viewed a circle moving vertically downward to cross a horizonal line (Fig. 1). They were instructed to press a button when the circle overlapped the target line if the circle was white, but do nothing if the circle was black (the actual meaning of the two colors was counterbalanced across participants). The initial color of the circle varied from block to block so that participants were either initially cued to respond, or to not respond. In each case, however, the color of the circle changed in a subset of trials (30%) at a random time (50-500ms) before hitting the target line, so that participants needed to either rapidly abort an initially prepared response, similar to conventional stopping paradigms (Slater-Hammel, 1960; Verbruggen et al., 2019) (Response-to-No-Response condition; R-to-NR), or they had to rapidly initiate a response that they were not initially intending to (No-Response-to-Response; NR-to-R).

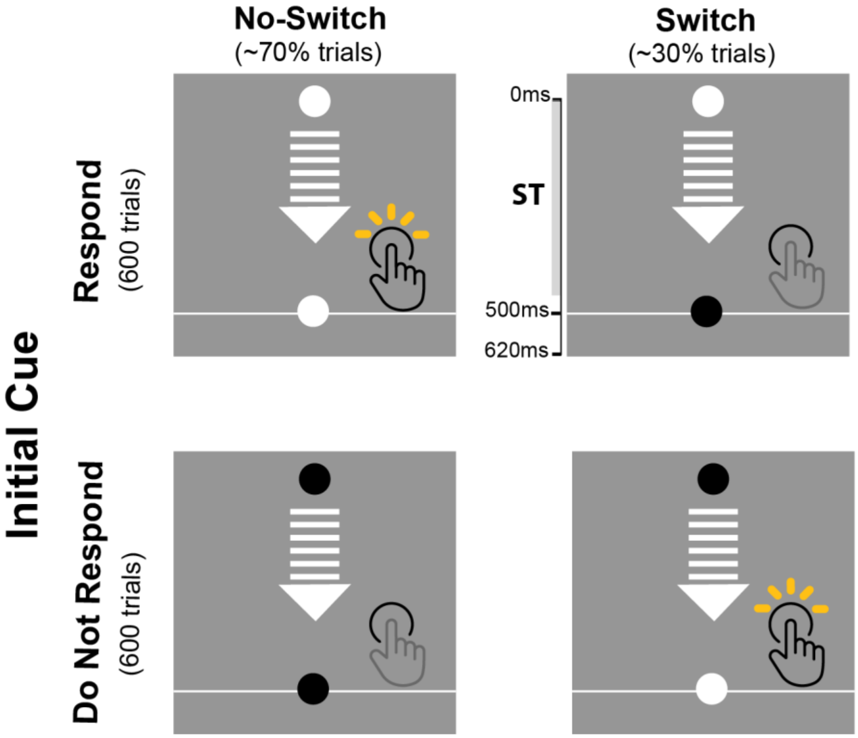

Fig. 1. Experiment procedure. Participants were asked to either press a key or do nothing when a moving circle reached the target line. Whether or not a response was required in a given trial depended on the color of the circle (e.g., white = respond; black = do not respond). The actual circle color was counterbalanced across participants, controlling for the potential perceptual differences of black and white colors. The circle always started with the same color within each block of 100 trials. In ~70% of trials, the circle remained the same color throughout. However, in another ~30% of trials, the circle changed color before it hit the target line, forcing participants to cancel a preplanned response or to initiate a response when the circle crossed the line. By manipulating the time at which the circle color changed in each condition, we were able to compare people's ability to stop themselves from generating a planned response ("R-to-NR" condition; upper panel) to the ability to rapidly generate a response ("NR-to-R" condition; lower panel). Participants completed 12 blocks of 100 trials, generating 204 switch trials in each condition.

The performance of one exemplar participant is shown in Fig. 2B. In switch trials of the R-to-NR condition, when the time available to abort an impending response (the *forced response time* or *forced RT*; Fig. 2A; Methods) was very short (< 100ms), this participant almost always failed to abort their prepared response. However, at longer RTs (> 300ms), this participant was able to correctly cancel the response in almost all trials. At intermediate RTs, the participant was sometimes successful and sometimes unsuccessful in canceling their response. Similar but complementary behavior was observed in the NR-to-R switch trials. This participant failed to initiate a timely response if the circle changed color shortly before it crossed the line. When allowed longer time to react to the color change, however, they always correctly generated a response.

From this raw response data, we constructed a speed–accuracy trade-off for each condition (R-to-NR and NR-to-R), based on a 50 ms sliding window on the allowed RT, which describes the probability of correctly aborting a response (R-to-NR) or the probability of generating a successful response (NR-to-R) as a function of allowed RT (Fig. 2C). For the exemplar participant in Fig. 2B, the centers of the speed–accuracy trade-off function, representing the average time required to either cancel (R-to-NR condition) or initiate (NR-to-R) a response, were both located around 280 ms. The speed–accuracy trade-offs averaged across all participants (n = 35 out of 36; Methods; Figs. S1 and S2) showed the same pattern (Fig. 2D; Dashed lines), indicating that equal amounts of time were required to initiate a response and to cancel the initiation of a response.

To quantitatively estimate how fast participants could cancel a response in the R-to-NR condition, we considered a simple model in which we assumed that the cancellation of a response could be thought of as a discrete event occurring at a random time $T_{NR} \sim \mathcal{N}(\mu_{NR}, \sigma_{NR}^2)$ after the circle changed color (Haith et al., 2016). On a given trial, if the time needed to cancel a response (T_{NR}) was shorter than the allowed RT, participants would successfully avoid generating a response. But, if the required time, T_{NR}, was longer than the allowed RT, participants would fail to cancel the impending response. This led to a predicted probability of being correct that increased smoothly as a function of allowed RT, as observed in the data. An analogous model was applied to the NR-to-R condition. In this case, the decision to initiate a response was assumed to be made at a random time $T_R \sim \mathcal{N}(\mu_R, \sigma_R^2)$ after the circle changed color and, on a given trial, they would succeed at correctly generating the response at the right time only if T_R is shorter than the allowed

RT on that trial. Thus, μ_{NR} and μ_R represented the average times required either to withhold a movement or initiate a movement, respectively.

We fitted these models to each participants' data via maximum likelihood estimation, yielding model fits that closely matched the empirical data (Fig. 2D; Solid lines). By comparing the estimates of μ_{NR} to μ_R across participants, we found that the time required to cancel the initiation of an impending response (μ_{NR}: 294.5 ms ± 28.3 ms; mean ± s.d.) and the time required to initiate a response (μ_R: 289.0 ms ± 24.4 ms; mean ± s.d.) were not significantly different from one another ($\mu_{nr} - \mu_r$: 5.5 ms ± 24.4 ms, Cohen's d = 0.207, t34 = 1.34, paired t-test, p = 0.19, 95% CI: [-2.8ms, 13.9ms]; Fig. 2E) and they were also highly correlated (ρ = 0.58, p < 0.001). The equivalence and noninferiority test demonstrated that these two speeds were statistically equivalent ($p_{lower} < 0.001$; $p_{upper} = 0.004$, 90% CI: [-1.5ms, 12.5ms]). Thus, when we compared response initiation and cancellation under experimental conditions that were as closely matched as possible, we found no evidence to support the assertion that cancelling a response is faster than generating one.

Because μ_{NR} and μ_R represent the RT at which performance accuracy reaches the center of the speed–accuracy trade-off function, their estimates were sensitive to exactly how we defined whether a trial was performed correctly or not. In our initial analysis, a late response (if the circle was below the target line without overlapping it at the time the response was made) of an NR-to-R trial was considered as incorrect, whereas in R-to-NR trials, it was impossible to distinguish between trials in which a response was cancelled at the correct time, and trials in which a response was delayed before being cancelled. This unavoidably imposed a stronger requirement of timing precision on initiating a response than cancelling a response (asymptotic accuracy: β = 0.89 ± 0.06 vs. β = 0.96 ± 0.05; t34 = -5.26, paired t-test, p < 0.0001, 95% CI: [-0.09, -0.04]; Methods). After relaxing the timing requirement for considering a response as being correctly initiated, which matched the asymptotic accuracies between conditions (0.94 ± 0.06 vs. 0.96 ± 0.05; t34 = -1. 9, paired t-test, p = 0.07, 95% CI: [-0.04ms, 0.01ms]), we found that the updated μ_R (286.3 ms ± 23.6 ms) was only 2.7 ms different from the original analysis and was only 8.2 ms shorter than μ_{NR} ($\mu_{NR} - \mu_R$: 8.2 ms ± 23.5 ms, Cohen's d = 0.31, t34 = 2.06, paired t-test, p = 0.047, 95% CI: [0.0ms, 16.3ms], power = 0.5; Fig. S3). In addition, the difference between μ_{NR} and the updated μ_R fell within the equivalence bound between -16.5 ms and 16.5 ms ($p_{lower} < 0.001$; $p_{upper} = 0.02$, 90% CI: [1.5ms, 14.9ms]). Our conclusion about action cancellation and initiation having a similar time course was not strongly affected by late responses.

