eLife

Reviewed Preprint
v1 • September 1, 2026
Not revised

For correspondence:
bompasae@cardiff.ac.uk

Competing interests: No competing interests declared

Funding: See page 25

Reviewing editor: Emilio Salinas, Wake Forest University School of Medicine, United States

© 2026, Statham et al. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Non-decision time: the elephant in the executive control room

Heather Statham, Phil Schmid, Petroc Sumner, Aline Bompas

Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology Cardiff University, Cardiff, United Kingdom

eLife Assessment

In this **fundamental** study, the authors argue that the most widely used behavioral measure of the ability to stop an action is largely determined by the time it takes for visual signals to reach the brain and for movements to be executed, rather than by the control processes it is assumed to index. The evidence is **convincing**, combining reanalyses of seven existing datasets with a preregistered new experiment, and its force comes from the authors doing three things at once: identifying the problem, showing that it is present wherever it can be measured, and offering remedies applicable to both archived and future data. Various concerns were raised as to the interpretation of the data and their robustness to experimental manipulations, but if these were properly addressed, the implications of the work could extend across psychology, psychiatry and neuroscience.

https://doi.org/10.7554/eLife.111819.1.sa4

Abstract

Response inhibition is seen as a key facet of executive control, investigated in thousands of studies across psychology, psychiatry and neuroscience. It is most frequently studied using the stop-task, and its main outcome measure, the stop-signal reaction time (SSRT), is overwhelmingly interpreted as indicating top-down inhibitory ability or speed. Here we provide converging evidence that undermines this central assumption, from reanalyses of archival datasets and a novel *preregistered* study. We find that low-level visuo-motor delays account for 2/3rds of manual SSRT duration and 40% of its individual differences, and predictably affect SSRT when manipulated. Sensory and motor contributions must be ruled out before linking SSRT results to inhibition or cognition. We evaluate a method for doing so, then introduce and evaluate a new measure, the selective stopping delay, that factors out visuomotor delays and captures the original aim of the SSRT to measure how fast response inhibition can be deployed.

Introduction

The ability to inhibit an ongoing action is thought to be impaired in nearly all psychopathologies, including schizophrenia[1], attention deficit hyperactivity disorder (ADHD)[1–3], obsessivecompulsive disorder (OCD)[4] and Parkinson's disease[5]. It is thought to be partly heritable[6] and has been linked to grey and white matter properties and integrity in the frontal lobe and activity in the basal ganglia[7,8]. All these conclusions strongly rely on one task, the stop signal task[9] (Fig. 1), which requires speeded responses to targets, and their inhibition in a minority of trials where a stop-signal follows target onset. From this task, a stop-signal reaction time (SSRT) is extracted and interpreted as reflecting a person's ability for top-down inhibitory control[9–11]. The SSRT has become the most widely used behavioural measure of response inhibition, permeating multiple research fields including cognitive, developmental, social and clinical Psychology, Psychiatry and cognitive Neuroscience in human and animal models (over 8500 publications[a]

Search for "stop-signal reaction time" on Dimensions.ai, completed on the 14th of April 2026

, including large multi-modal research programs, such as the multi-million US$ ABCD study following over 10,000 children going through adolescence[12]).

Here we argue that this corpus of work may have been misinterpreted because the SSRT is systematically influenced by low level sensory and motor transmission times, arguably more so than by inhibition or cognitive processes. These humble peripheral transmission times are not a constant that can be ignored in drawing conclusions about cognition. Sensory and motor speed and abilities vary across people and increase with old age. They are common side-effects of many drugs and are prevalent across many mental health and clinical conditions where SSRT increases have been systematically documented, including Parkinson's disease[13], ADHD[14], or psychosis[15].

SSRT's repeatability, interpretation and assumptions have been questioned before, attracting best practice recommendations and sophisticated refinements of the conceptual framework and modelling[16–21]. These refinements have preserved – and therefore reinforced – the message that SSRT, once appropriately estimated, essentially reflects response inhibition, or cognitive abilities at least. However, none of these adequately characterise or control for the overlooked effect of peripheral delays. Due to its low-level nature, this confound presents an existential problem for the large body of research and theory relating to cognitive inhibition.

We first establish this confound (Fig. 2), through reanalyses of 7 datasets[22–25], and a novel pre-registered study on 37 participants (all summarised in Table 1). We conclude that many SSRT differences reported in the literature and interpreted as differences in top-down inhibition are likely to be driven by peripheral delays that are neither top-down nor inhibitory. We then offer and evaluate two ways to correct for this problem. The first is simply to subtract a measure of visuomotor delays from the SSRT. This can be helpful but does not address another common misunderstanding: the SSRT, although conceived as a delay and expressed in msec, cannot be a direct measure of the *speed* of inhibition, since it also directly depends on how often participants try to stop[18]. This is an issue for studies using SSRT to define time-windows for neurophysiological analysis (see Discussion). We therefore propose an alternative measure, the selective stopping delay (Fig. 3), designed to directly and specifically estimate how fast selective inhibition *can be* deployed, irrespective of how often it is deployed or how fast on average, irrespective of incompressible sensory and motor delays, and without relying on simplifying modelling (and sometimes unvalidated) assumptions[26]. In other words, we introduce a measure aimed at delivering what the SSRT is often wrongly assumed to deliver. For this, we rely on a stimulus-selective stopping task, where participants are asked to stop to some signals while ignoring others. We again use archival data and the new preregistered study to provide estimates of selective stopping delay from manual and saccadic modalities (Fig. 4–5) and show it is immune to confounds affecting SSRT.

The present work relies on analyses of RT distributions. In the stop-task, distributions of response latencies in go trials (Fig. 1A) and failed stop-signal trials (Fig. 1C) overlap up to a divergence point (blue dot on Fig. 1D, see supplementary Fig. 1 for empirical distributions). This indicates that failed stop trials with an RT up to the blue dot escape all interference from the stop signal (Fig. 1E), while responses that would have been slower are susceptible to it, getting delayed or cancelled as a result (Fig. 1F). SSRT does not correspond to any visible landmarks in the distribution but is calculated by identifying the quantile on the go RT distribution that matches the participant's proportion of failed stops. The SSRT is relative to signal onset, and therefore obtained by subtracting the stimulus onset asynchrony (SOA) between the target and the signal. The first observable influence of the stop signal on the RT distribution, the blue dot, is also tightly locked to signal onset and occurs after a delay that we refer to as T_0.

From Fig. 1D, it seems clear that SSRT and T_0 are inevitably connected since, on good quality data at least, SSRT should always follow T_0. This matters because: 1) T_0 measures the visuomotor deadtime[22,27,28], i.e. the sum of incompressible visual and motor delays that constitute the minimum non-decision time in all visually-guided motor responses, 2) T_0 is long, varies

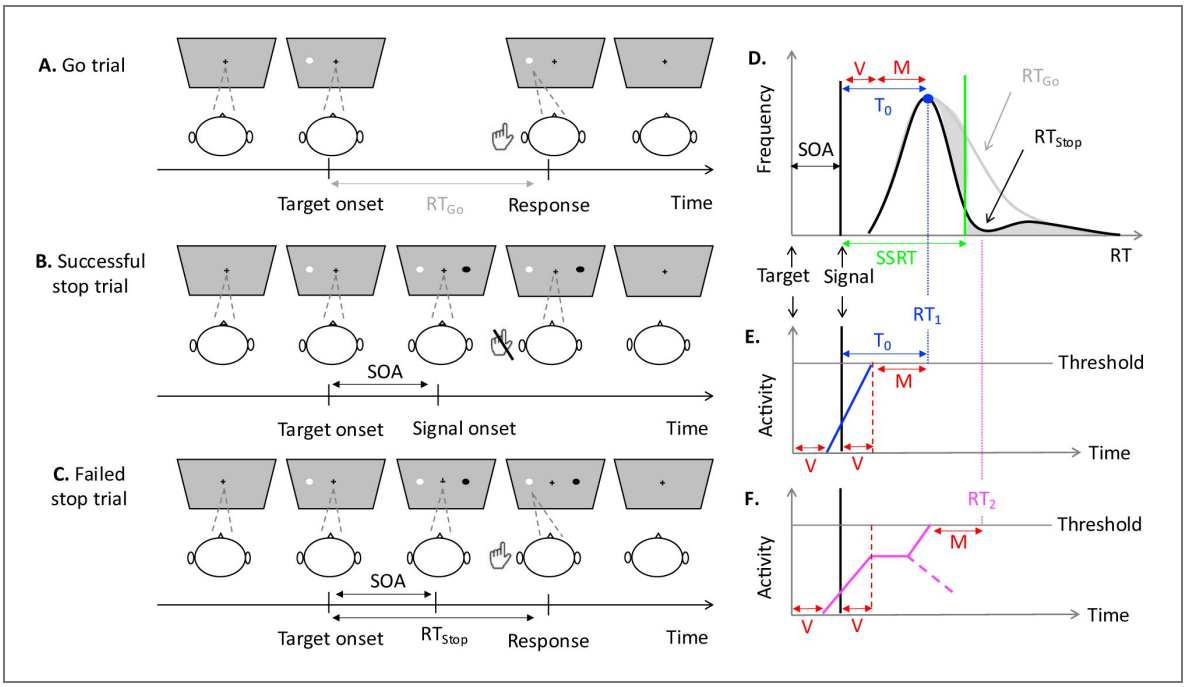


Fig. 1. Trial types in an example saccadic or manual standard stop-task, and RT-derived indices.

A. Participants are given the primary task of moving their eyes or pressing a button as fast as possible to indicate the side of a peripheral target (e.g. a white circle). **B-C.** On a minority of trials, a stop-signal (e.g. a black circle) is presented with a variable delay (stimulus onset asynchrony, SOA) after the target. Participants are instructed to withhold their response, which they sometimes do (B) and sometimes fail to (C). **D.** Distributions of reaction time on signal-absent trials (RT_{Go}, grey curve) and failed stop-trials (RT_{Stop}, black curve). The SSRT is the time delay between the signal onset and the vertical green bar (indicating the RT up to which the area under the RT_{Go} curve equals the area under the entire RT_{Stop} curve, i.e. the RT that equalises the two grey-shaded areas). Visuomotor deadtime (T_0) is the delay between the signal onset and the blue dot (where the RT_{Stop} distribution diverges from the RT_{Go}). In practice though, T_0 is estimated on RT locked on signal onset after pooling across all SOAs (see Methods). **E.** Hypothetical decision mechanism leading to an RT equal to (RT_1 blue line) or faster than SOA + T_0. The action decision activity triggered by the target has reached the threshold before the signal can interfere with it, i.e. before SOA + V (red dashed line, V = visual transmission delay for target and stop signals). These trials correspond to RTs ≤ SOA + V + M (M = motor execution delay), all unaffected by the signal. **F.** Trials where the activity is below the threshold by SOA + V are exposed to automatic inhibition from the signal, leading to responses getting delayed (RT_2, pink full line) or cancelled (pink dashed line). Note that the above logic is agnostic to the specific profile of the accumulation process or the interference.

systematically and often largely across participants, and predictably across empirical manipulations, 3) SSRT follows T_0 because it necessarily includes those same peripheral delays and therefore inevitably inherits their large within and between-participant variability.

That T_0 measures visuo-motor deadtime was initially established from the automatic interference of task-irrelevant visual distractors on saccadic and manual reaction times, supported by model simulations[27,28] (although automatic sensory-motor interference patterns are seen across many other contexts e.g. reading[29], free viewing[30], nystagmus[31], microsaccades and auditory stimuli[32]). Trials with a reaction time = SOA + T_0 (Fig. 1E) are those when the fastest signal-related information travelled from the retina to decision-related brain areas (visual deadtime) just when the decision to respond to the target has reached a point of no-return when it can no longer be inhibited or cancelled. Since the behavioural consequences of this interference (a dip in the RT distribution) only becomes visible after the additional motor-execution delay, T_0 (the onset of this dip) is equal to the sum of the fastest visual and motor delays i.e., the visuomotor deadtime. Consistent with this low-level nature, T_0 is predictably affected by visual and motor manipulations, and immune to cognitive manipulations[22]. Critically, it does not differ when interference is produced by stop-signals or task-irrelevant distractors (signals that are meant to be ignored)[25].

