When going can be faster than stopping: Rethinking response inhibition as a flexible decision about whether or not to act

Yue Du, Alexander D. Forrence, Delaney M. Metcalf, Adrian M. Haith
Department of Neurology, Johns Hopkins University School of Medicine

The ability to cancel or inhibit a voluntary action moments before it is initiated is widely considered to be a fundamental aspect of action control[1]. A pedestrian about to step into the street must swiftly prevent themselves from doing so if a fast car suddenly approaches. Experimental paradigms such as the stop-signal task have clearly established that suppressing an action can be achieved very rapidly, with people able to cancel a prepared movement even if the signal to cancel is received during the reaction time before movement[2]. Classical accounts have posited the existence of a dedicated mechanism, along with dedicated neural pathways[3] for inhibiting movement that acts as an "emergency brake" to prevent a movement from being initiated. In computational terms this has been characterized in terms of a race between separate "stop" and "go" processes to reach a threshold – with a movement being prevented if the stop process hits the threshold before the "go" process triggers initiation of the movement.

Here, we propose that the ability to withhold a prepared movement does not simply arise from an "emergency brake" mechanism for inhibiting movement, but instead reflects a more general decision about whether or not to act. Recent work has established that, in cued reaching tasks, movement preparation occurs independently of movement initiation[4]. That is, deciding *what* action to take (preparation) is independent from deciding *when* to act (initiation). The dual nature of such decisions has long been appreciated in the context of self-generated behaviors[5]. These theories also include deciding *whether* to act as a third decision that is also independent of the other two. We therefore supposed that, even in reactive tasks, preventing initiation of a movement might reflect a decision process independent of both what action to take and when to take it. This idea differs from prevailing accounts of action inhibition which imply that whether or not an action occurs is closely coupled to when it might occur. Viewing stopping as a decision about whether to act also makes a critical prediction: that not only could participants rapidly abort a pre-planned movement that they were about to execute, they should also be able to rapidly deploy a prepared action that they had previously intended to withhold. Here, we tested this prediction using a timed-response approach where participants (N=36) had to decide whether or not to press a key at a prescribed time, given the color of a cue, eliminating any decision about when to act.

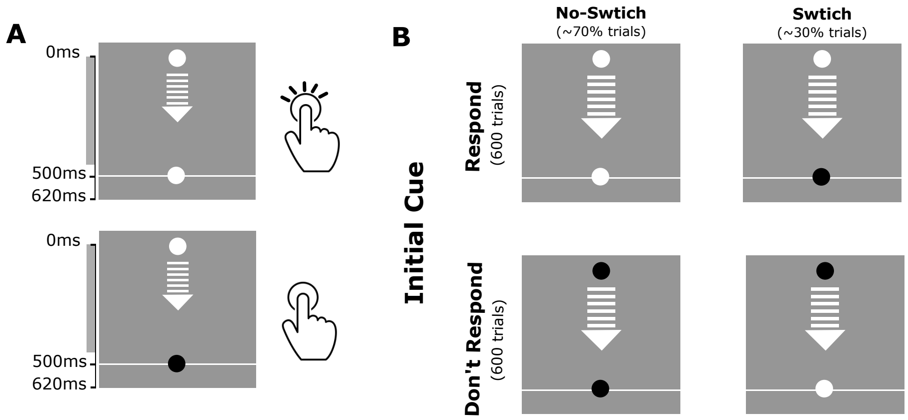
Participants viewed a circle moving vertically downward to cross a horizonal line. They were asked to either press a button or do nothing when it reached the target line, depending on the color of the circle (e.g., white = response; black = no response; color counterbalanced across participants; Fig. 1). Unlike most stopping tasks that merely focus on stopping, a trial could begin as either a "Response" or "No Response" color, varying block-wise. In the majority of trials (~70%), the color remained the same throughout the trial. On occasional "switch" trials (~30%), however, the circle changed color before hitting the target line, forcing participants to revise their decision about whether or not to press the button when the circle crossed the line – either from intending to respond to not responding ("R-to-NR" condition, similar to conventional stopping paradigms) or from intending to not respond to responding ("NR-to-R" condition). We varied the amount of the time available to revise this "whether" decision by manipulating when the color changed (between 50 ms and 500 ms before the targeted line). This allowed us to construct a speed-accuracy tradeoff which relates available decision time (DT) to corresponding probability of correct "whether" decisions.