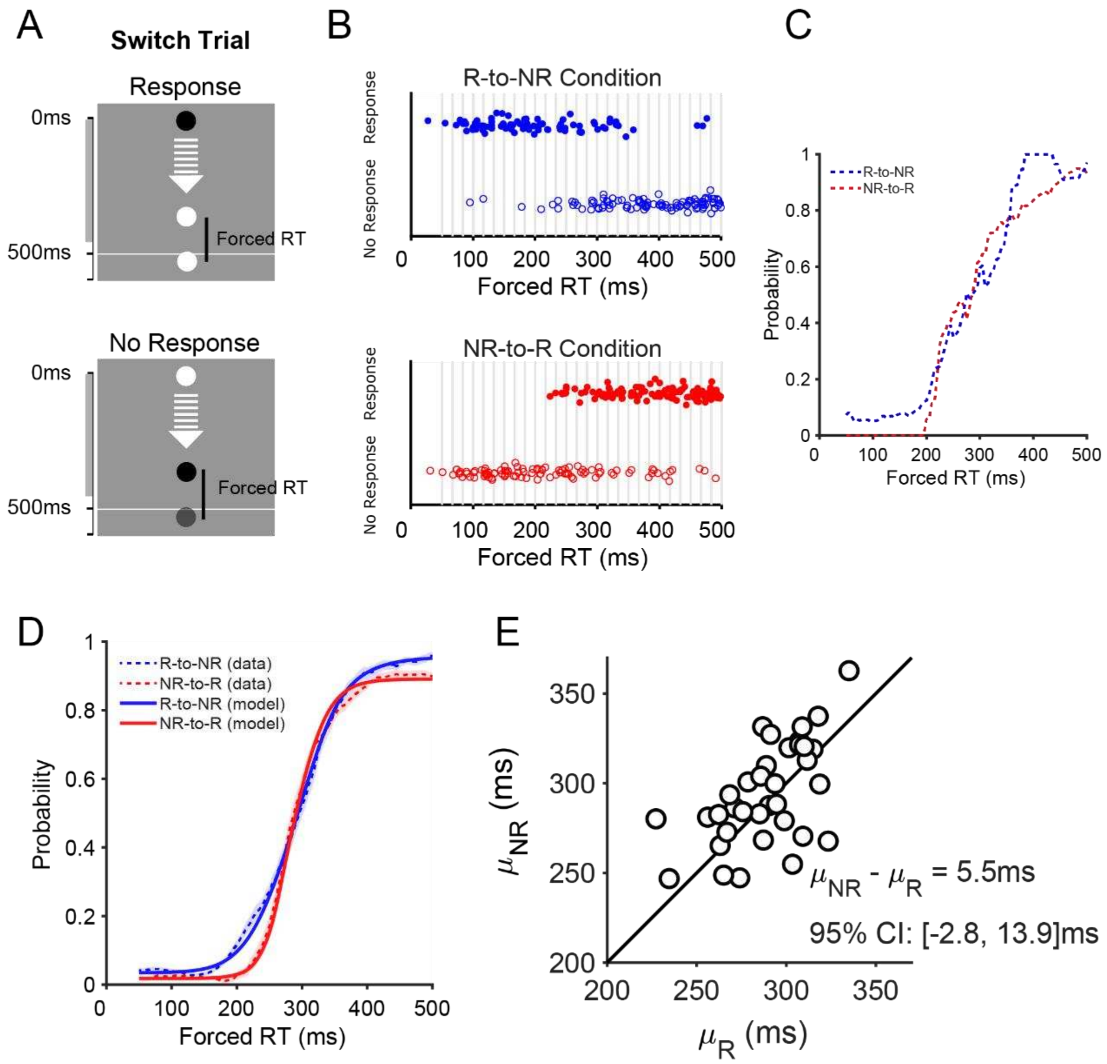

Fig. 2. Cancelling an impending response is not faster than initiating one. A) The forced RT was quantified as the time elapsed from color change to the time of the button press when participants generated a response. When no response occurred, the forced RT was approximated as the time interval between the color change to the typical time of button presses in comparable trials (see Methods for more details). B) Behavior of one exemplar participant. In trials in which only a very short forced RT was allowed, this participant consistently made the wrong choice as to whether to respond or not. When a longer forced RT was allowed, this participant was able to consistently make the correct choice to respond or not. Vertical jitter was added to allow individual data points to be seen more easily. C) The raw data were used to construct speed–accuracy trade-offs, showing the probability of a correct choice as a function of forced RT. D) Mean speed–accuracy trade-offs for each condition across all participants (dashed lines) were well captured by a computational model (solid lines) in which we assumed that the decision to respond or not could be thought of as a discrete event occurring at a random time $T \sim \mathcal{N}(\mu, \sigma^2)$ after the circle changes color. E) We

used this model to estimate the average time needed to cancel an intended response μ_{NR}in R-to-NR condition and the average time to initiate a response μ_R in the NR-to-R condition. Across all participants, μ_{NR} (294.5 ms ± 28.3 ms; mean ± s.d.) was not significantly different from μ_R (289.0 ms ± 24.4 ms; mean ± s.d.).

Similar to other stop-signal tasks, the preceding analyses relied on designating a surrogate RT for trials where a response is missing. To eschew the need to devise a proxy RT when participants failed to respond, we focused only on trials in which a response was generated and analyzed the actual time of responses (i.e., the time at which a response is made after the trial onset) as a function of the intended forced RT (i.e., time interval between color change and the target line) in the NR-to-R condition. We observed that responses were generated around the target line when the intended RT time was long (e.g., > 350 ms), but tended to be made very late when the intended RT was short (e.g., < 200 ms). Participants seemed to time their responses (i.e., by delaying when to initiate the response) given long intended RT and switch to performing the task in a reactive mode, responding as soon as possible after the color change, even if this was too late. We assumed that the timing of this switch strategy coincided with the minimum time at which participants could successfully initiate a movement on time. To estimate the timing of this switch μ_T, we fit a simple model of the timing of participants' responses as a function of intended forced RT (Methods). The model comprised two linear components: one to represent accurate timing at longer RTs, and one to represent delayed, reactive timing at very short RTs (Fig. 4A; Solid light blue line). The estimates of μ_T (302.3 ms ± 39.0 ms; mean ± s.d.) were in a close agreement with the original estimate of μ_R based on approximated RT (ρ = 0.65; p < 0.0001) and, importantly, did not differ significantly from the time that participants needed to withhold a previously intended response, μ_{NR} estimated in the R-to-NR condition ($\mu_{NR} - \mu_T$: -7.7 ms ± 37.7 ms, Cohen's d = -0.22, t34 = -1.25, paired t-test, p = 0.218, 95% CI: [-20.3ms, 4.8ms]; Fig. 4C). Results from the equivalence and noninferiority test confirmed that these two speeds were similar to each other ($p_{\text{lower}} = 0.003$; $p_{\text{upper}} < 0.001$, 90% CI: [-18.4ms, 2.8ms]). This result accords with our initial analysis that cancelling a response is not faster than initiating a response.