The realisation that the SSRT is necessarily and largely occupied by sensory and motor delays was first spelt out in Boucher et al. (2007b)[33], and further developed in Salinas & Stanford (2013)[34] and Bompas et al. (2020)[25]. However, this early modelling work was limited to saccades produced by a small number of monkeys and humans. Generalising this assumption, we predict that SSRT and T_0 should be clearly correlated across participants, and that SSRT should be sensitive to those same sensory and motor factors that affect T_0. Our account further predicts that experimental manipulations which lead to slower and more variable visual encoding (e.g. dim stimuli, vs bright) or motor execution (e.g. button press vs eye movement responses) will increase the variance in SSRT that is attributable to peripheral delays. The first section of the results confirms all these predictions, revealing just how much of the variance in SSRT is explained by T_0, and calling for a fundamental shift in how stop-task data ought to be analysed and interpreted. The second and third sections of the results offer solutions to this issue.

Results

SSRT strongly reflects visuomotor deadtime

Fig. 2 illustrates the strong relationship between SSRT and visuomotor deadtime during the stoptask. Fig. 2A shows the extremely high shared variance (85%) between measures when variance arises from a combination of empirical design and individual differences across 7 archival datasets. This relationship was preregistered and further confirmed for our novel data with 80% shared variance between SSRT and T_0 when pooling across manual and saccadic modalities (Fig. 2B, see supplementary Tables 1–2 and supplementary Fig. 2–3 for other outcome measures).

When estimating the shared variance driven solely by individual differences, the relationship is clearest in our manual dataset: T_0 (192 ms on average) occupies 2/3rd of SSRT (266 ms) and explains 40% of individual variance in SSRT. Critically, comparing our data to Hedge et al (2018)[16], manual SSRT turns out to be more highly correlated to visuomotor deadtime than to Go/No-go commission errors, another popular measure of motor response inhibition (27% shared variance), and any other executive control indices derived from the Flanker or Stroop tasks (all less than 2% shared variance). In saccades, the novel data shows a similar, yet weaker pattern: saccadic T_0 (60ms) occupies 1/3rd of SSRT (174ms) and explains 7% of the variance in SSRT. This reduced correlation is most likely due to extremely low individual differences in saccadic T_0 (all tightly packed between 40 and 70ms). We also preregistered, and confirmed, that SSRT should correlate with median RT from separate speeded blocks, since these inherit differences in visuomotor delays too (see supplementary Fig. 4).

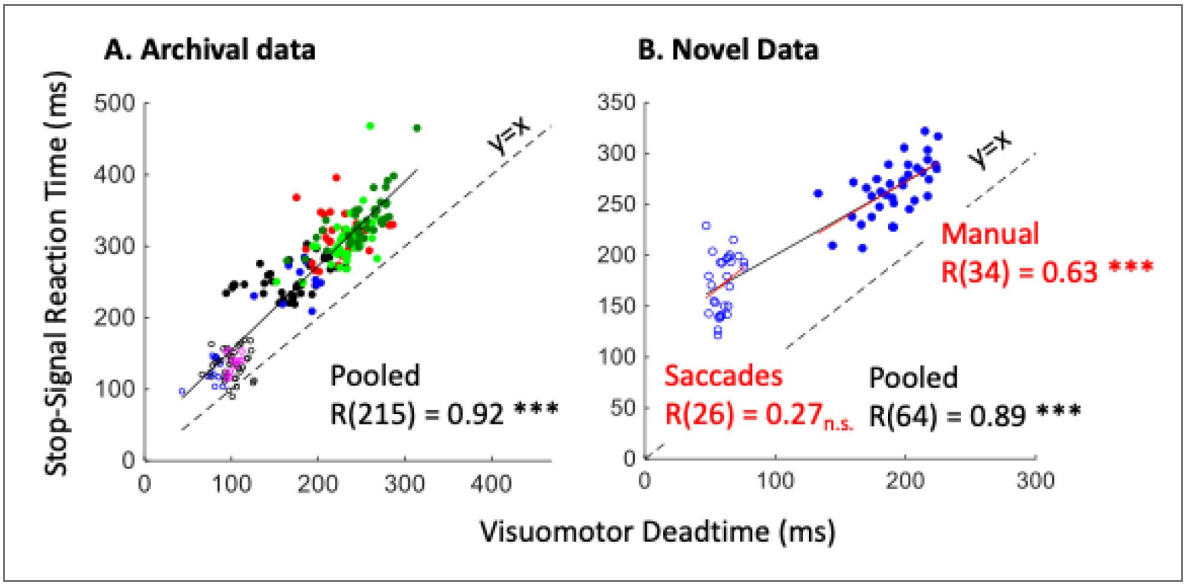


Fig. 2. Stop signal reaction time correlates with visuomotor deadtime in archival (A) and novel (B) datasets.

Data points correspond to individual estimates from manual (full circles) and saccadic (empty circles) blocks. The continuous lines show the linear regression, within (red) or across (black) datasets, the dashed lines show unity for reference. **A.** Pooled data across seven datasets: Campbell et al., (2017)[23] in black, Bompas et al., (2020)[25] experiments 1–3 in magenta, Boucher et al., (2007)[24] in blue, Bompas et al., (2025)[22] experiment 1 in green (bright and dark green for bright and dim signals) and experiment 2 in red. **B.** Data from the stop-blocks in the novel experiment. R-values are provided on each panel, with their degrees of freedom within brackets. *** indicates p-values < 0.0001, n.s. indicates a p-value > 0.05 and inconclusive BF10 (between 0.33 and 3).

As predicted, within-participant manipulations of visual contrast and execution modality that are known to affect T_0 also affected SSRT, and conditions associated with longer and more variable T_0 showed larger correlations between T_0 and SSRT. SSRT and T_0 from dim (yet still clearly visible) visual signals were longer and more strongly correlated than from bright signals (compare dark green and light green circles on Fig. 2A, differences > 18 ms, paired t-test $p < 10^{-6}$, dim r(38) = 0.75, bright r(38) = 0.54), consistent with a direct effect of visual delays on SSRT. Manual SSRT and T_0 were markedly longer and more strongly correlated than saccadic indices (compare full and empty circles on Fig. 2B: differences > 90 ms, paired t-test $p < 10^{-18}$), consistent with a direct effect of motor execution delays on SSRT, these delays being much longer and more variable across participants for button presses compared to saccades[24,28,35].

Correcting SSRT for visuomotor delays

On our novel dataset, regressing SSRT against T_0 across participants gave slopes of 0.75 for manual and 1.04 for saccades, consistent with a simple additive effect attenuated by measurement noise in the manual dataset. Therefore, it appears justified to subtract T_0 from SSRT to remove or reduce the low-level sensorimotor contribution and create a metric more strongly associated with cognitive processes. Testing for the effect of any manipulation on SSRT-T_0 should reveal its (over or under-additive) impact on response inhibition, beyond its direct effect on peripheral delays. For instance, in the exp. 1 from Bompas et al. (2025)[22], subtracting T_0 from SSRT reduced the difference between dim and bright signals from 29 to 11ms, which remained significant (z = 2.81, p = 0.005, n = 40). Therefore, although 2/3rd of the effect of signal contrast on SSRT was attributable to visuomotor deadtime, contrast also affected subsequent decisional processes, possibly including the speed, strength or likelihood of inhibition. The difference between manual and saccadic SSRT changed polarity when subtracting T_0 from SSRT in the novel dataset, from 92 to -37ms, but again remained significant (t(27) = 6.77, $p < 10^{-7}$), suggesting response inhibition is harder for saccades.

We encourage researchers to extract T_0 from their data and test how the additive contribution of visuomotor deadtime to SSRT might impact conclusions. Supplementary Fig. 5 illustrates the impact of trial numbers on T_0 estimates. Based on this, we recommend using at least 250 signal-present trials per T_0 estimates. This can be achieved individually from 1000 trials in total if signals appear on 25% of trials (about 30 minutes of data collection). Pending this, supplementary Fig. 6 further shows that pooling trials across 2 or more observers is a sensible way to achieve sufficient trial numbers to test for group or condition effects on T_0, with confidence intervals derived from randomised subsampling.

The selective stopping delay (ΔT)

Even corrected for visuomotor delays, the interpretation of SSRT remains problematic: although SSRT is often assumed to reflect a time delay (i.e. the time required for the brain to trigger topdown inhibitory commands), this is in fact misleading since SSRT is directly confounded by the probability to trigger these commands (itself not a delay)[18]. Low probability of stopping may just as likely indicate poor compliance due to low motivation, poor understanding or memory of the instructions, or poor ability to sustain cognitive demand over time. Therefore, SSRT may be *influenced* by the speed of reactive response inhibition, but it cannot offer an absolute measure of such delays. If SSRT is not a delay measure, SSRT-T_0 is not either and only addresses one limitation of SSRT (its dependency on visuomotor deadtime).

Below we explore a model-free, data-driven approach to estimating the *fastest* translation from exogenous to top-down inhibitory commands. In the standard stop-task, successful stopping may result from multiple influences, including proactive and automatic reactive processes[36,37]. In contrast, when to-be-ignored signals (henceforth 'ignore signals') and stop signals are randomly interleaved, as in the stimulus-selective stopping task (Fig. 3), their differential post-T_0 treatment provides direct evidence of reactive top-down processes. We reasoned (and preregistered) that the similarity of the initial interference between signal-stop and signal-ignore trials allows for the calculation of two new indices: T_S and ΔT (Fig. 3B). T_S denotes the selective stopping time: the point at which the behavioural effects of stop and ignore signals begin to

diverge, i.e., the earliest time point at which an individual is able to deploy instruction-related stopping signals. ΔT represents the selective stopping delay, T_S - T_0, which, we suggest, reflects the minimum time required at the decision level to transform a sensory input into the top-down command needed for selective action inhibition.

Fig. 4 presents the first characterisation of ΔT from selective stopping tasks in the manual and saccadic domains. For ΔT to serve as a valid measure of delay, two conditions had to be met, both were preregistered and confirmed. First, the initial interference was identical between stop and ignore signals, i.e. they produced indistinguishable T_0 (manual mean difference = 5.6ms, t(34) = 1.10, p = 0.28, BF_{01} = 3.2; saccadic mean difference = 1ms, t(31) = 0.86, p = 0.39, BF_{01} = 3.7), although only weakly correlated (manual: r(33) = 0.29, p = 0.05, BF_{10} = 1.04; saccades: r(30) = 0.39, p = 0.015, BF_{10} =3.22). This justifies the pooling of stop and ignore trials to get an overall, more robust, estimate of T_0. Second, ΔT consistently yielded positive values, consistent with it reflecting the additional time required to shift from exogenous (automatic) to endogenous (instruction-relevant) commands (manual T_S and T_0 differences: z = 4.39, p < 0.001, n = 36; saccadic: t(28) = 13.89, p < 0.001). On an individual level, T_S always exceeded T_0 for saccades, and on most cases for manual responses (except for 7 participants with small negative differences, see Fig. 5).

Saccadic ΔT were longer than manual ΔT (t(27) = 7.90, p < 0.0001), by 71ms on average, and consistently across almost all participants. Slower executive control for saccades (i.e. the opposite conclusion as the one suggested by SSRT) could contribute to explain the reduced stopping accuracy (36% on average) compared to manual responses (52%, t(36) = 6.49, p < 0.001).