In both types of switch trials, participants almost always failed to make the right decision when DT was very small (<100ms), but almost always made the right decision when allowed sufficient time to decide (>300ms) (Fig. 2A & 2B). The speed-accuracy tradeoffs, generated using a 50 ms sliding window, revealed that participants were faster at switching from not responding to responding than the other way around (larger area under the curve in the NR-to-R than the R-to-NR condition; $t_{34} = 8.73$, $p < 10^{-8}$, CI: [28.75, 46.20]; Fig. 2C). Using a simple parametric model, in which we assumed that the decision to respond or not could be thought of as a discrete event occurring at a random time, we estimated that the time required to go from not responding to responding (263.4 ms ± 25.2 ms; mean ± s.d.) was around 30ms shorter than the time required to go from responding to not responding (296.7 ms ± 28.8 ms ; mean ± s.d.; t34 = -8.22, paired t-test, $p < 10^{-7}$, 95% CI: [-41.4ms, -25.0ms]; Fig. 2D). These results demonstrate that, contrary to prevailing accounts, people are in fact able to commit to generating an action *more* rapidly than they are able to withhold an action.

Important potential confounds to these results are that randomly guessing whether to act and strategically delaying responses to buy additional time for correct action cancellation could corrupt our estimates of decision speed (Fig. 3B). However, analyses on the probabilities of choice and the distributions of response time (RT) confirmed that participants established a default tendency to act or not within each condition (Fig. 3A). Meanwhile, they did not strategically delay their responses (Fig. 3C & D). Another potential caveat in the R-to-NR condition (as with any stopping task) is that the actual available DT was not observable. We initially defined the DT as the time interval between color change and the target line when there was no response, inherently assuming that the prepared but later cancelled response would have had accurate timing, if it was misfired (Fig. 4A). This was clearly not the case: The actual DT depends on the virtual RT if the cancelled response had been produced. We approximated the actual DT by the RT in trials when the circle stayed in the Response color (Fig. 4A, pink distribution). The resultant speed-accuracy tradeoff based on the adjusted DT hardly differed from the original analyses (Fig. 4B & 4C).

In summary, our results challenge previous accounts that stopping behaviors are operated via a rapid and dedicated inhibitory process. Instead, it reflects a more general decision about whether or not to act that operates in conjunction with the decision about what and when in defining our behavior. This view, together with recent findings in reactive tasks[4], parallels theories of self-generated behavior which posits distinct "what", "when" and "whether" decisions underlying intentional behavior[5].

Fig.1. Experiment procedure. A) 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 circle color was counterbalanced across participants, controlling for the potential perceptual differences of black and white colors. B) The circle always started with the same color within each block of 100 trials. In ~70% trials, the circle remained the same color throughout. However, in another ~30% trials, the circle changed color before it hit the target line, forcing participants to change their decision about whether or not they would generate a response when the circle crossed the line. By manipulating the amount of time (between 50 ms and 500 ms) available to revise this "whether" decision, 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 when they at first had no intention to ("NR-to-R" condition; lower panel). Participants completed 12 blocks of 100 trials, generating 204 switch trials in each condition.