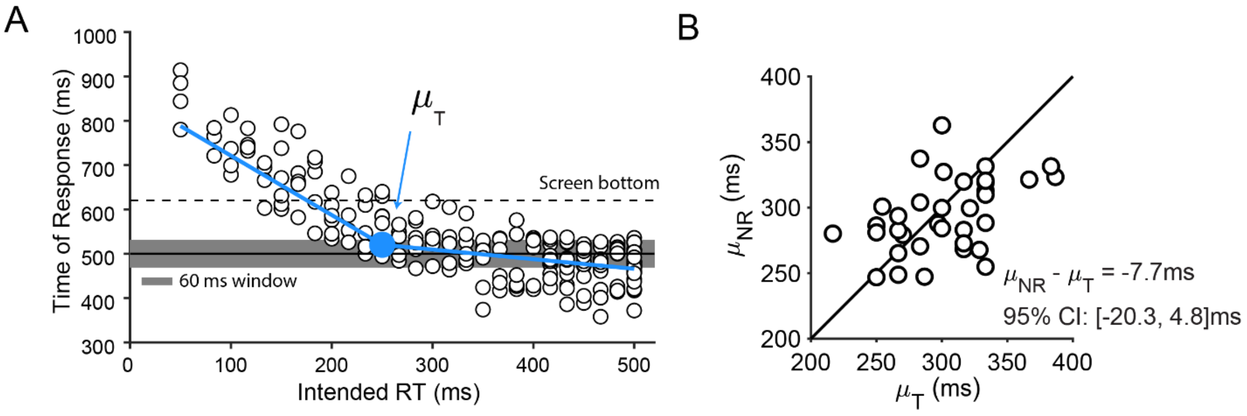

Fig. 3. Alternative analysis based on the pattern of delayed responses. A) The time of response data as a function of intended RT from the same exemplar participant as in Fig. 2B. The time of response was measured as the time interval between trial start and the time at which a response is made. Intended forced RT was the time interval between color change and the target line. Since this analysis focused only on trials in which a response was actually generated, it did not rely on approximating unobserved RT. We fitted the time of response data with two linear functions of intended forced RT, which intersected at μ_T – a parameter that represents the minimum forced RT at which an accurately timed response could still be made. B) Across all participants, the estimates of μ_T (302.3 ms ± 39.0 ms; mean ± s.d.) – the alternative estimate of time required to initiate a response – were not significantly different from μ_{NR} (294.5 ms ± 28.3 ms; mean ± s.d.), the original estimate of the time required to cancel a response.

Discussion

By matching experimental conditions as closely as possible between cancelling an impending action (i.e., "stopping") and generating an action (i.e., "going"), we found that the time for these two processes is comparable, both around 290 ms (with a ~33 ms delay caused by visual display, which was not accounted for in our calculations), suggesting that stopping and going can occur equally rapidly. This result is echoed by previous evidence from two independent research fields showing that simple reaction time (Luce, 1991; Welford, 1980) and the time to stop an action (He et al., 2021; Leunissen et al., 2017; Logan and Cowan, 1984; Matzke et al., 2021) are both around 200 – 250 ms, and both can reduce to around 150 ms triggered by an unexpected event (Carlsen et al., 2004; Haith et al., 2016; Wessel and Aron, 2017). Our finding questions the consensus view from previous experimental, computational, and theoretical work that a rapid, dedicated inhibition mechanism exists to act like an "emergency brake" on response initiation and prevent an unwanted response to be produced (Aron et al., 2014; Boucher et al., 2007; Dunovan et al., 2015; Logan and Cowan, 1984; Slater-Hammel, 1960; Verbruggen et al., 2019; Wiecki and Frank, 2013).

If stopping is not any faster than going, how are we ever able to prevent ourselves from making an unwanted response? One explanation is that response initiation is typically delayed to allow time to make a decision about whether to act or not. Although delaying is a common strategy in proactive stopping (Hannah and Aron, 2021), it has not thought to be a relevant factor in reactive inhibition. However, even in seemingly reactive tasks, reaction times to initiate a movement are delayed beyond the time of movement preparation by up to 100 ms (Carlsen et al., 2004; Haith et al., 2016; Valls-Solé et al., 1995). This flexibility in deciding when to initiate a movement is well aligned with the concept of "freedom for immediacy" – the capacity to decouple our responses to external stimuli from the appearance of the stimulus itself (Haggard, 2008; Haith et al., 2016). The systematic delay between preparation (i.e., what to act) and initiation (i.e., when to act) is thought to exist, at least in part, to avoid the risk of initiating a response before it has been fully prepared. Likewise, in the context of stopping behaviors considered here, participants may make the decision to respond quite quickly after seeing the go cue (i.e., whether to act), but may delay initiation of their response to avoid the risk of initiating a movement that should instead be canceled. Indeed, reaction times in previous stop-signal tasks are often prolonged (e.g., Leunissen et al., 2017) and accompanied by delayed motor cortex excitability (Rawji et al., 2022), suggesting that participants deliberately delay their responses so as to allow time for the movement to be aborted if needed (Gulberti et al., 2014; Özyurt et al., 2003) and can even flexibly adjust their reaction speed to shift the balance in favor of responding or stopping behavior (Corneil et al., 2013; Leotti and Wager, 2010).

By constraining the decision about when to act as much as possible, we observed comparable and highly correlated speeds between action cancellation and action initiation, suggesting that stopping and going may reflect two opposing states of a single process supporting a whether decision, rather than reflecting two distinct processes. The *whether* decision about acting or not acting, together with the decision about *when* to act and the decision about *what* to act are recognized as three independent aspects of action control in self-generated behaviors (Brass and Haggard, 2008; Haggard, 2008). Recent behavioral and neurophysiological work has established that, even in reactive tasks, the process of deciding *what* action to take is mechanistically distinct from the process of deciding *when* to act (Ames et al., 2019; Elsayed et al., 2016; Haith et al., 2016; Haith and Bestmann, 2020; Kaufman et al., 2016; Lara et al., 2018). Although both the *when* and *whether* decisions are likely involved in classic stop-signal tasks, it remains unclear whether they are mechanistically dissociable. It also remains to be determined whether changing the decision about *whether* to act or not requires inhibiting the initial choice or just simply switching from one another (see also, MacLeod et al., 2003). We suggest that such processes ought to be considered as bi-directional rather than as a dedicated function of stopping.

Behavior in stop-signal tasks is often interpreted through a race between separate “stop” and “go” processes to reach a threshold (Boucher et al., 2007; Logan and Cowan, 1984). Our findings raise a question about what mechanisms the “go” process represents. In conventional stop-signal tasks, the go process starts accumulating earlier upon an early go signal than the stop process. The stop process, however, with more rapid accumulations, reaches the threshold earlier despite its later start, reflecting its assumed privileged function over the go process. However, the “go” process might conflate multiple distinct aspects of action control, namely the decision about *whether* to respond, *when* to initiate the response, or *what* response to initiate if the task imposes multiple response choices (e.g., Verbruggen et al., 2019). As we have shown, constraining the freedom of when to initiate a response as much as possible yielded an equal amount of time required to cancel an impending action as to generate an action. This result is difficult to be explained by race models, in which a core assumption is that the “stop” process accumulates faster than the “go” process. It is also not clear how this model can accommodate tasks in which a response is required to be made at a pre-determined time. One approach is to model the go process as evolving with a speed that spans the entire duration between the start of the trial and the pre-determined response time (Dunovan et al., 2015), but this rather artificially depends on the experimental setup. In general, the coupling of whether or not an action will occur and the timing of the action in the race model is at odds with recent findings that movement preparation and initiation are separable processes, both neurally and behaviorally (Ames et al., 2019; Elsayed et al., 2016; Haith et al., 2016; Haith and Bestmann, 2020; Kaufman et al., 2016; Lara et al., 2018). Movement preparation is often swift and consumes little time (Carlsen et al., 2008; Carlsen and MacKinnon, 2010; Lara et al., 2018; Wong et al., 2015) and it is not necessary for a go process to reach a fixed threshold level before a response is initiated (Jagadisan and Gandhi, 2017).