Saccadic and manual ΔT were correlated, weakly but significantly (Fig. 5B, see Fig. 6C for all cross-modality correlation coefficients and Bayes Factors). Excluding participants with negative ΔT values or who complied poorly with the selective stopping instruction (i.e. whose stopping accuracy was less than 20% higher than their omission rate) only made our results marginally clearer but did not change our conclusions (cross-modality difference in ΔT: t(16) = - 7.4, p < 0.0001, $BF_{10} > 10^4$; cross-modality correlation in ΔT: r(15) = 0.54, p < 0.05, BF_{10} = 2.09).

The finding that individual differences in the speed of selective inhibition are partly shared across action modalities is consistent with the conclusion one would have drawn from SSRT (cross-modality correlation: r(27) = 0.65, p < 0.001). The stronger correlation observed with SSRT compared to ΔT may be driven by those factors that affect SSRT but not ΔT, such as visuomotor deadtime (weakly but positively correlated across modalities, consistent with shared visual deadtime) or stopping accuracy (strongly correlated across modalities, consistent with shared inhibition abilities, strategy or compliance).

This research started by highlighting two main limitations of the SSRT: its dependency on visuomotor delays and stopping accuracy. How does ΔT compare with SSRT on these fronts? Fig. 6 shows the coefficients and Bayes Factors for all cross-variable correlations in the novel dataset. Manual SSRT correlates with all the other variables of interest: T_0 and median RT, but also T_S and stopping accuracy (all p < 0.05). Manual ΔT obviously correlates with T_S, but does not correlate with the other variables, and also does not correlate with SSRT (all p > 0.1, supplementary Fig. 8A). Saccadic data show the same pattern, albeit less clearly (p = 0.055, BF_{10} = 0.98, see supplementary Fig. 8B).

Furthermore, ΔT also does not correlate with the difference in stopping accuracy and omission rate in ignore trials, an index of how differentially participants are responding to stop and ignore trials (manual: r(34) = 0.005, p = 0.98, BF_{10} = 0.13; saccadic: r(27) = -0.14, p = 0.46, BF_{10} = 0.19, including all participants before making any exclusions, see supplementary Fig. 8C). This confirms that how often a participant selectively stops does not influence ΔT, consistent with its intending aim to be a pure delay measure.

Discussion

Our results unambiguously demonstrate the dependency of SSRT on visuomotor delays, with implications for a large body of published work. Fortunately, the measure we introduce to quantify the issue (visuomotor deadtime) can be extracted on nearly all existing stop-task datasets

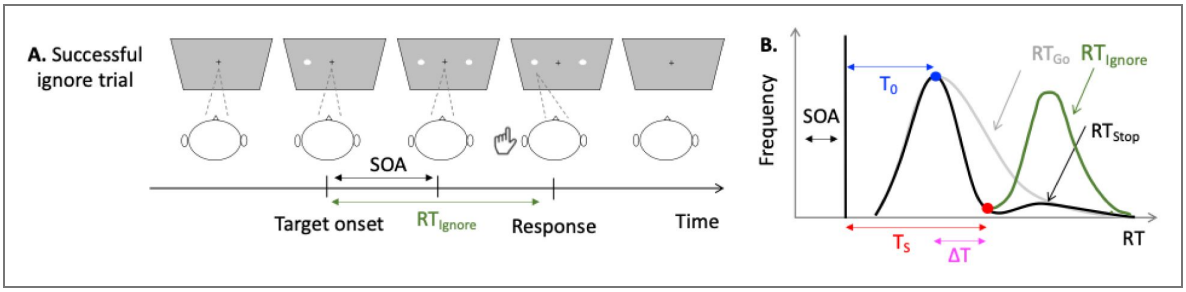


Fig. 3. Trial types in an example saccadic or manual ignore trial, and RT-derived indices.

A. Example ignore trial, interleaved with go and stop trials (see Fig. 1) as part of a selective stopping task. The participant is instructed to respond to the first stimulus and ignore the second one. **B.** Reaction time distribution on signal-ignore trials (RT_{Ignore}, dark green curve) shows the same initial decrease as failed signal-stop trials (both starting at the blue dot T_0) but shows a later rebound. The red dot indicates the RT at which the RT_{Stop} and RT_{Ignore} distributions diverge. T_S is this RT minus the SOA. In practice though, like T_0, T_S is estimated on RT locked on signal onset after pooling across all SOAs. The selective stopping delay ΔT is the difference between T_s and T_0. See supplementary Fig. 1 for empirical curves.

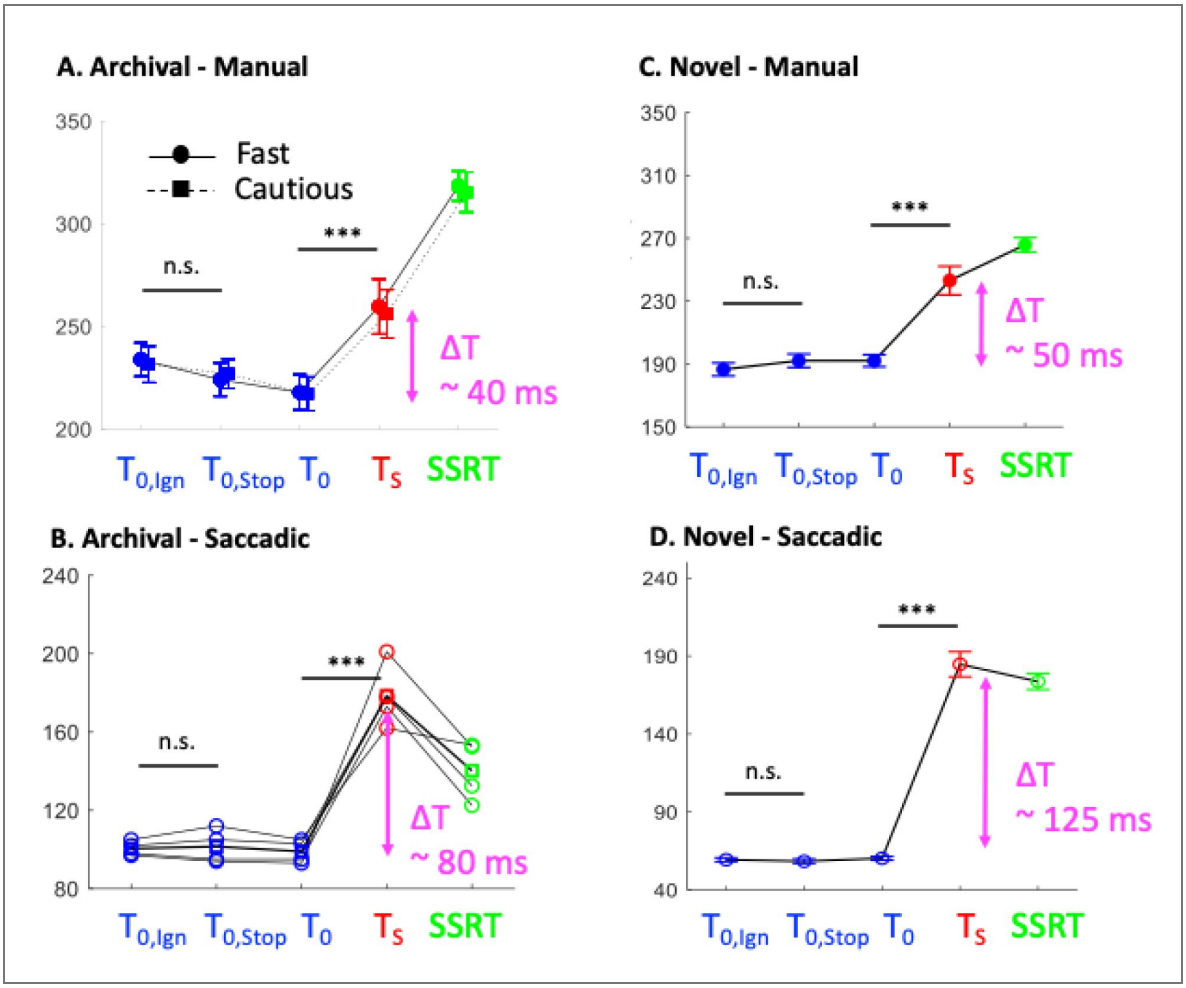


Fig. 4. Temporal indices from the selective stopping task, from two archival datasets (A-B) and the novel data (C-D).

A. Group averages and standard errors from 14 participants in Bompas et al. (2025)[22] experiment 2 (see supplementary information Fig. S7A for task description; fast and cautious blocks serve here as internal replication). **B.** Individual indices (empty circles and thin lines) and group averages (squares and bold line) over the 4 observers from Bompas et al. (2020)[25] experiment 3 (see supplementary information Fig. S7B for task description). **C-D.** Group averages and standard errors for manual and saccadic indices from our novel experiment. Visuomotor deadtime indices (blue) are when the RT_{Ignore} ($T_{0,Ign}$), RT_{Stop} ($T_{0,Stop}$) or pooled $RT_{Ignore+Stop}$ (T_0) distributions first drop below the RT_{Go} distribution. As pre-registered, they do not differ significantly (n.s.). T_S is when RT_{Stop} first drops below RT_{Ignore}, and, as pre-registered, is significantly higher than T_0 (*** $p < 0.001$). SSRT is the stop-signal reaction time calculated from Stop and Go trials.

Fig. 5. Individual indices from the novel experiment.

A. Same conventions as Fig. 2B in main text. **B.** The red line indicates where x=0. Values falling below this are non-plausible and, because they are small, probably due to noise in T_0 and/or T_S estimates (estimated independently).