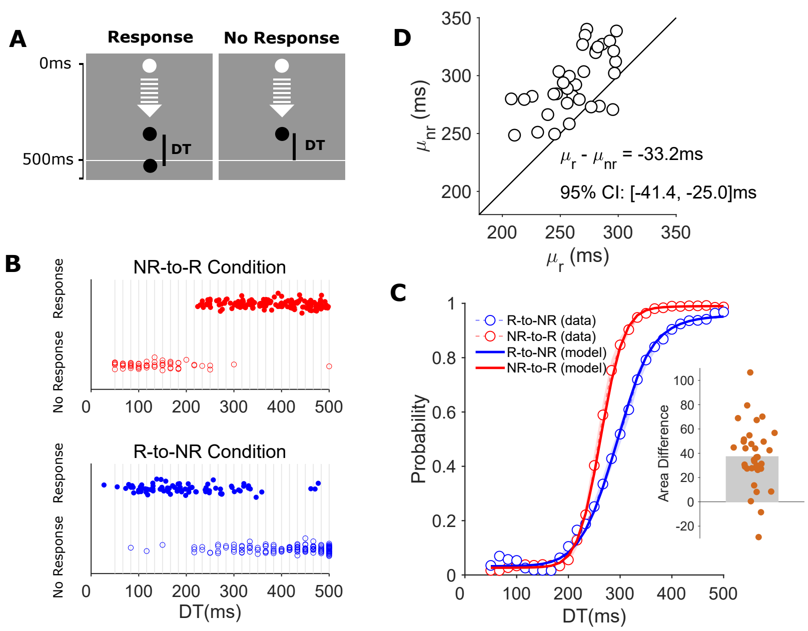


Fig.2. A) The actual decision time (DT) was quantified as the time elapsed from color change to the time of the button press when participants generated a response, and from the color change to the time the center of the circle crossing the target line when participants did not. B) Behavior of one exemplar participant. In trials in which only a very short DT was allowed, this participant consistently made the wrong choice as to whether to respond or not. When a longer DT was allowed, this participant was able to consistently make the correct choice to respond or not. C) Speed-accuracy tradeoffs showing the probability of a correct decision as a function of DT (circles). The approximated area (inset) under the speed-accuracy tradeoff curves was larger in the NoResponse-to-Response (NR-to-R) condition (red) than the Response-to-NoResponse (R-to-NR) condition (blue). A simple computational model where we assumed that the decision to respond or not could be thought of as a discrete event occuring at a random time $T \sim N(\mu, \sigma^2)$ well captured the speed-accuracy tradeoff in both conditions (solid lines). D) We used our model to estimate the average time μ needed to revise the "whether" decisions in R-to-NR and NR-to-R conditions respectively. On average, the time required for switching from not responding to responding (μ_r: 263.4 ms ± 25.2ms; mean ± s.d.) was shorter than the time needed for switching from responding to not responding by about 30 ms (μ_{nr}: 296.7ms ± 28.8 ms; mean ± s.d.)

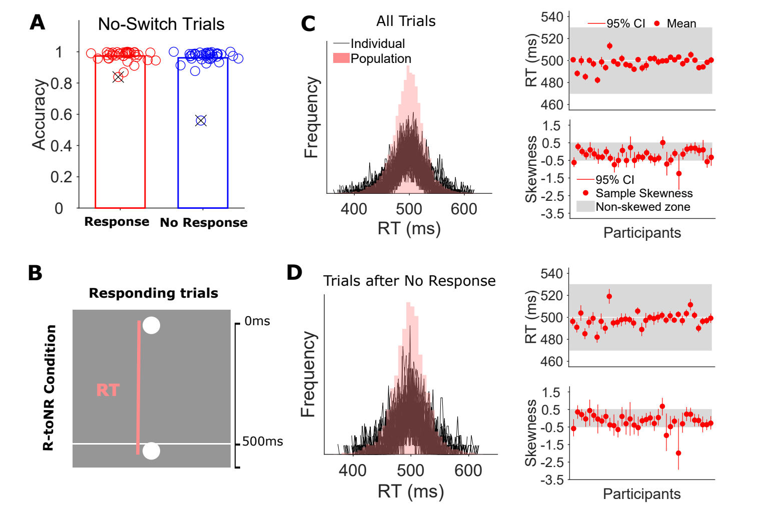


Fig.3. We considered two potential caveats to our results. A) Participants may have tried to anticipate switch trials, which might influence our estimate of μ. However, the high accuracy of choices in no-switch trials revealed that they did not behave randomly except one participant, who we excluded from further analysis. B) Participants may have strategically delayed all responses to gain additional decision time, however, we found that participant's response times (RT) were accurate; bootstrapped 95% CI of individuals' RT bias and skewness contained zero, demonstrating that participants did not delay their responses. D) Similar analyses suggested that participants also did not strategically delay their responses when the preceding trial required the response to be withheld.

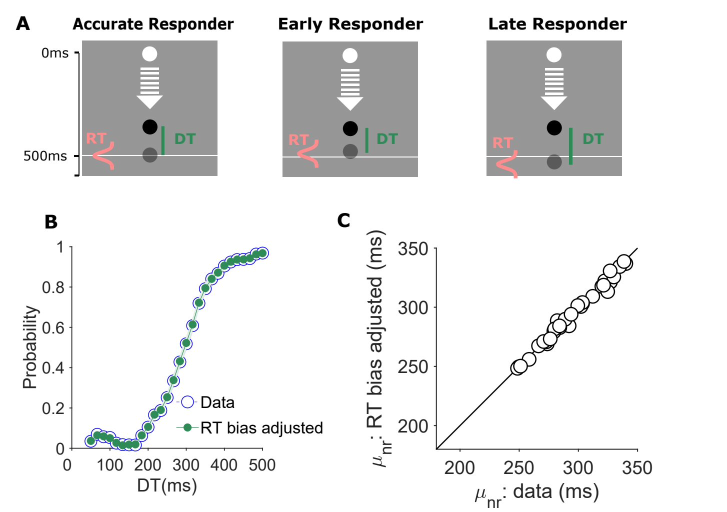


Fig.4. A) In the Response-to-NoResponse (R-to-NR) condition, the actual DT was not observable. Our estimate depended on defining DT as the time interval between color change and the circle crossing the target line in trials where there was no response. This inherently assumes that the prepared but later cancelled response would have had accurate timing, if it has been generated. In reality, participants tended to respond consistently earlier or later than the target line. To account for this, we further approximated the actual DT for each individual using their RT in trials when they did response. B) and C) This adjustment had very little effect on the inferred speed-accuracy tradeoff.

Reference: 1. Logan & Cowan, *Psychol Rev* (1984) 2. Verbruggen et al., *Elife* (2019) 3. Wiecki & Frank, *Psychol Rev* (2013) 4. Haith et al., *J Neurosci* (2016) 5. Brass & Haggard, *The Neuroscientist* (2008)