In light of our experimental results, the putative neural mechanisms thought to support reactive action inhibition may need to be revisited. In humans, reactive stopping behaviors are thought to be controlled through a particular prefrontal-cortex – basal-ganglia hyperdirect pathway. This pathway is thought to serves as an “emergency brake” that can abruptly abort a no-longerwanted response (Aron, 2007; Aron et al., 2014; Dunovan and Verstynen, 2016; Hannah and Aron, 2021; Jahanshahi et al., 2015; Wessel and Aron, 2017; Wiecki and Frank, 2013). The special function has been often studied through the lens of stop-signal tasks and race models which, as we have argued, contain an embedded bias towards stopping being more urgent and, thus, faster than going. In some cases, circuits thought to be important for stopping have been found to be engaged during both response selection and production (Filevich et al., 2012; Mostofsky and Simmonds, 2008). We suggest that to better understand the function of the hyperdirect pathway, it is critical to examine going and stopping under the same controlled experimental conditions as we have shown here.

In summary, we demonstrate that cancelling an impending response is not any faster than initiating a response. This challenges prevailing beliefs that the ability to prevent oneself from initiating an action has a privileged status and dedicated neural mechanisms. Instead, we propose a more parsimonious explanation that the ability to cancel an intended response may reflect a more general decision making about whether to act or not. This parsimonious explanation extends the recent finding in reactive tasks that deciding when and how to act are independent of one another, by adding a third component as whether or not to act in defining our behaviors. This parallels theories of volitional control which posits distinct “what”, “when” and “whether” decisions underlying self-generated intentional behavior (Brass and Haggard, 2008).

Methods

Participants**:** Thirty-six right-handed participants (15 female; 1 non-binary) between 18 and 41 years of age took part in the study. The experimental procedure was approved by the Johns Hopkins School of Medicine Institutional Review Board. All participants gave written informed consent and received $15 per hour for their participation. Data from one participant who did not follow the task instructions well were excluded from analyses (Fig. S1 and S2).

General procedures**:** Participants sat in front of a laptop with a gray screen and with a key pad next to it. The key pad was positioned so that participants could comfortably rest the index finger of their right hand on a mechanical key mounted on the key pad. On every trial, a white target line was placed with the same distance from the bottom of the screen and a circle was displayed at the top center of the screen (Fig. 1). Once the trial started, the circle moved downwards vertically and participants were asked to either press the key or do nothing when the moving circle reached the target line. The circle stopped moving once a response was registered, or it kept moving toward the bottom of the screen.

For trials where a response was required, a red cross mark was shown if the circle did not intersect the line when it was stopped by a response or if the circle left the screen with no response having been generated, while a green check mark appeared if the circle intersected the target line when it was stopped by a response. The feedback was used to encourage participants to respond with accurate timing and minimize tendencies to delay their response in order to gain more time to make decisions. In trials in which no response was required, a green check mark was displayed if the circle left the screen without a response having been made, while a red cross was displayed if any response was generated at all. Whether or not a response was required in a given trial depended on the color of the circle (white or black). Since perceptual processing plays a critical role in motor response inhibition (Salinas and Stanford, 2013), we counterbalanced the association between response and the color of the moving stimulus so that for half of the participants, white color cued a response (i.e., press the key) and black color indicated that no response was needed (i.e., do not press the key), while this association was reversed for the other half participants, so as to control for the potential perceptual differences between black and white colors.

Criterion task**:** Before the experimental trials began, participants completed two criterion blocks. In these two blocks, all trials required a response in order for participants to become familiar with the timing requirement of the response. The meaning of the color used in this task was consistent with that used in subsequent tasks for each individual. In the first and easier criterion block, the moving circle started from the top center of the screen and dropped toward the bottom of the screen with a constant speed, which took 900 ms in total. A white target line was placed 750 ms from the top and thus 150 ms from the bottom. The circle diameter was sized such that it took 120 ms for the circle to move across the target line. The block ended with five consecutive correct responses (i.e., any part of the circle stopped on the line). Participants then performed the second and more difficult criterion block which matched the conditions of the main experiment, i.e., the diameter of the circle was reduced to 60 ms and it took 500 ms from the trial onset to the center of the circle intersecting the target line and another 120 ms to the bottom of the screen. Similarly, 5 consecutive correct trials were required to end this block.

After successfully completing these two criterion blocks, participants then performed the main task with a response-to-no response (R-to-NR) condition and a no response-to-response (NR-to-R) condition, the order of which were counterbalanced across participants. Each condition consisted of 6 blocks and each block had 100 trials.

R-to-NR Condition**:** This task is also known as the adaptive stop-signal task (Coxon et al., 2006; Leunissen et al., 2017; Slater-Hammel, 1960) and has been used to examine how fast participants can decide to cancel a prepared response that was originally planned to be executed. The moving circle always started with the color that cued a response (white for half of participants and black for the other half). In a random ~30% of trials (204 out of 600 trials), the circle turned to the not-responding color while it was moving towards the target line. The time of color switch before the center of the circle intersected the target line was randomly drawn from a uniform distribution between 50ms to 450 ms with a step size of 16.7ms. The choice of this step size was constrained by the refresh rate of the monitor, which was 60 Hz. Thus, there were 28 possible time points at which the circle color changed. The closer the time point was from the targe line, the shorter time available to make a decision.

NR-to-R Condition: This task, conceptually similar to the timed-response task commonly used in motor reaching task (Ghez et al., 1997; Haith et al., 2016), was used to examine how fast participant can initiate a response. Trials started with the circle defaulted to the not-responding color (white for half of participants and black for the other half) and switched to the responding color in a random subset of trials. Consistent with the R-to-NR condition, the proportion of color-switch trials was ~30% of trials (204 out of 600 trials) and the time of color change ranged from 50 to 450 ms. These switch trials and their corresponding color change times were matched between these two conditions on a trial-by-trial basis. One participant showed clear evidence of guess the required response in the NR-to-R condition (Figs. S1 and S2) and we therefore excluded this participant from subsequent analysis.

Data analysis:

Speed–accuracy trade-off: Our primary analysis was focused on the trials in which the color switched in both conditions. We assessed the time course over which participants were able to abort an impending action in the R-to-NR condition by constructing a speed–accuracy trade-off relating the time available to cancel a response and the probability of the response being successfully cancelled. For visualization purposes, we estimated this speed–accuracy trade-off using a 50 ms sliding window on the time available to cancel a response (i.e., forced RT, see below for details). Similarly, the speed–accuracy trade-off can also reveal how rapidly a response can be initiated in the NR-to-R condition.

Response correctness: In our original speed–accuracy trade-off analysis (Fig. 2), a trial in which a response was required was considered to be successful only if a response was made while the circle was not below the target line (i.e., within a -30 ms time window below the target line). In practice, participants often did generate a response, but did so after the circle was no longer overlapping the line. We designated these trials as failures and considered them to be equivalent to not generating a response at all. However, we designated a trial that does not require a response as correct if participants did not press the button before the trial ended. This non-produced response may be more than 30 ms after the target line if it were generated.

This designation of correctness included a strong requirement of timing accuracy when initiating a response in the NR-to-R condition but not when cancelling a response in the R-to-NR condition (Fig. 2C). This asymmetry in the analysis may have affected our estimation of the relative timing of going and stopping. To match the timing requirement, we further used a -35 ms time window below the target line. Results from these two different criterions for correctness are consistent with one another (Fig. S3 and Main Text).

Forced RT: By manipulating the time at which circle changed its color (between 50 ms and 450 ms before the targeted line), we forced participants to cancel an intended response in the R-to-NR condition or press a button in the NR-to-R condition within a particular amount of time, referred to as forced RT. In the R-to-NR condition, when participants failed to cancel an impending response, the forced RT was quantified as the time elapsed from color change to the time of the button press. When no response was generated, the actual forced RT was not observable and so instead the forced RT was approximated by the intended forced RT, i.e., the time interval between the center of the circle and the target line at the moment of color change. In the NR-to-R condition, the forced RT was calculated as the time interval between color change and the time of response if participants pressed a button before or when the circle reached the target line, whereas it was approximated as the interval between color change and the target line (i.e., intended forced RT) if participants did not generate a response or the response was made later than the target line (i.e., the circle was no longer overlapping the target line).

When using the intended RT described above, we inherently assumed that the not-produced response would have had accurate timing, if it was produced, which, however, is not true in reality.

All participants had an idiosyncratic tendency to respond consistently earlier or later than the target line (Fig. S4). To better approximate the true RT, we first calculated how much later or earlier each participant responded to trials in non-switch trials in which the circle was in the responding color throughout the trial (i.e., 396 out of 600 trials, about 70% in the R-to-NR condition). From these measurements, we randomly drew a sample and added it to an intended RT to approximate the unobservable true forced RT. We repeated this bootstrapping process 1000 times for each individual and modelled the mean speed–accuracy trade-off.