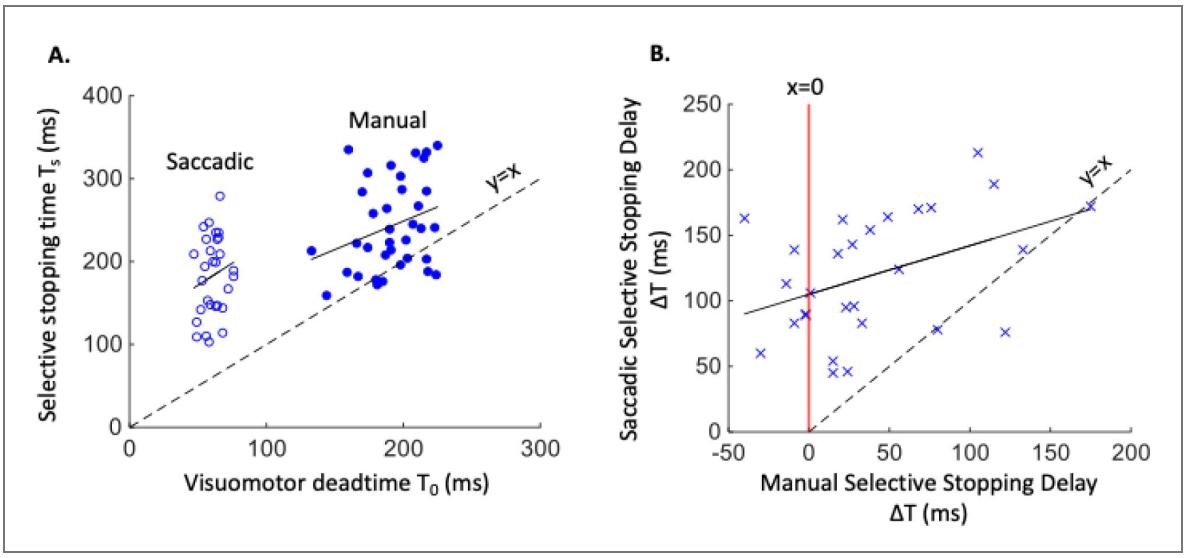
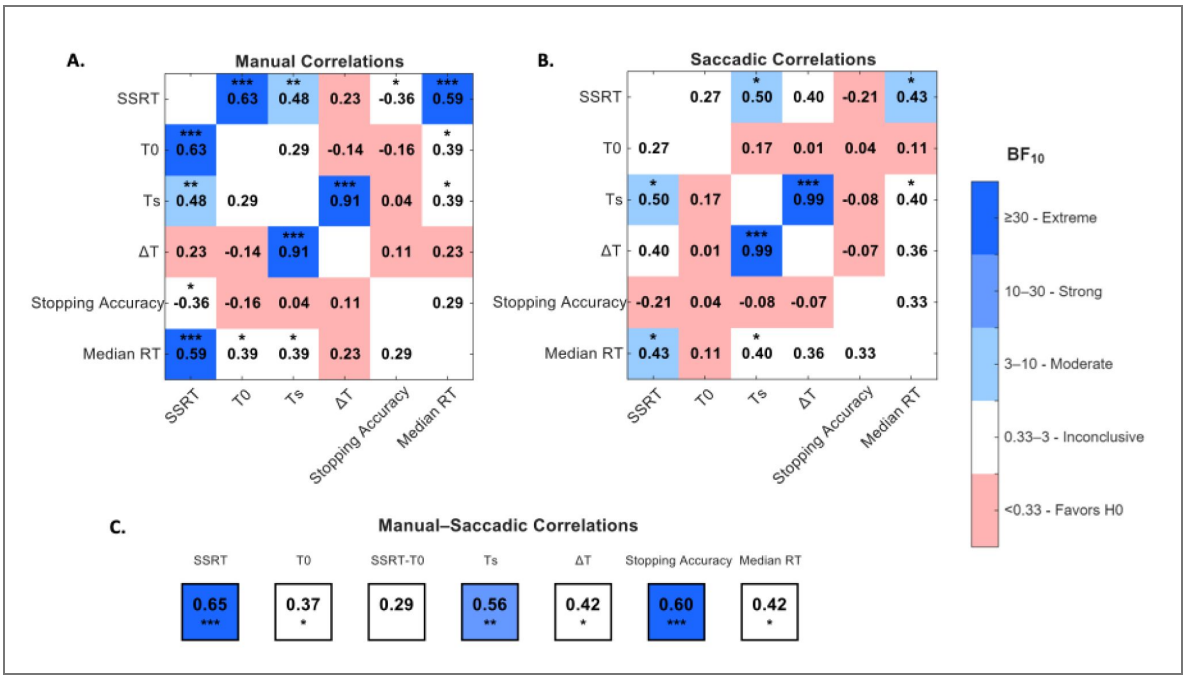


Fig. 6. Pearson's correlations for within-modality and cross-modality variable pairings.

Colour scale reflects the strength of Bayes evidence for the hypothesis against the null (BF_{10}). Asterisks reflect frequentist significance (p-values < 0.05*, < 0.01**, < 0.001***). All measures included are post participant exclusions. BF and p-values for correlations linking SSRT to T_0, and SSRT to median RT (A-B) are presented as one-tailed, consistent with the confirmatory and directional nature of the hypotheses (see Methods). All other correlations are 2-tailed.



to support a re-evaluation of previous conclusions (all analysis code is openly available, and applicable even on low trial numbers, see supplementary Fig. 5–6 and our offer of assistance in supplementary material). Presently, we do not think that viable empirical alternatives exist, including from simple decision models which ability to extract meaningful non-decision time parameters has been seriously questioned[22].

As an example amongst many where this idea could be applied, McPhee & Hendershot (2023)[38] gathered 26 studies investigating the effect of alcohol on SSRT, to conclude that "those participants who received alcohol demonstrated larger SSRTs (i.e., poorer inhibitory control) compared to those participants who received either a placebo or control" (p.10). This includes the Campbell et al. (2017)[23] study, where our reanalysis of the data suggests a post-alcohol increase in visuomotor deadtime compared to the sober blocks (+13*, 7, 52* and 25* ms for the saccadic-ignore, saccadic-stop, manual-ignore and manual-stop, * indicates $p < 0.05$). T_0 increases were commensurate with reported SSRT increases (Campbell's Fig. 4, small and insignificant for saccades, large and significant for manual), consistent with these effects being fully driven by visuomotor delays and unrelated to inhibition. If the same pattern is confirmed across the other datasets, this would overturn what has been considered until now a well-established conclusion.

Similarly, previous findings relating to clinical populations may need reconsideration. ADHD has long been characterized by impaired inhibition, evidenced via longer SSRTs. The original meta-analysis by Lipszyc and Schachar (2010) included 68 studies, where all but one study found longer SSRT values for those with ADHD compared to control groups[1]. Senkowski et al's., (2023)[3] meta-analysis investigated an additional 19 papers, and again, all but one showed prolonged SSRTs for the ADHD groups. Regardless, if peripheral visuomotor delays partly explain these differences in 'inhibitory control', this substantial body of literature may need reinterpreting. Accounting for such delays could reveal broader deficits, and thus, call for the reevaluation of how we profile, assess, and intervene with neuropsychological disorders like ADHD.

Relevant to the vast majority of human research using manual responses, our results show SSRT's dependency on motor delays is pervasive. This may be counterintuitive because SSRT depends on the number of withheld actions which are, by definition, not executed. However, SSRT can only be estimated from *observable* responses. As a result, any factors that increases the time it takes for an action to be executed or registered (e.g. a slow peripheral device, muscle weakness, or motor pathway damage) will directly increase RT_{Go} and therefore SSRT, even if these factors have no impacts on executive control *per se*.

The confounding effect of visual delays on SSRT is more intuitive, yet still remained to be quantified and its implications fully spelled out. Visual salience has been shown to impact both neuronal latency and SSRT[39,40], but these effects have been interpreted as evidence that salience influences inhibitory control, via its effect on signal detection, seen as an integral precursor to the inhibitory process. This aligns with the recommendation by Verbruggen et al. (2019)[20] to use salient stop-signals, to reduce the risk that the signal is missed, and avoid within or between-group differences in SSRT being confounded by differences in "perceptual processes" (see also Verbruggen et al., 2014[41]). Although our findings align with previous observations, we highlight that confounds from visual delays can be general (rather than specific to stop-signal processing within the context of inhibitory control) and very low-level (visual deadtime takes place before perceptual processing). For studies exclusively concerned by individual differences within a healthy adult population with normal vision, tested with the highest-standard laboratory equipment, design and analysis, our results suggest that confounds from visual delays are minimal when using salient stimuli, since individual differences in saccadic T_0 (the best proxy for visual deadtime due to minimal contribution of motor delays[22]) are small, and so is its shared variance with saccadic SSRT. For any other research, the potential confound from visual delays should be carefully considered. This conclusion most likely generalises to all other sensory modalities.

Despite the lack of consideration paid to peripheral delays[22], experts have long acknowledged the need to rely on additional control measures to support SSRT-related interpretations. However, all come with their own issues. For instance, a lack of difference in RT_{Go} does not provide sufficient reassurance that a difference in SSRT is about inhibition (e.g. our dim versus bright signal

comparison affects SSRT and leaves RT_{Go} unchanged, yet is unrelated to inhibition). Moreover, while visuomotor differences will undoubtedly impact RTs, so can proactive slowing, itself closely coupled to accuracy, making changes (or lack thereof) in RT_{Go} hard to interpret. Other control measures include mean RT in tasks that do not require stopping, such as speeded RT tasks (without signals or with only ignore signals) or the dual-response task (where signals attract an extra motor response, instead of inhibition, e.g.,[42]). Because they are less sensitive to strategic influences, RTs from these tasks constitute better control measures for low-level confounds. However, they require additional testing and, for the dual-response task, attracts response execution interference, adding extra complexity[43]. Moreover, they still reflect the sum of decision and non-decision time, and can therefore remain hard to interpret. In contrast, T_0 can be estimated on all existing datasets, and specifically controls for peripheral confounds.

For research that requires specific timing estimates, we encourage the use of the selective stopping task. This is particularly relevant for neuroimaging studies searching for the neural basis of reactive inhibitory control in high-temporal-resolution data such as single cell recordings, electroencephalography (EEG), magnetoencephalography (MEG) or intracranial recordings data. Many papers use SSRT to define time windows of interest, or correlate the latency of neural markers with SSRT to support claims that these markers are associated with the speed of inhibition (e.g.,[44–50]). Since SSRT is not a measure of delay, defining time windows around it is conceptually flawed. Moreover, since SSRT necessarily contains large motor delays, neural mechanisms of reactive stopping must unfold several tens of milliseconds prior to SSRT[22]. This is typically not the case for the P300, the most commonly assumed marker of inhibition, often found to occur *after* SSRT[46,50,51] or synchronously[45]. T_0, TS and estimates of motor output time can together allow for the definition of more principled temporal anchors for isolating neural markers of inhibition (see the preregistered MEG study by Statham et al., 2025[52]).

One limitation of the stop-task (standard of selective) is its reliance on participants' compliance with the dual-task expectations between going and stopping. We found that even young healthy controls often get excluded from SSRT analyses (in line with recommendations[20]) because they stop too little or too much. In contrast to SSRT, the selective stopping delay is designed to reflect how fast selective inhibition *can* be deployed, rather than how often it happens, and indeed is not affected by response probability on stop or ignore trials. It can therefore be estimated on all participants irrespective of task performance, which makes it more inclusive. Still, since dip extraction is data driven and reliant on sufficient trial numbers, the more often participants deploy selective inhibition (i.e., respond differently to stop and ignore trials), the closer estimates will be to true (i.e. fastest possible) values. To improve participants' compliance, future research could implement trial-by-trial feedback, such as time penalty following incorrect responses (e.g. stopping on an ignore trial or responding to stop trials with an SOA of zero).

Another limitation of both the standard stop and selective stopping task is the need to collect a high number of trials to get robust estimates (although a large proportion of SSRT studies arguably rely on too few trials to guaranty decent repeatability[16,17]). This need for large trial numbers is exacerbated in a selective stopping design (because it has one more condition) and by the distributional analyses our work relies on. Supplementary materials Fig. 5–6 directly address this concern and illustrates how pooling can be used to mitigate the consequences of low samples.

Our results show that the selective stopping delay behaves as expected, producing mostly positive values and free from the confounds that affect SSRT. However, truly validating this new measure will require future work, for instance to identify cognitive factors that affect TS while leaving T_0 unchanged. This could include manipulating the instruction or reward structure to encourage caution versus speed, or increasing the complexity of the stimulus-response mapping to make the task cognitively harder. Importantly, ΔT includes delays related to perceptual decision making (i.e. discriminating stop from ignore signals) and is therefore not a pure measure of topdown *stopping* delays. This does not appear as a weakness to us, as surely the ability to stop selectively to some signals (e.g., red traffic light) but not others (green traffic light) is an essential part of any ecologically valid measure of stopping. An important follow-up question would be to what extent TS is specific to selective stopping or also predicts equivalent temporal landmarks in the dual-task

variant (that requires an additional response but no stopping) or in the stop-change task (that requires stopping and action updating). While these variants have been directly compared to the stop task before[43], data analysis never involved distributional analyses enabling that level of understanding.

Last, our results invite an update of the DINASAUR model[25,27,28] with a new endogenous stopping delay parameter. Our previous modelling assumed a single delay between exogenous signals and endogenous signals for going and stopping, fixed at 25ms for saccades[25] and 0 for manual responses[28]. Previous modelling by Logan et al. (2015)[53] suggests 16 and 42ms for translating fixation signals into stopping commands for saccades in two monkeys ($D_{control}$ — D_{fix} in Blocked Input 2.0). Our results show the delay to discriminate between two visual signals and trigger a stopping command could vary all the way from very short to 200ms across participants and action modalities. The refinement of biologically inspired generative models of behaviour holds the key to further our understanding of the highly non-linear mechanisms leading to complex behavioural signatures like those observed in the selective stop task.

Higher level cognitive abilities have evolved from basic sensory and motor functions, and still heavily rely on them, as do all the cognitive tasks and tests developed by researchers and clinicians to evaluate these abilities. When changes in basic visual or motor functions are well documented, parsimony should lead us to hypothesise that they drive changes in SSRT, until proven otherwise. This is likely the case with many mental health challenges, clinical conditions, age, drugs, general intelligence, brain structure or genes. It may take many years to revisit the literature, but this exercise may reveal rich insights.

Methods

Table 1 summarizes the main features of all the datasets included in this article, and group averages of their main outcome measures. The design for archival datasets was covered in Bompas et al. (2025)[22]. All new analyses of archival data were conducted in the same way as for the novel data. This methods section focuses on the novel dataset, preregistered on the Open Science Framework (*https://osf.io/2yu7k*). The preregistration outlined the study design, desired sample size, hypotheses, and planned analyses. All data and code are available in *this OSF folder*.

Participants

Data was collected from 37 participants, all of whom were recruited through volunteer sampling. This sample size is similar to previous experiments reported in Bompas et al. (2025). Eligibility criteria were to be between 18 and 50 years old, have normal or corrected to normal vision, and no motor impairment. Participants were offered compensation for their participation either by course credits or financial payment of £10 per hour. Ethical approval was granted by Cardiff University's School of Psychology Ethics committee. No personal information, including age and sex, was recorded. All participants took part in all conditions.

Equipment and stimuli

The selective stopping task was coded using MATLAB (version 2018b, The MathWorks, Inc.) with Psychtoolbox (version 3). Stimuli were presented on an Asus VG248QE (1920p by 1080p) monitor with the refresh rate set to 100 Hz. Manual button press responses were recorded using the side buttons on Nata Technologies LXPAD-2×5-10M 5-button, 2-hand system button boxes. Participants used their left and right thumbs. Eye movements were tracked using an EyeLink 1000 Plus eye-tracker (SR Research) with a sampling rate of 1000 Hz for each eye (binocular mode). Targets and signals were full contrast white and black dots appearing at 6.4° eccentricity from a black central fixation cross on a middle grey background. Participants viewed the stimuli from a combined chin and headrest mount which was set 60 cm away from the screen.

Publication	Exp	N	Go-signal	Stop-signal	Ignore-signal	Action modality	Instruction	$T_{0,Stop}$	SSRT	Median RT	$T_{0,Ignore}$	T_0	T_S	ΔT
Boucher et al. (2007)	1	7	Peripheral small black square	Central small brightly coloured square		Saccades	Unspecified	79	122					
						Saccades bimodal		77	126					
						Joystick		186	238					
						Joystick bimodal		172	264					
Campbell et al. (2017)		40	Peripheral small white circle	Central small black circle	Central small black circle	Saccades	Neutral	99	134	161	107			
						Button presses		170	255	306	195			
Bompas et al. (2020)	1	4	Peripheral small white circle	Central small white circle		Saccades	Neutral	104	140		97			
	2	4						101	128	155	91			
	3	4		Central small white or black circle	Central small black or white circle			102	140		101	99	179	80
Bompas et al. (2024)	1	40	Peripheral small black square	Central bright ~3 deg rectangle		Button presses	Speed	233	308	421				
				Central dim ~3 deg rectangle				251	337	436				
	2	14	Peri-central black letters	Peri-central, different letter from go-signal	Peri-central, same letter as go-signal	Button presses	Speed	224	318		234	218	260	42
							Accuracy	227	315		232	217	256	39
Novel		37	Peripheral small black or white	Peripheral, different polarity from go-signal	Peripheral, same polarity as go-signal	Saccades	Speed	59	174	182	58	60	185	122
						Button presses		187	266	314	192	192	243	50

Table 1. Main features of each dataset analysed in the current article and their mean values in msec for all indices of interest.

Grey cells indicate non-applicable values. $T_{0,Stop}$, $T_{0,Ignore}$ and T_0 indicate visuomotor deadtime estimated from stop trials only, ignore trials only, or all signal trials. T_S is the selective stopping time. ΔT is the selective stopping delay, i.e. the time difference between T_0 and T_S. Median RT is estimated from separate blocks without stop signals. Novel refers to the preregistered dataset collected in the article, for which full details can be found in the Methods section.

Procedure

Each trial started with a central fixation cross (Fig. 7). A second later, a target dot appeared either left or right of the cross. In go-trials, no other stimuli were presented. Speeded blocks were made up purely of these go-trials. In stop/ignore trials, a second dot appeared on the opposite side of the cross to the target. The SOA between target and signal onset were 80, 100, 120, 140 and 160ms, presented in equal proportion. In all trials, the fixation cross was removed 500ms after target onset, signaling to participants the right moment to blink. This departed from the preregistered methodology which did not specify a fixation cross offset period, in order to minimize trial loss due to blinking around target onset. Targets/signals remained on screen until the end of the trial, which was 1 second after target onset. Half of the blocks had a white target and ignore signal with a black stop signal, and vice versa for the remaining blocks.

For each modality, a total of 1000 go-trials, 500 ignore trials, 500 stop trials, and 400 speeded go-trials were collected. Data was collected in two sessions, each split into 12 blocks and lasting two hours. The blocks alternated between manual and saccadic responses. Only target position within a block was randomized across trials using MATLAB's built-in default random number generator. Block order for each session was identical across participants. All four speeded blocks (one for each target polarity and response modality) were presented in session 2.

If eye tracking recalibration was needed, this was done so at the beginning of each block. For manual blocks, participants were instructed to keep their eyes on the central fixation cross for the duration of the trial. At target onset, they were instructed to rapidly press the button corresponding to the target side. In saccadic blocks, participants were told to keep their eyes on the cross until the target appeared, and to rapidly move their gaze to the target then back to the centre before the start of the next trial. In speeded blocks, participants were informed that no other stimuli would appear. In selective stopping blocks, participants were told that on some trials, a second dot could appear on the opposite side shortly after target onset. If the second dot was the same colour as the target, they should ignore it; if it was not the same colour, they should try to withhold their response. It was heavily emphasized that participants should prioritize speed over accuracy, and expect to fail on about half of the stop trials (as recommended in [54]). Feedback was given after each manual block as to whether they were doing well or should stop more or stop less (if their stopping accuracy was below 15% or above 85%, respectively).

Analysis

Saccades were detected using standard SR research Eyelink parser settings which included a velocity criterion of 30°/s and an acceleration criterion of 8,000°/s[2]. We set fixations to merge with neighbouring ones if their duration was less than 250ms and their separation was less than 1°.

No effect of stimulus contrast polarity (black or white) was observed in median RTs for manual responses in the speeded blocks ($t(34) = 1.41$, $p = 0.17$, $BF_{10} = 0.45$). A difference was found for saccadic responses with responses to white targets being 10ms faster than to black ($z = 2.33$, $p = 0.02$, $BF_{10} = 3.26$). Since the difference was small (comparable to our bin size, and therefore our temporal resolution), it was still deemed appropriate to pool data prior to extracting the main indices from the selective stop blocks.

For consistency across archival datasets (mix of standard and selective tasks), visuomotor deadtimes presented Fig. 2A are all $T_{0,Stop}$, i.e. they are estimated from go and stop trials only. T_0,Stop in the new data were extracted using the same procedure as in Bompas et al., (2025)[22]. RTs from unsuccessful stop trials with a correct response direction (towards the target) were aligned to the stop-signal onset and pooled across all SOAs. These were binned using a bin size of 10ms and then were smoothed with a Gaussian kernel (n = 25 ms, σ = 5 ms) to yield 1ms resolution. For comparison, a surrogate go distribution was created by subjecting go trial RTs to the same alignment, binning, and smoothing steps. The point of maximum interference (T_{Max}) was defined as the peak of the distraction ratio ((N_{go}–N_{stop})/N_{go}), which had to be at least 20% for a dip to be deemed present. Dip onset time ($T_{0,Stop}$) was then identified by moving backward from T_{Max} until the ratio dropped below 2%, or when (N_{go}–N_{stop}) fell below 1. $T_{0,Ignore}$ was extracted in the same

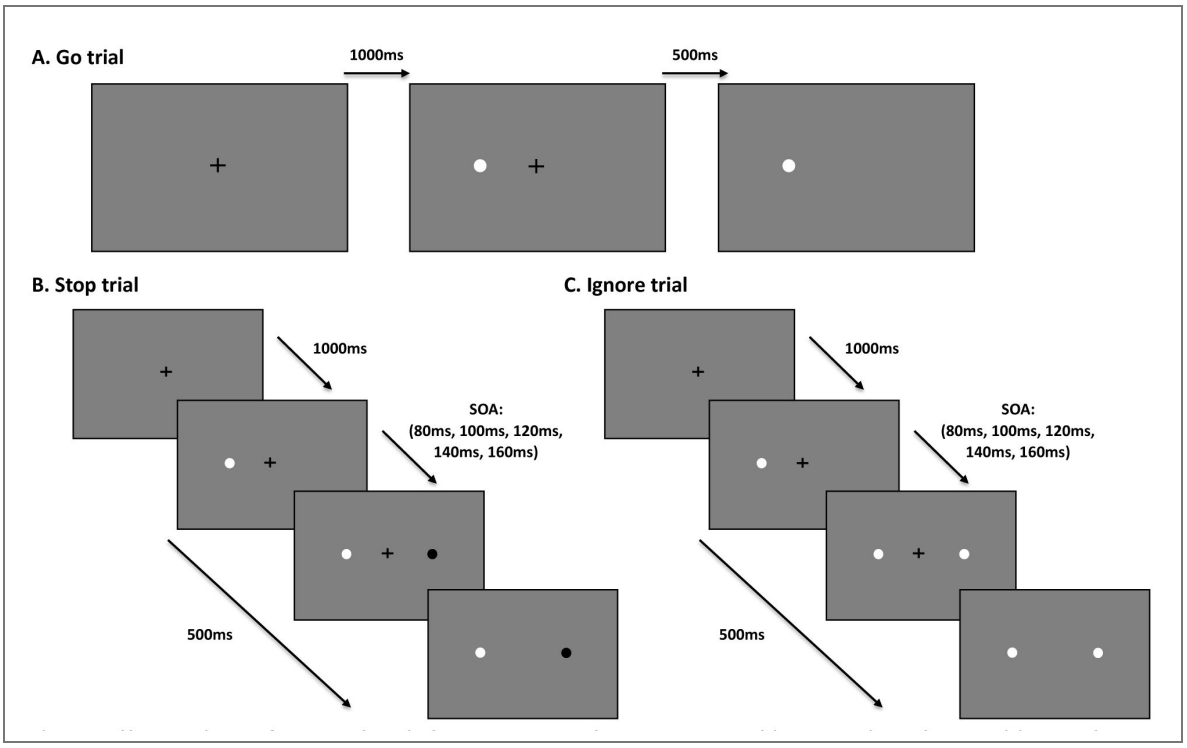


Fig. 7. Illustration of example trial sequences when target and ignore signal are white and stop signal is black.

A. Go trial. **B.** Stop trial. **C.** Ignore trial. All trial types have a total length of 2000 msec (trial end is 500 msec after the fixation cross is removed).

way but by comparing the go and ignore distributions. T_0 was extracted in the same way but comparing the go and pooled ignore and stop distributions. Except when directly comparing $T_{0,Ignore}$ and $T_{0,Stop}$, all mention of visuomotor deadtimes in the novel datasets refer to T_0. T_S was extracted in the same way by comparing the stop and ignore distributions. Departure points were visually inspected and adjusted if it was deemed that the algorithm had placed them in inappropriate places e.g. a lot later than when the distributions visually begin to diverge. They could also be excluded if the dip was difficult to distinguish from noise, or its onset was too ambiguous. ΔT was then calculated as $T_S - T_0$.