Modelling speed–accuracy trade-off: To quantify the speed–accuracy trade-off, we assumed that cancelling an intended response in the R-to-NR condition occurred at a random time $T_{NR} \sim \mathcal{N}(\mu_{NR}, \sigma_{NR}^2)$. A response would be correctly aborted with probability β_{NR} (close to 1) if forced RT was long than T_{NR} and with a probability α_{NR} (close to 0) if the available forced RT was shorter than T_{NR}. Thus, the probability, in trial i, of observing a correct response cancellation (c = 1), given the preparation time (t^i) is given by:

$$p_{NR}\left(c^i \middle| t^i\right) = \alpha_{NR} p\left(t^i \le T_{NR}\right) + \beta_{NR} p\left(t^i > T_{NR}\right)$$

$$= \alpha_{NR}(1 - \Phi_{NR}\left(t^i | \mu_{NR}, \sigma_{NR}^2\right)) + \beta_{NR}\, \Phi_{NR}\left(t^i | \mu_{NR}, \sigma_{NR}^2\right)$$

where $\Phi_{NR}\left(t^i \mid \mu_{NR},\ \sigma_{NR}^2\right)$ is the cumulative normal distribution of T_{NR}.

Similarly, in the NR-to-R condition, the probability of correctly initiating a response given the preparation time (t^i) is:

$$p_R\left(c^i \middle| t^i\right) = \alpha_R p\left(t^i \le T_R\right) + \beta_R p\left(t^i > T_R\right) = \alpha_R(1 - \Phi_R\left(t^i\right)) + \beta_R\, \Phi_R\left(t^i\right)$$

where $\Phi_R\left(t^i\right) = \Phi_R\left(t^i \mid \mu_R,\ \sigma_R^2\right)$ is the cumulative normal distribution of T_R.

We estimated the parameters using maximum likelihood estimation with the MATLAB function fmincon.

Modelling time of response: Time of response was the time elapsed from the trial onset to the time at which the response was made, if any. In our speed–accuracy trade-off analysis, we relied on a proxy RT for trials in which a response was not produced. To avoid this reliance, we also estimated the speed of making a response in the NR-to-R condition by fitting the time of response, y, with two linear functions, which intersected at an intended forced RT of μ_T:

$$y = \begin{cases} \mu_0 + \beta_1(t - \mu_T) + \mathcal{M}(0, \sigma_{rt}^2, \delta), & t < \mu_T \\ \mu_0 + \beta_2(t - \mu_T) + \mathcal{N}(0, \sigma_r^2), & t \ge \mu_T \end{cases}$$

In this model, we took the time of transition between these components, μ_T, as the minimum time at which participants could initiate a response on time following the color change. At μ_T, a response needed to be made without any delay so that it could land on the target line. μ_0 is the time of response when $t = \mu_T$. We assumed that for $t < \mu_T$, participants behaved in a reactive manner to the appearance of stimulus and that, therefore, the time of response would follow a typical reaction time distribution. In particular, we observed that some participants occasionally generated times of response that were longer than most of their responses, so we assumed that the residual term in the upper equation followed an exponentially modified Gaussian distribution $\mathcal{M}(0, \sigma_{rt}^2, \delta)$. This choice did not lose its generality if participants did not generate some uncommon late responses, because when δ is close to zero, $\mathcal{M}$ approaches to be a Gaussian distribution. When $t \geq \mu_T$, participants would time the response and press the button at around the target line. Thus, we assumed that residual term of the lower equation in this case followed a Gaussian distribution $\mathcal{N}(0, \sigma_r^2)$. Parameters μ_0, β_1, $\beta_2, \mu_T, \sigma_{rt}, \sigma_r, \delta$, were estimated by maximum likelihood estimation with MATLAB function fmincon. To avoid a local minimum estimation, we ran the maximum likelihood estimation with 100 random starting values. A parameter recovery analysis indicated that our model fitting yielded unreliable estimation of true parameters (Fig. S5). We found, based on parameter recovery, that it was better to constrain σ_{rt} to be greater than 0 (lower bound of 0.005), in order to avoid poot quality fits. For the same reason, we also regularized the fits by penalizing the log-likelihood with:

$$LL^* = LL - \gamma(\sigma_{rt} - 0.03)^2 - \gamma(\delta - 0.03)^2 - \gamma(\mu_0 - 0.5)^2$$

σ_{rt} and δ were included to avoid unrealistic estimation of $\sigma_{rt} \approx 0$ and $\delta \approx 0$ (Fig. S5). We chose 0.03 for σ_{rt} and δ as it was the mean value of initial estimation across participants. We also regularized μ_0 and set it to 0.5 s because our data showed that participants tended to respond around the target line given a long enough RT (Fig. S4). We set $\gamma = 2000$, which avoided overfitting these three parameters to the particular value we selected. Parameter recovery demonstrated that this regularized fitting procedure led to reliable estimation of the true parameters when applied to synthetic data (Fig. S6).

Statistical analysis: Data (e.g., μ_{NR} vs. μ_R) were analyzed using paired t-test at the significant level of $\alpha = 0.05$ after examining the normality of samples. Because non-significant outcomes from hypothesis testing does not necessarily mean two samples are not different from one another, we further conducted the equivalence and noninferiority test (Lakens, 2017; Walker and Nowacki, 2011). Power analysis indicates that the sample size n = 35 (out of 36) had 80% power to detect an effect size of 0.7 between conditions (Lakens, 2017). Therefore, we set the upper and lower bounds of the equivalence and noninferiority test as 0.7 and -0.7, corresponding to the equivalence bound between 16.5 ms and -16.5 ms in the unit of reaction time.

Author Contributions

Y.D., A.D.F., and A.M.H. conceptualized the experiment; A.D.F. programmed the task; Y.D., A.D.F., and D.M.M. collected data; Y.D. performed data and statistical analyses; Y.D. prepared the figures; Y.D. drafted the manuscript; Y.D., A.D.F., D.M.M., and A.M.H. revised the manuscript and approved final version of the manuscript.

Competing Interests

The authors declare no competing interests.

References

Ames KC, Ryu SI, Shenoy KV. 2019. Simultaneous motor preparation and execution in a last-moment reach correction task. *Nature communications* **10**:1–13.

Aron AR. 2007. The neural basis of inhibition in cognitive control. *The neuroscientist* **13**:214–228.

Aron AR, Robbins TW, Poldrack RA. 2014. Inhibition and the right inferior frontal cortex: one decade on. *Trends in cognitive sciences* **18**:177–185.

Boucher L, Palmeri TJ, Logan GD, Schall JD. 2007. Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychological review* **114**:376–397.

Brass M, Haggard P. 2008. The what, when, whether model of intentional action. *The Neuroscientist* **14**:319–325.

Carlsen AN, Chua R, Inglis JT, Sanderson DJ, Franks IM. 2008. Motor preparation in an anticipation-timing task. *Experimental brain research* **190**:453–461.

Carlsen AN, Chua R, Inglis JT, Sanderson DJ, Franks IM. 2004. Prepared movements are elicited early by startle. *Journal of motor behavior* **36**:253–264.

Carlsen AN, MacKinnon CD. 2010. Motor preparation is modulated by the resolution of the response timing information. *Brain research* **1322**:38–49.

Corneil BD, Cheng JC, Goonetilleke SC. 2013. Dynamic and opposing adjustment of movement cancellation and generation in an oculomotor countermanding task. *Journal of Neuroscience* **33**:9975–9984.

Coxon JP, Stinear CM, Byblow WD. 2006. Intracortical inhibition during volitional inhibition of prepared action. *Journal of neurophysiology* **95**:3371–3383.

Dunovan K, Lynch B, Molesworth T, Verstynen T. 2015. Competing basal ganglia pathways determine the difference between stopping and deciding not to go. *Elife* **4**:e08723.

Dunovan K, Verstynen T. 2016. Believer-skeptic meets actor-critic: rethinking the role of basal ganglia pathways during decision-making and reinforcement learning. *Frontiers in neuroscience* **10**:106.