SSRT was calculated in a way to produce outcomes identical to Verbruggen's[20] recommended integration method. For each subject, modality, and SOA, a scaling factor is computed (total number of valid go trials / total number of valid stop-signal trials in that SOA). Valid trials are those that survived our eye-tracking related exclusions (see below), irrespective of whether a response was produced or not. We then sort the go trial RT and find the RT corresponding to the probability of responding on stop trials (adjusted by the scaling factor) and subtract the SOA. An overall SSRT for each subject and modality is then calculated as a weighted average across SOAs, weighted by the number of valid stop-signal trials in each SOA.

Hypotheses and statistical tests

The following hypotheses for the novel dataset were all inspired by results on the archival datasets, as outlined in the preregistration.

Confirmatory hypotheses:

- SSRT will correlate with T_0 across participants
- SSRT will correlate with median RT from speeded blocks
- $T_{0,Stop}$ and $T_{0,Ignore}$ will correlate
- $T_{0,Stop}$ and $T_{0,Ignore}$ will not differ
- T_S will be longer than T_0, and therefore ΔT will be positive overall.

Although the pre-registration of the top three hypotheses did not explicitly state "positively" nor mentioned if tests would be 1 or 2-tailed, this was an oversight, since these follow-up from positive correlations observed on archival data (Fig. 2A, and supplementary Fig. 4A, and Fig. 8B-C in Bompas et al. (2025)). Accordingly, we provided one-tailed p-values and Bayes Factors for these. One-tailed tests were also used or the $T_S > T_0$ hypothesis, pre-registed as directional. Two-tailed tests were used for the $T_{0,Stop} = T_{0,Ignore}$ hypothesis.

Exploratory and Non-preregistered tests:

All exploratory and non-pre-registered tests were reported 2-tailed. This includes all analyses on archival data. On the novel dataset, we pre-registered one exploratory analysis: the cross-modal comparison of ΔT.

All correlations were assessed using Pearson's coefficient, whilst tests of differences were assessed using paired t-tests. Pearson correlation coefficients and paired-sample t-statistics were computed using standard MATLAB functions ('corr' and 'ttest'). Corresponding Bayes factors were extracted using the 'bayesFactor' MATLAB package[55]. Bayes factors were computed using the Jeffreys-Zellner-Siow (JZS) approach with a scale of 1, assuming equal prior probabilities for the null and alternative hypotheses. The computation was based on the observed test statistic (t or r) and sample size (n), numerically integrating over the variance parameter *g*, which follows an inverse-gamma prior[56]. No Markov chain Monte Carlo sampling was used. We predominantly report BF_{10} which indicates the ratio of the likelihood of the observed data under each hypothesis compared to the null hypothesis. Where BF_{01} was reported, it signified 1/ BF_{10}. One-tailed BF_{10} values were estimated using $BF_{10,two\text{-}tailed} \times 2$. Normality was tested using MATLAB functions to perform Shapiro-Wilk parametric hypothesis tests. Non-parametric tests of differences, i.e. Wilcoxon signed ranks tests, were used when variables being tested violated normality. This was only the case for manual T_S in the novel dataset and, in the archival analyses, SSRT and corrected SSRT values (SSRT-T_0) for both contrasts from the Bompas et al., (2025, Exp 1) data, and the manual SSRT, speeded median RT, and corrected SSRT values from the Campbell et al., (2017) (all $p < 0.05$).

Trial exclusion criteria

Trial exclusion followed the pre-registered pipeline, except for two departures mentioned below. All reported results using the novel dataset use the adapted preprocessing criteria. Criteria for trial exclusions were 1) poor fixation around target onset, 2) unclear saccadic behaviour, 3) saccades taking place during manual blocks or 4) manual RT < 120 ms (anticipatory), or > 700ms.

Criterion 1 was assessed between -250 and +70ms around target onset and fixation was classified as poor if blinks or saccades of more than 1 degree occurred. As a result, saccadic reaction times shorter than 70ms (likely to be anticipatory) were excluded.

Criterion 2 was assessed between +70ms and +700ms after target onset and was met if the first saccade detected was larger than 1 degree but smaller than half the target eccentricity (i.e. a saccade was performed but not clearly directed to the visual stimulus). As such, the maximum saccadic RT was 700ms. This deviated from our preregistered criterion which stated 500 ms cutoff, because many participants produced saccades with RTs > 500 ms, particularly on saccadic ignore trials (12 participants were originally lost to omission rates > 20% with the 500 ms cutoff).

Criterion 3 was assessed within manual blocks only and trials were excluded if any saccade larger than 1 degree was detected within –250 to +70ms around target onset. This again departed from the pre-registration which originally specified a detection window of –250ms to +700ms. Visual inspection of the data showed participants to be making blinks shortly after making their manual response on many trials, and these got excluded by the original detection window. The shortened window allowed participants to blink or spontaneously look at the target throughout the duration of manual trial, other than in the time around target onset.

Participant exclusion criteria

Supplementary Table 2 presents number of participants contributing to each index after applying all exclusions. Participants were excluded from modality specific analyses if their primary task accuracy (left-right discrimination of target) for that modality was less than 80% (only 1 participant lost for saccadic speeded median RT calculation based on low saccadic primary accuracy on speeded blocks).

Participants were excluded from SSRT analyses in either modality if they failed to stop on more than 80% or less than 20% of stop trials across all SOAs (manual: 1 excluded, saccadic: 8 excluded). This was due to SSRT estimations being less reliable the more the probability of responding on a stop-signal trial deviates from 50%[20]. All but one of those excluded in the saccadic data for this criterion showed manual stopping accuracies within range, which would suggest the low saccadic stopping accuracies are more likely due to difficulties inhibiting saccades specifically, rather than a lack of instruction understanding. We did not exclude participants from SSRT calculations if their mean RT_{Stop} was equal or longer than their mean RT_{Go} (which only happened for 4 participants in the novel saccadic dataset, see supplementary Fig. 2A). From our perspective, the assumption of contextual independence between go and stop commands is not theoretically warranted[19,57]. T_0's logic explicitly relies on the competition between alternative response options, and mean RT do not adequately capture the non-normal, often multimodal, RT distributions resulting from the highly non-linear dynamics of the selective stop task.

Participants stopping on less than 20% of stop trials or more than 20% of signal-ignore trials across all SOAs in either modality were excluded from T_S (and therefore ΔT) analyses in this modality (manual: 0 excluded, saccadic: 1 excluded). Lastly, participants with more than half of any trial type excluded or missing were excluded from analyses involving this trial type (only 1 participant was excluded from the manual speeded condition, predominantly due to the participant making ambiguous eye movements (criterion 2 below)).

Participants were excluded from T_0 (and therefore ΔT or SSRT-T_0) analyses for a modality if their T_0 estimates were shorter than 35ms for saccades and 95ms for manual responses (i.e. 40 and 100 ms minus the 5 ms smoothing window), considered biologically implausible (following 22,28).

Supplementary information

Table 1. Additional outcome measures from the novel dataset.

Group average and standard deviation within brackets for each trial type for the novel data, before participant exclusions. Go, Ignore and Stop trials correspond to the three trial types during the selective stop task. Speeded Go are from separate blocks with no signals. Primary accuracy: percentage of all valid responses that were directed to the target (for stop trials, this applies only to failed stop trials).

Trial Type	Primary Accuracy (%)		Stopping Rate (%)		Mean RTs (ms)	
	Manual	Saccadic	Manual	Saccadic	Manual	Saccadic
Go	98.4 (2.1)	97.3 (3.8)	1.3 (2.0)	2.4 (3.5)	408 (33)	283 (49)
Ignore	94.0 (4.3)	94.9 (3.6)	2.8 (3.2)	4.4 (4.3)	443 (43)	323 (68)
Stop	99.1 (1.0)	96.1 (6.0)	52.2 (14.7)	36.4 (17.7)	374 (29)	264 (42)
Speeded Go	98.4 (1.2)	97.6 (6.2)	0.74 (1.0)	2.1 (6.1)	326 (33)	188 (27)

Table 2. Number of participants contributing to indices of interest in the novel data after all exclusions.

Measure	N	
	Manual	Saccadic
T_0	37	36
T_s	36	30
SSRT	36	29
ΔT	36	29

Assistance with extracting T_0

We welcome requests for assistance with extracting T_0 values from stop-task data shared with us (*bompasae@cardiff.ac.uk*) if they are formatted according to the present guidance and accompanied by a file describing how the data were obtained and pre-processed, with a link to the article if published. Data should include a list of all trials (1 row per delivered trial), with key descriptors as columns or fields, including:

- participant number or code
- delay between target and signal (in ms, use NaN or leave empty on Go trials)
- reaction time from target onset (in ms, use NaN or leave empty if no response recorded)
- accuracy of the primary response (whether response matches target requirement, not stopping accuracy)
- any other conditions leading to distinct T_0 estimates

If the task involves multiple response (eg two button presses, or a button press and an eyemovement), separate RT and accuracy descriptors should be added. Ground for trial rejection (e.g. poor fixation) should be added as a separate descriptor. Ideally, all delivered trials should be included in the list, irrespective of their accuracy or whether they have been rejected. Pending this, the number of trials delivered within each condition needs to be provided. This is essential to scale the signal and go reaction time distributions so numbers are directly comparable. Accuracy and rejection should be assessed in the exact same way across conditions (most importantly irrespective of whether a trial contains a signal or not).

Fig. 1. Reaction time distributions pooled across all participants.

Each index is estimated from RT locked on signal onset, after pooling across the 5 SOAs and 37 participants, separately for the manual (left) and saccadic blocks (right). Same conventions as Fig. 1D and 3B from main paper. Index calculation uses the same algorithm as individual estimates, but without Gaussian smoothing.

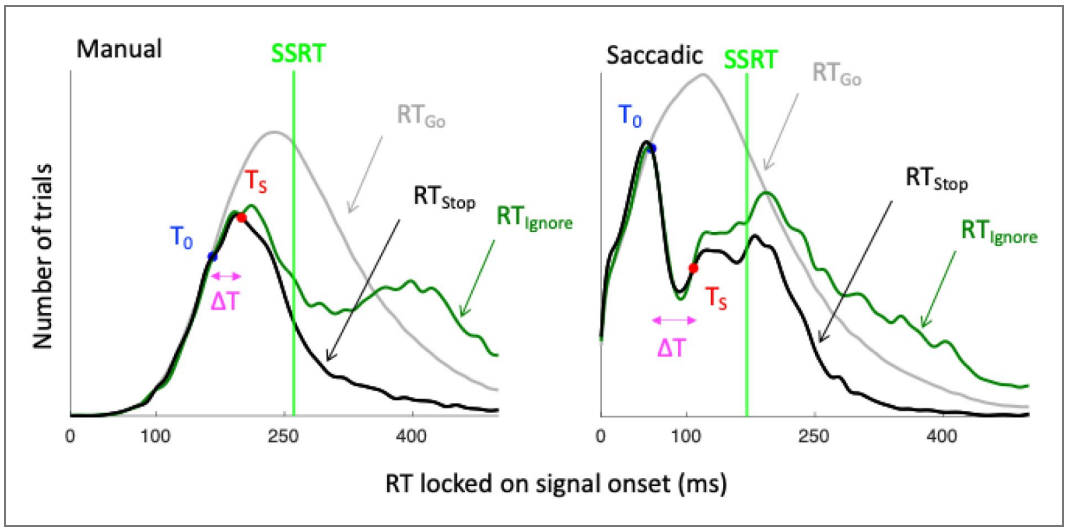


Fig. 2. Individual outcome measures from the novel selective stop-task.

Red reflects participants who were excluded from SSRT and/or ΔT analyses due to meeting exclusion criteria or with ambiguous T_0 or T_S. Pale blue are participants included in both SSRT and ΔT analyses (manual N = 35, saccadic N = 24). Bold blue is the group-level means calculated using participants plotted in pale blue.