Elsayed GF, Lara AH, Kaufman MT, Churchland MM, Cunningham JP. 2016. Reorganization between preparatory and movement population responses in motor cortex. *Nature communications* **7**:1–15.

Filevich E, Kühn S, Haggard P. 2012. Negative motor phenomena in cortical stimulation: implications for inhibitory control of human action. *cortex* **48**:1251–1261.

Ghez C, Favilla M, Ghilardi MF, Gordon J, Bermejo R, Pullman S. 1997. Discrete and continuous planning of hand movements and isometric force trajectories. *Experimental Brain Research* **115**:217–233.

Gulberti A, Arndt P, 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. *Frontiers in human neuroscience* **8**:61.

Haggard P. 2008. Human volition: towards a neuroscience of will. *Nature Reviews Neuroscience* **9**:934–946.

Haith AM, Bestmann S. 2020. Preparation of movement In: Poeppel D, Mangun G, Gazzaniga M, editors. The Cognitive Neurosciences. MIT Press.

Haith AM, Pakpoor J, Krakauer JW. 2016. Independence of movement preparation and movement initiation. *Journal of Neuroscience* **36**:3007–3015.

Hannah R, Aron AR. 2021. Towards real-world generalizability of a circuit for action-stopping. *Nature Reviews Neuroscience* **22**:538–552.

He JL, Hirst RJ, Puri R, Coxon J, Byblow W, Hinder M, Skippen P, Matzke D, Heathcote A, Wadsley CG. 2021. OSARI, an Open-Source Anticipated Response Inhibition Task. *Behavior research methods* 1–11.

Jagadisan UK, Gandhi NJ. 2017. Removal of inhibition uncovers latent movement potential during preparation. *Elife* **6**:e29648.

Jahanshahi M, Obeso I, Rothwell JC, Obeso JA. 2015. A fronto–striato–subthalamic–pallidal network for goal-directed and habitual inhibition. *Nature Reviews Neuroscience* **16**:719–732.

Kaufman MT, Seely JS, Sussillo D, Ryu SI, Shenoy KV, Churchland MM. 2016. The largest response component in the motor cortex reflects movement timing but not movement type. *ENeuro* **3**.

Lakens D. 2017. Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social psychological and personality science* **8**:355–362.

Lara AH, Elsayed GF, Zimnik AJ, Cunningham JP, Churchland MM. 2018. Conservation of preparatory neural events in monkey motor cortex regardless of how movement is initiated. *Elife* **7**:e31826.

Leotti LA, Wager TD. 2010. Motivational influences on response inhibition measures. *Journal of Experimental Psychology: Human Perception and Performance* **36**:430–447.

Leunissen I, Zandbelt BB, Potocanac Z, Swinnen SP, Coxon JP. 2017. Reliable estimation of inhibitory efficiency: to anticipate, choose or simply react? *European Journal of Neuroscience* **45**:1512–1523.

Logan GD, Cowan WB. 1984. On the ability to inhibit thought and action: A theory of an act of control. *Psychological review* **91**:295–327.

Luce RD. 1991. Response times: Their role in inferring elementary mental organization. Oxford University Press, USA.

MacLeod CM, Dodd MD, Sheard ED, Wilson DE, Bibi U. 2003. In opposition to inhibition.

Matzke D, Strickland LJG, Sripada C, Weigard AS, Puri R, He J, Hirst R, Heathcote A. 2021. Stopping timed actions. *PsyArXiv*. doi:https://doi.org/10.31234/osf.io/9h3v7

Mostofsky SH, Simmonds DJ. 2008. Response inhibition and response selection: two sides of the same coin. *Journal of cognitive neuroscience* **20**:751–761.

Özyurt J, Colonius H, Arndt PA. 2003. Countermanding saccades: evidence against independent processing of go and stop signals. *Perception & Psychophysics* **65**:420–428.

Rawji V, Modi S, Rocchi L, Jahanshahi M, Rothwel J. 2022. Proactive inhibition is marked by differences in the pattern of motor cortex activity during movement preparation and execution. *Journal of Neurophysiology* 819–828.

Salinas E, Stanford TR. 2013. The countermanding task revisited: fast stimulus detection is a key determinant of psychophysical performance. *Journal of Neuroscience* **33**:5668–5685.

Slater-Hammel AT. 1960. Reliability, accuracy, and refractoriness of a transit reaction. *Research Quarterly American Association for Health, Physical Education and Recreation* **31**:217–228.

Valls-Solé J, Solé A, Valldeoriola F, Muñoz E, Gonzalez LE, Tolosa ES. 1995. Reaction time and acoustic startle in normal human subjects. *Neuroscience letters* **195**:97–100.

Verbruggen F, Aron AR, Band GP, Beste C, Bissett PG, Brockett AT, Brown JW, Chamberlain SR, Chambers CD, Colonius H. 2019. A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *Elife* **8**:e46323.

Walker E, Nowacki AS. 2011. Understanding equivalence and noninferiority testing. *Journal of general internal medicine* **26**:192–196.

Welford A. 1980. Choice reaction time: Basic concepts. *Reaction times* 73–128.

Wessel JR, Aron AR. 2017. On the globality of motor suppression: unexpected events and their influence on behavior and cognition. *Neuron* **93**:259–280.

Wiecki TV, Frank MJ. 2013. A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychological review* **120**:329–355.

Wong AL, Haith AM, Krakauer JW. 2015. Motor planning. *The Neuroscientist* **21**:385–398.

Supplement Figures:

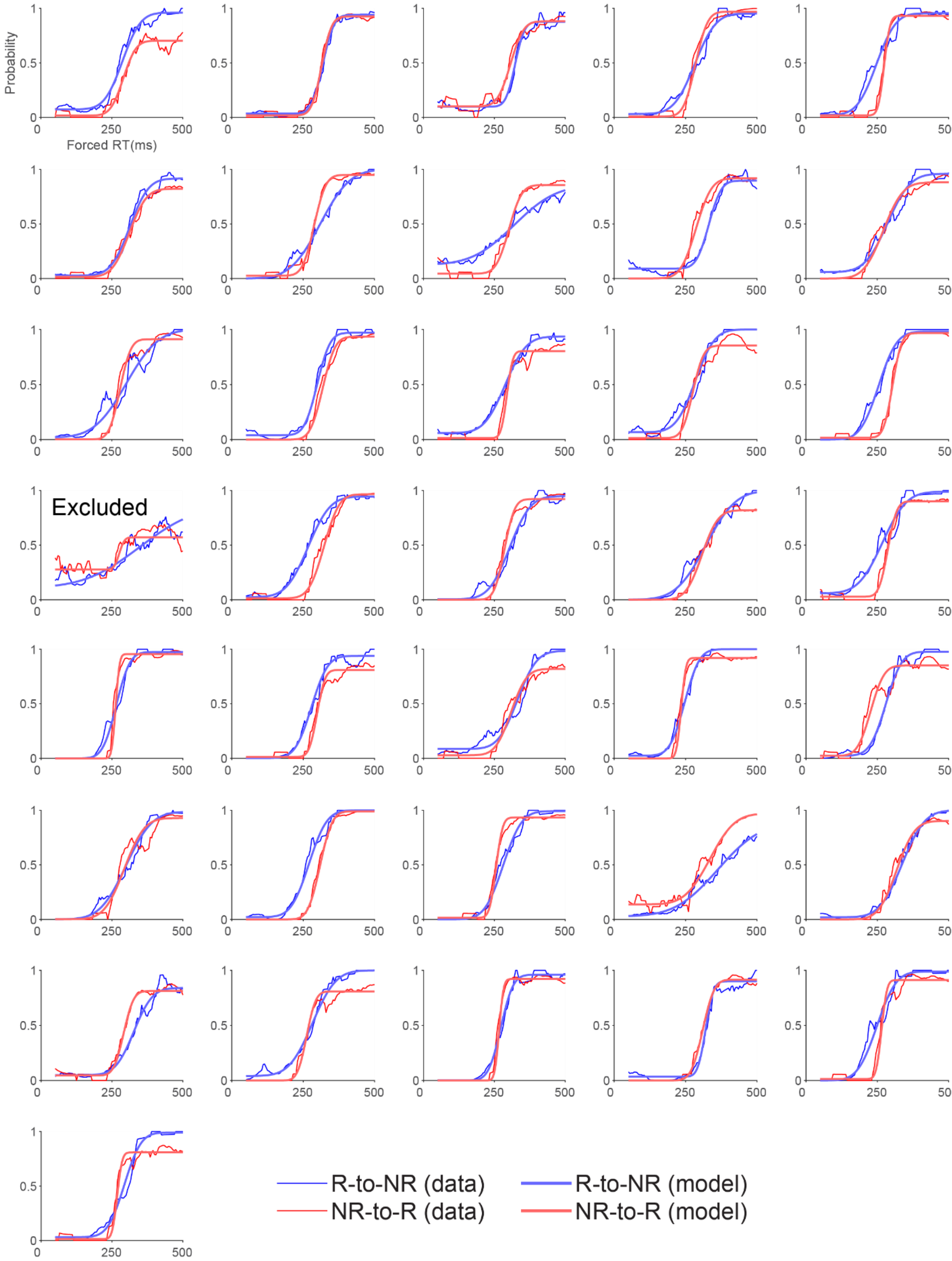

Fig. S1. Speed–accuracy trade-off for each individual. Most participants' performance (thinner lines) was similar to the exemplar participant shown in Fig. 2. The accuracy was close to zero when forced RT was very short (i.e., < 100 ms) and it increased with longer forced RT. However, one participant exhibited abnormally high accuracy even when forced RT was less than 100 ms, suggesting that this particular participant randomly guessed whether to initiate or cancel a response instead of following the instruction for each condition. We excluded this participant from further analysis.

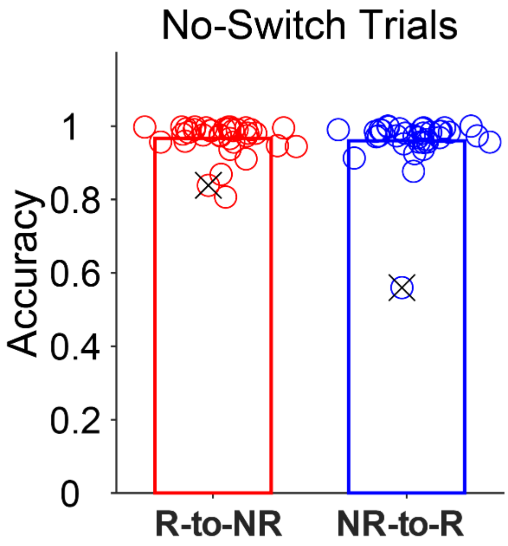

Fig. S2. Response correctness in trials that stimulus circle did not change color. The high accuracy in no-switch trials (i.e., Making a response in the R-to-NR condition or not making a response in the NR-to-R condition) revealed that participants did not behave randomly except one participant, who exhibited a chance-level performance in the NR-to-R condition. This participant was the one who also produced higher accuracy in switch trials even when forced RT was very short shown in Fig. S1. Since we aimed to compared the individual-wise performance between two tasks, this participant's data from both conditions (crossed circles) were excluded for our analyses reported in main texts.

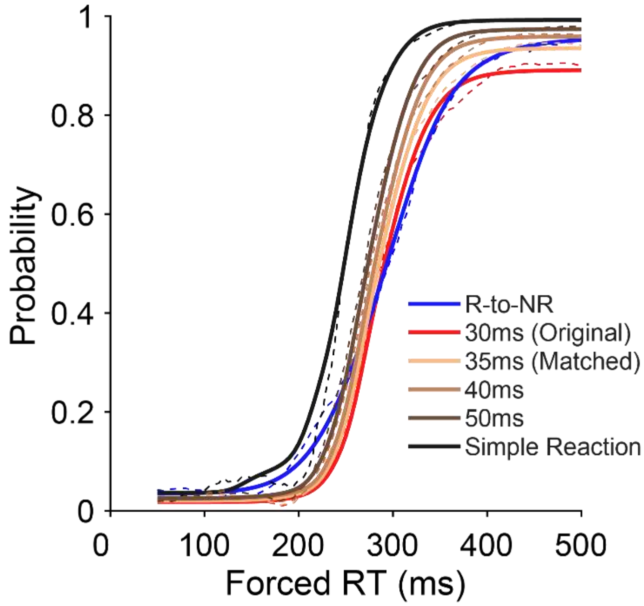

Fig. S3. The effect of correctness criterion on the estimated speed of response initiation. A) The shape of the speed–accuracy trade-off in the NR-to-R condition depended on how we defined a correct responding trial. The original tolerance was 30ms (i.e., a half size of the stimulus circle). Correctness defined on this small-time window did not include trials where a response was made 30 ms or more later below the target line, resulting in a lower accuracy in the NR-to-R condition than the R-to-NR condition when RT was long enough (red vs. blue). To match accuracy rates across conditions, we instead used a timing tolerance of 35 ms that yielded a similar accuracy level between conditions for trials with long RT (light brown). In this case, we still found that the timing of response cancellation was not different from the timing of response generation (blue vs. light brown). In addition, the original estimation of response initiation speed was around 285 ms. This slow speed was, at least partially, caused by the timing requirement of initiating a response on time in the NR-to-R condition. Further broadening timing tolerances from 35 ms to 50 ms led to superior speed–accuracy trade-offs with faster mean speeds of responding than the original (brown lines). The speed was even faster (248.7 ms ± 31.7 ms) when all trials with a response at any time were considered correct (the ’simple reaction’ mode; black line). Solid lines: model fitting; Dashed lines: data.

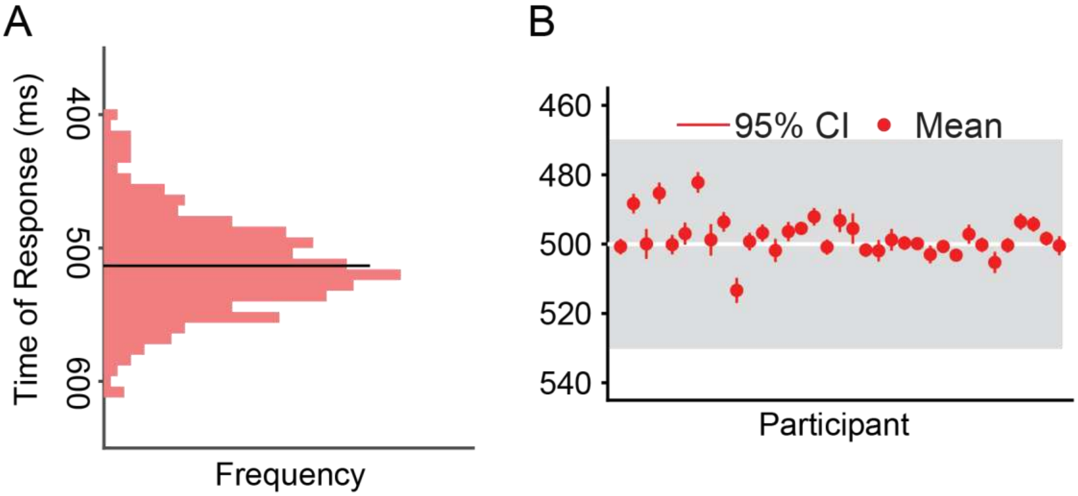

Fig. S4. Idiosyncratic tendencies to respond consistently earlier or later than the target line. A) Data from the same exemplar participant as in Fig. 2B. The time of response, measured for trials that the circle started and stayed as the same color that required a response in the R-to-NR condition, was calculated as the time interval between trial onset and the time at which a response was made. This participant tended to respond consistently slightly later than the target line (i.e., 500 ms). B) The mean time of response for each individual participant. All participants had an idiosyncratic tendency to respond consistently earlier or later than the target line, although the circle still overlapped the target line when they responded (grey area representing the diameter of 60 ms of the circle).