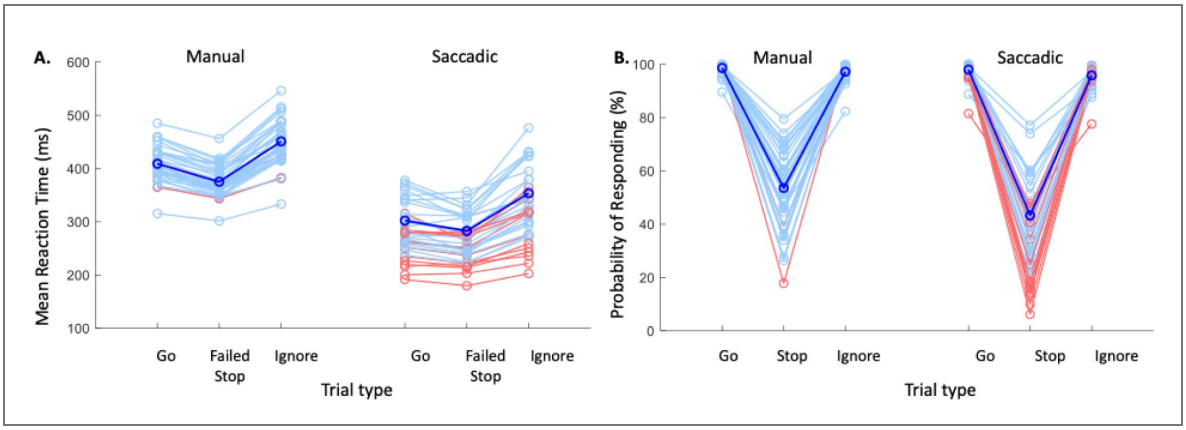


Fig. 3. **Individual stopping accuracy across SOAs.**

Same conventions as in Fig. 2.

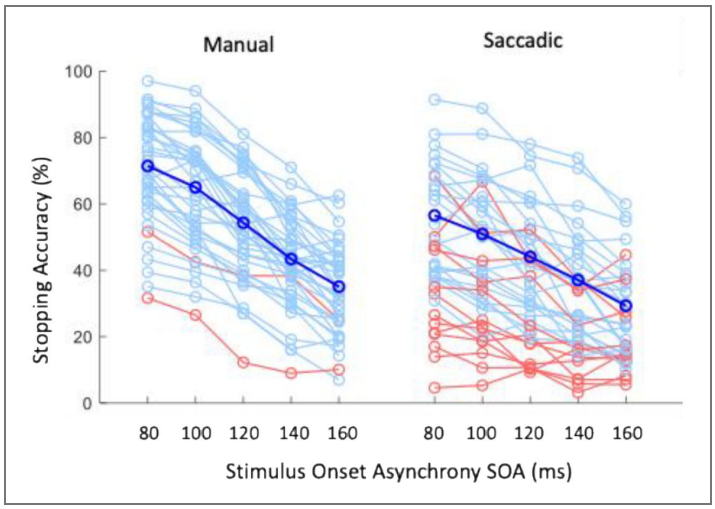
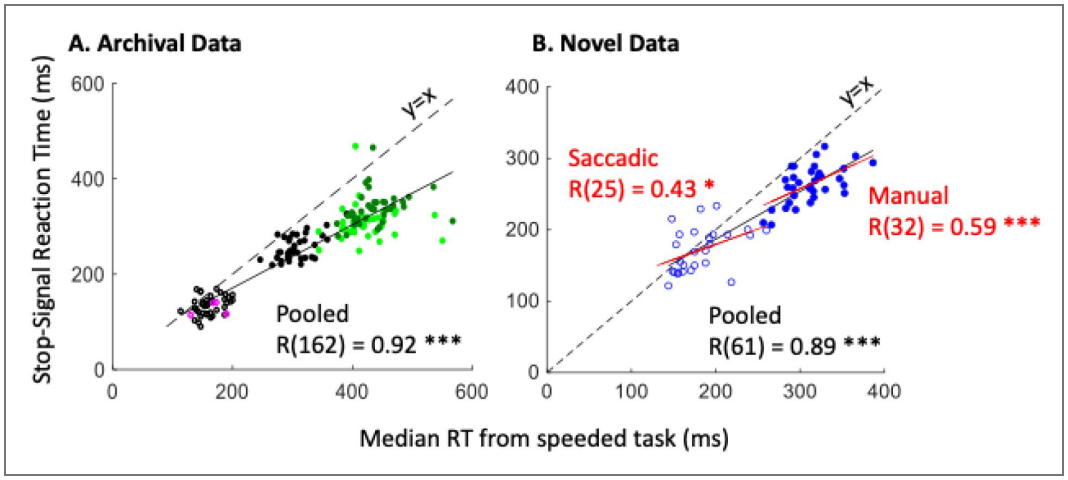


Fig. 4. **Stop signal reaction time against speeded reaction time for archival (A) and novel data (B).**

Median RTs are calculated from no-signal trials in ignore blocks for Campbell et al. (2017) (black circles) and Bompas et al., (2020) experiments 1–3 (magenta), and from go-only blocks for Bompas et al., (2025) experiment 1 (RT to bright and dim stimuli shown as bright and dark green circles) and the novel data. Data points correspond to individuals. The continuous lines show the linear regression, the dashed lines show unity for reference. R-values are provided on each panel, with their degrees of freedom within brackets. * indicates p-values < 0.05 and *** indicates p-values < 0.0001. The grey dotted line indicates the linear regression when manual and saccadic responses are pooled.



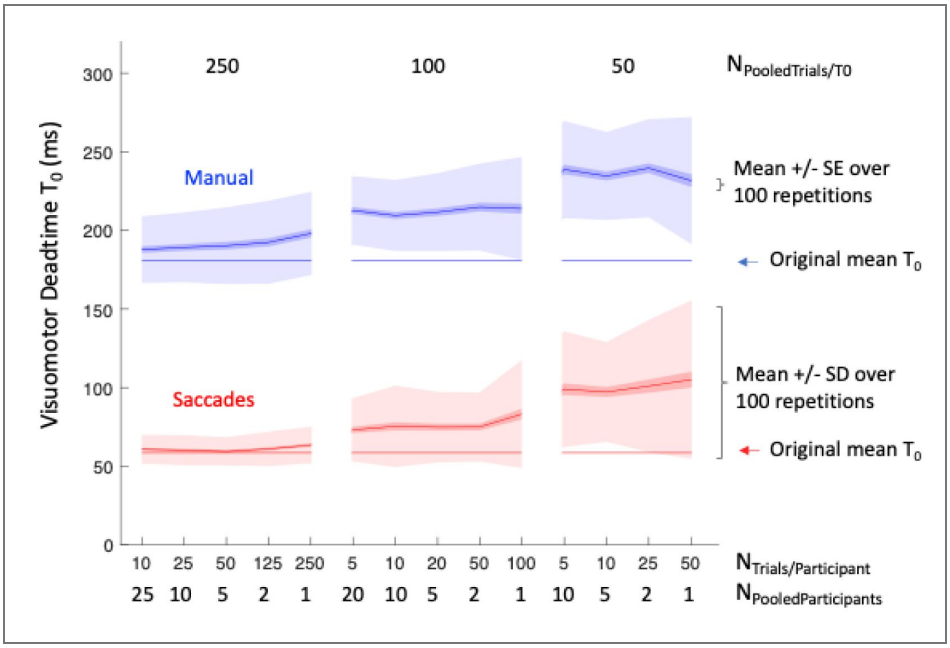


Fig. 5. Effect of trial number on T_0 estimates.

Mean, standard deviation (SD) and standard error (SE) are based on 100 independent bootstrapping of the novel dataset. Each repetition involved subsampling without replacement 250, 100 or 50 stop-signal trials ($N_{PooledTrials/T0}$) and twice more go trials, and extracting a T_0. Trials were obtained from various combinations of number of participants ($N_{PooledParticipants}$) and number of trials per participant ($N_{Trials/Participant}$). $N_{PooledParticipants}$ of 1 means data were not pooled across participants. Stop trials were drawn across the 5 SOAs in equal proportion. For instance, each repetition contributing to the left-most values were obtained by randomly selecting 25 participants out of 37, then randomly selecting and pooling 10 stoptrials (2 trials at each SOA) out of 500, and 20 go trials out of 1000 from each selected participant. Analyses were run separately for manual (blue) and saccadic (red) data. Original mean T_0 (straight horizontal lines) are the mean over all 37 participants, using all 500 stop-trials and 1000 go trials, serving here as ground truth. Reducing the number of trials available per estimate ($N_{PooledTrials/T0}$) leads to a clear increase in the mean and variability of these estimates. Different strategies for achieving a given number of trials (combinations of $N_{PooledParticipants}$ and $N_{Trials/Participant}$) had no visible impact.

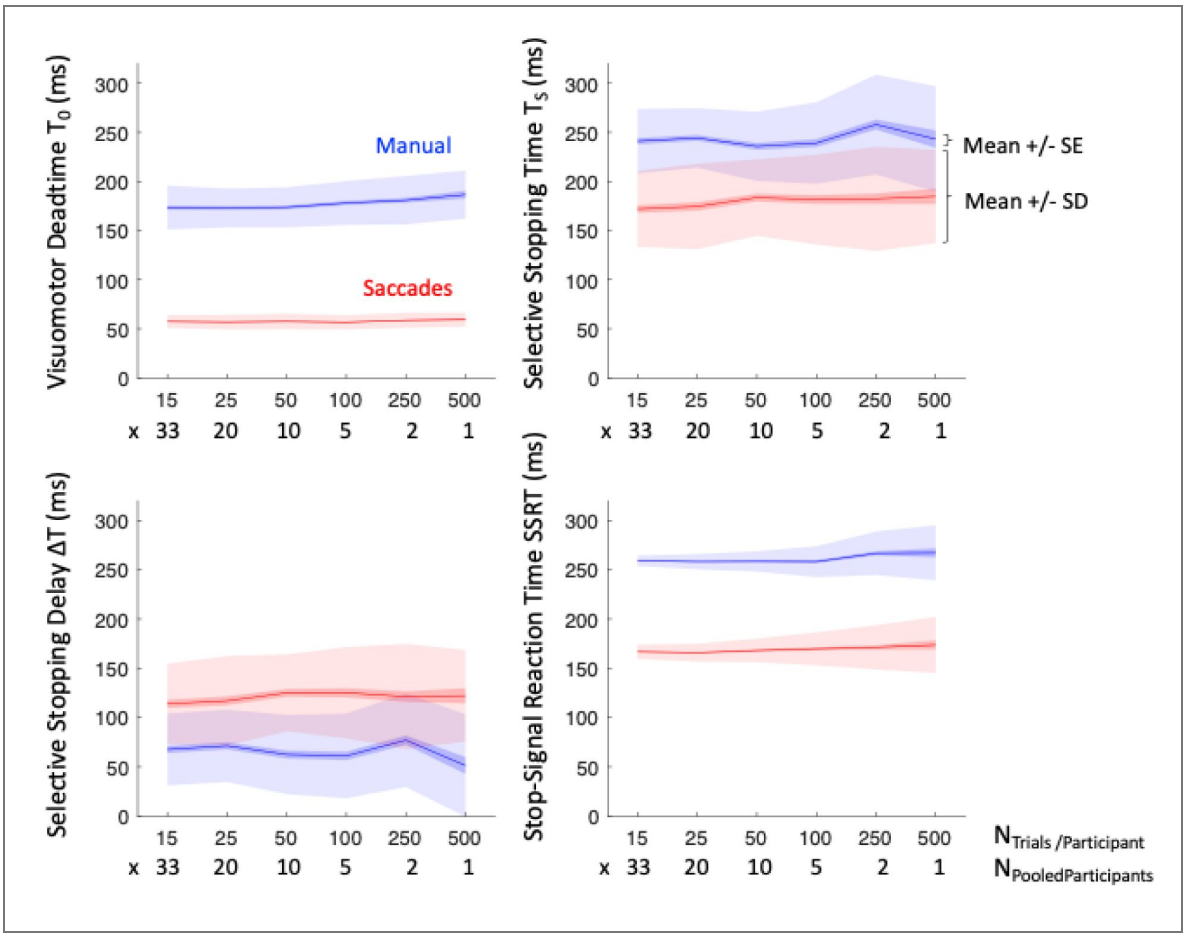


Fig. 6. Pooling participants is a sound way to obtain enough trials to support RT distributional analyses.