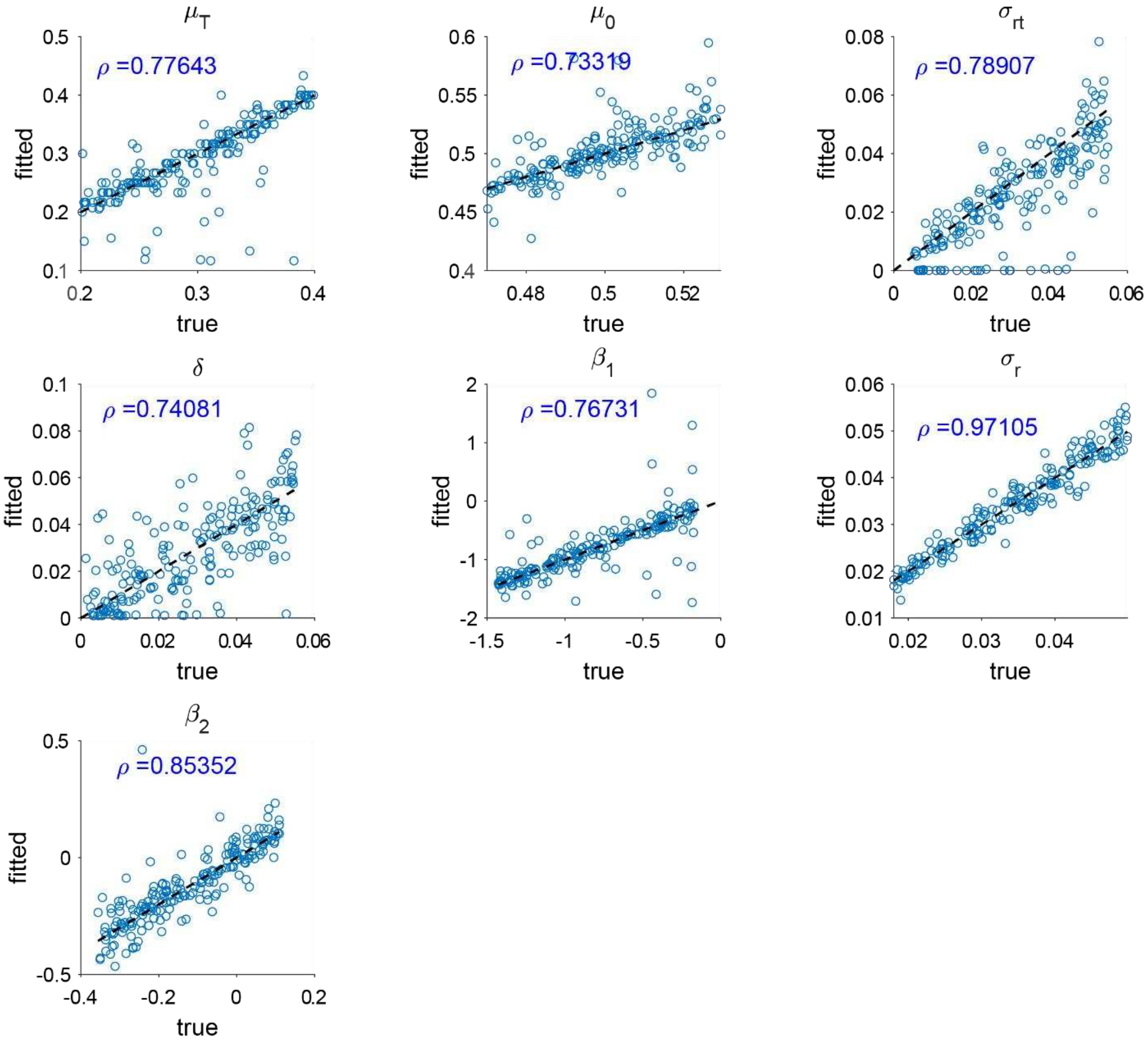

Fig. S5. Parameter recovery of the model used to fit time of response data without regularization terms. We used the model to generated synthetic datasets matching the amount of data collected from each participant. We generated datasets based on a range of true underlying parameter values and then used maximum likelihood estimation to try to recover the true underlying parameters. Each panel shows a different parameter (unit of measurement: second), with true value used in the simulation on the x-axis and the estimated value on the y-axis. For most parameters, including the key parameter of interest, μ_T, parameter recovery was not accurate.

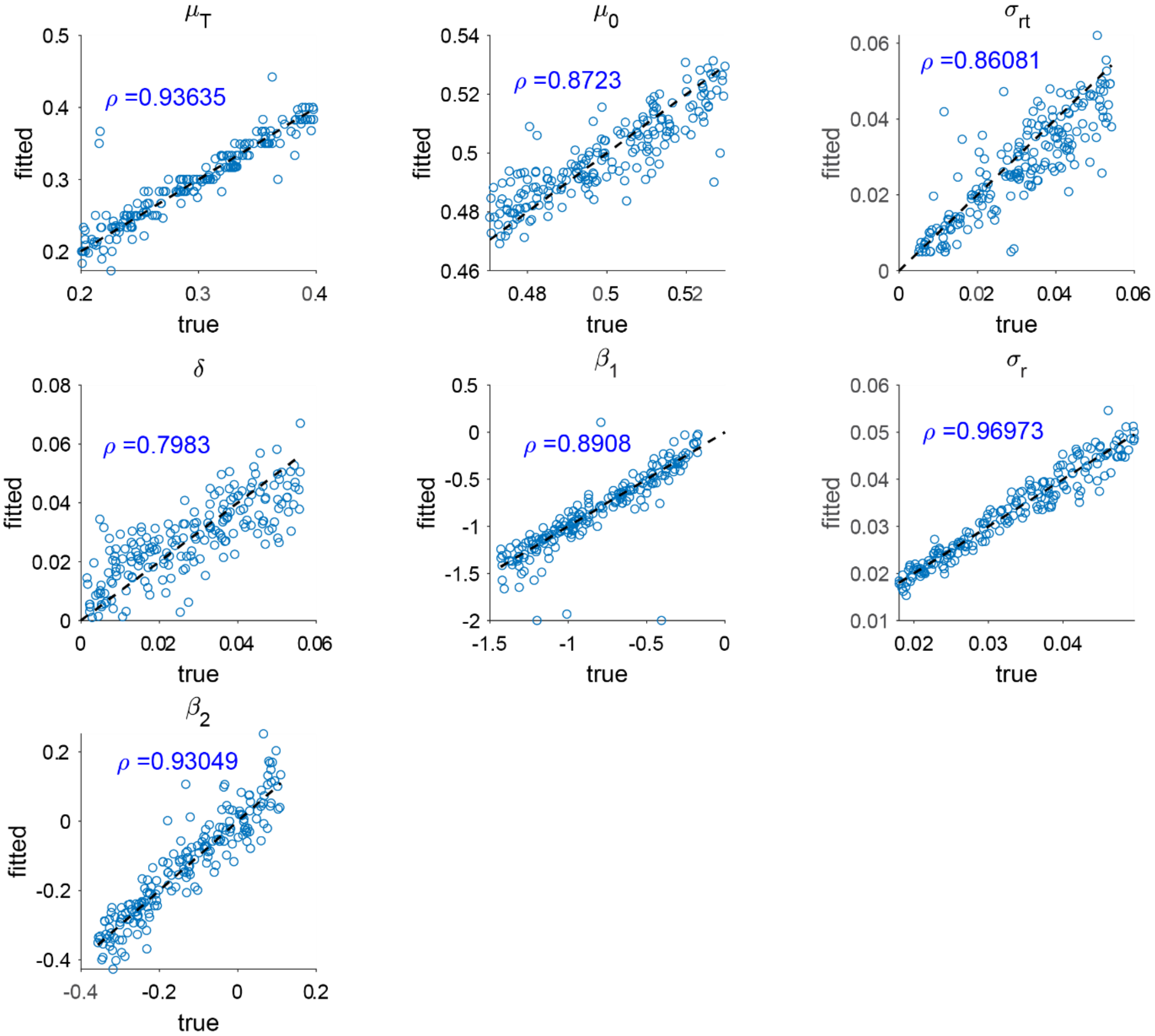

Fig. S6. Parameter recovery of the model used to fit time of response data with regularization terms. After adding regularization terms with respect to μ_0, σ_{rt}, and δ, the reliability of the parameter estimation became notably improved. The correlation between true values and estimated values of $\mu_T = 0.94$, which is much higher than the corresponding value of 0.78 in the model without regularization (Fig. S5).