Each panel shows a different index (T_0, T_S, ΔT and SSRT). Each index on each repetition is estimated from 500 stop, 500 ignore and 1000 go trials, obtained by randomly selecting and pooling N_{Trials} stop, N_{Trials} ignore and 2*N_{Trials} go, from $N_{PooledParticipants}$. Rightmost values (500 x 1) show the mean, standard error (SE) and standard deviation (SD) across all 37 participants, estimated using all available trials from each participant (same values as presented in the main article). For all other sampling levels (15 x 33 to 250 x 2), mean, SE and SD are based on 100 independent repetitions of randomly selecting trials and participants without replacement (same principle and convention as Fig. 5). T_0 and SSRT were extractedfrom stop and go trials only. T_S and ΔT were extracted based on all trial types.

Fig. 7. Example trials from the two archival selective stopping datasets from main paper Table 1 and Fig. 4.

A. Two examples for each of the three trial types involved in Bompas et al. (2024) experiment 2. Each block used a different triplet of letters (here O, Q and G), with two letters serving as targets (here O and Q), and the third serving as stop-signal (here G). Ignore trials presented the same target letter again at the alternative position, following a delay. Text in brackets shows the instruction associated with each trial type and were not visible to participants. **B.** Three trial types involved in Bompas et al. (2020) experiment 3. The colour of peripheral targets (black or white small disk) varied across blocks, followed by a signal (larger central disk) on 50% of trials. Ignore signals had the same polarity as the target, stop signals had the opposite polarity.

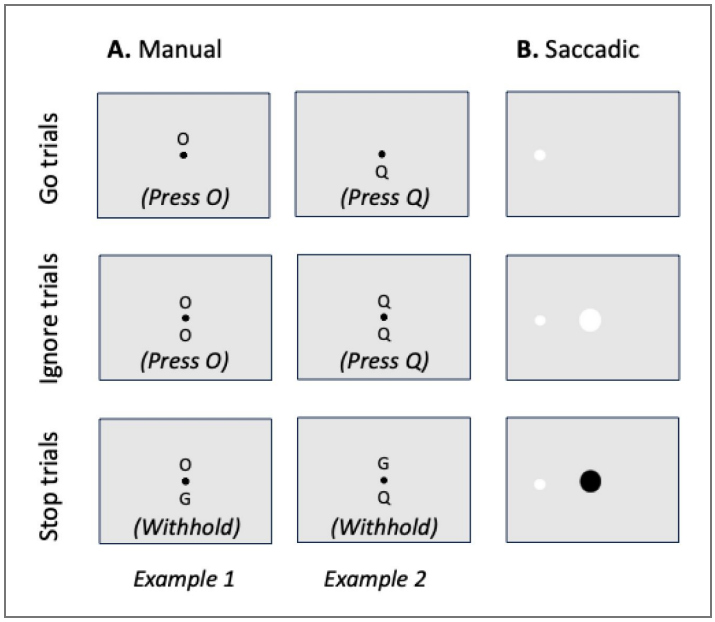
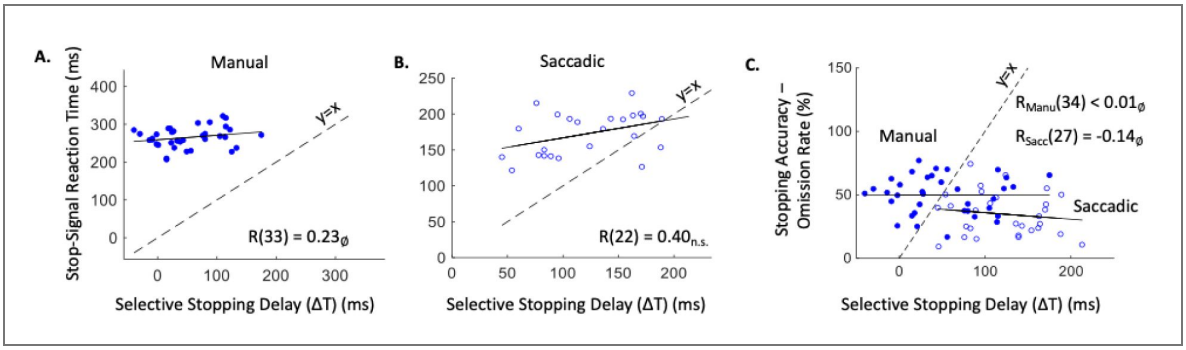


Fig. 8. Lack of correlation between ΔT, SSRT and selective stopping accuracy.

Dashed lines show the identity whilst solid lines are regression lines. Pearson R-values are provided on each panel, with their degrees of freedom within brackets. Φindicates a p-value > 0.05 and BF_{10} < 0.33 (evidence for the null). n.s. indicates a p-value > 0.05 and inconclusive BF_{10} (between 0.33 and 3).



Data availability

All anonymized partly preprocessed data, all task material and analyses scripts for the novel study are deposited in the Open Science Framework (*https://osf.io/n4xek*), together with the fully preprocessed data and scripts for the novel analyses on archival data. Raw eye-tracking files are available on request. Raw data from all archival datasets are available (*https://osf.io/gz9uc*) except for Boucher et al. (2007) which can be requested to the corresponding author, Aline Bompas, pending permission from their owners (Jeffrey Schall).

Acknowledgements

This research was funded by Cardiff University's School of Psychology. We thank Jenna Evans, Tessa Graftdijk for their help with data collection, and Christopher D. Chambers for providing pre-submission feedback on the manuscript.

Additional information

CRediT statement

Heather Statham led on the investigation, data curation, formal analysis and visualisation. She coled on software and supported the conceptualisation, methodology, writing–original draft, writing–review and editing. Phil Schmidt co-led on software and supported the methodology and writing–review and editing. Petroc Sumner supported the conceptualization and writing–review and editing. Aline Bompas led the conceptualisation, methodology, writing – original draft and writing – review and editing and supported the data curation, formal analysis and visualisation.

Funding

Funder	Grant reference number	Author
Cardiff University		Heather Rachel Statham Petroc Sumner Aline Bompas Phil Schmid

Author ORCID iDs

Aline Bompas: https://orcid.org/0000-0002-6957-2694

References

1. **Lipszyc J,** Schachar R (2010) Inhibitory control and psychopathology: A meta-analysis of studies using the stop signal task. *J Int Neuropsychol Soc* **16**:1064-1076 https://doi.org/10.1017/s1355617710000895 | PubMed
2. **Alderson RM,** Rapport MD, Kofler MJ (2007) Attention-Deficit/Hyperactivity Disorder and Behavioral Inhibition: A Meta-Analytic Review of the Stop-signal Paradigm. *J Abnorm Child Psychol* **35**:745-758 https://doi.org/10.1007/s10802-007-9131-6 | PubMed
3. **Senkowski D,** et al. (2024) Assessing Inhibitory Control Deficits in Adult ADHD: A Systematic Review and Meta-analysis of the Stop-signal Task. *Neuropsychol Rev* **34**:548-567 https://doi.org/10.1007/s11065-023-09592-5 | PubMed
4. **Mar K,** Townes P, Pechlivanoglou P, Arnold P, Schachar R (2021) Obsessive Compulsive Disorder and Response Inhibition: Meta-analysis of the Stop-Signal Task. *bioRxiv* https://doi.org/10.1101/2021.07.16.452538

5. **Manza P,** Amandola M, Tatineni V, Li CR, Leung H.-C (2017) Response inhibition in Parkinson's disease: a meta-analysis of dopaminergic medication and disease duration effects. *Npj Park Dis* **3**:23 https://doi.org/10.1038/s41531-017-0024-2 | PubMed
6. **Arnatkeviciute A,** et al. (2023) Trans-ancestry meta-analysis of genome wide association studies of inhibitory control. *Mol Psychiatry* **28**:4175-4184 https://doi.org/10.1038/s41380-023-02187-9 | PubMed
7. **He Q,** Geißler CF, Ferrante M, Hartwigsen G, Friehs MA (2024) Effects of transcranial magnetic stimulation on reactive response inhibition. *Neurosci Biobehav Rev* **157**:105532 https://doi.org/10.1016/j.neubiorev.2023.105532 | PubMed
8. **Hannah R,** Aron AR (2021) Towards real-world generalizability of a circuit for action-stopping. *Nat Rev Neurosci* **22**:538-552 https://doi.org/10.1038/s41583-021-00485-1 | PubMed
9. **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 https://doi.org/10.1037/0033-295x.91.3.295
10. **Lappin JS,** Eriksen CW (1966) Use of a delayed signal to stop a visual reaction-time response. *J Exp Psychol* **72**:805-811 https://doi.org/10.1037/h0021266
11. **Logan GD,** Cowan WB, Davis KA (1984) On the ability to inhibit simple and choice reaction time responses: a model and a method. *J Exp Psychol Hum Percept Perform* **10**:276-291 https://doi.org/10.1037//0096-1523.10.2.276 | PubMed
12. **Casey BJ,** et al. (2018) The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn Neurosci* **32**:43-54 https://doi.org/10.1016/j.dcn.2018.03.001 | PubMed
13. **Vidailhet M** (2024) Eyes as a window to brain pathology in parkinson's disease: a narrative review. *J Neural Transm* **131**:1155-1158 https://doi.org/10.1007/s00702-024-02820-z | PubMed
14. **Meachon EJ,** Klupp S, Grob A (2023) Gait in children with and without ADHD: A systematic literature review. *Gait Posture* **104**:31-42 https://doi.org/10.1016/j.gaitpost.2023.06.003 | PubMed
15. **Ferruccio NP,** et al. (2021) Neurological Signs at the First Psychotic Episode as Correlates of Long Term Outcome: Results From the AESOP-10 Study. *Schizophr Bull* **47**:118-127 https://doi.org/10.1093/schbul/sbaa089 | PubMed
16. **Hedge C,** Powell G, Sumner P (2018) The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behav Res Methods* **50**:1166-1186 https://doi.org/10.3758/s13428-017-0935-1 | PubMed
17. **Thunberg C,** Wiker T, Bundt C, Huster RJ (2024) On the (un)reliability of common behavioral and electrophysiological measures from the stop signal task: Measures of inhibition lack stability over time. *Cortex* **175**:81-105 https://doi.org/10.1016/j.cortex.2024.02.008 | PubMed
18. **Matzke D,** Love J, Heathcote A (2017) A Bayesian approach for estimating the probability of trigger failures in the stop-signal paradigm. *Behav Res Methods* **49**:267-281 https://doi.org/10.3758/s13428-015-0695-8 | PubMed
19. **Bissett PG,** Jones HM, Poldrack RA, Logan GD (2021) Severe violations of independence in response inhibition tasks. *Sci Adv* **7**:eabf4355 https://doi.org/10.1126/sciadv.abf4355 | PubMed
20. **Verbruggen F,** et al. (2019) A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife* **8**:e46323 https://doi.org/10.7554/eLife.46323 | PubMed
21. **Colonius H,** Diederich A (2018) Paradox resolved: Stop signal race model with negative dependence. *Psychol Rev* **125**:1051-1058 https://doi.org/10.1037/rev0000127 | PubMed
22. **Bompas A,** Sumner P, Hedge C (2025) Non-decision time: The Higgs Boson of decision. *Psychol Rev* **132**:330-363 https://doi.org/10.1037/rev0000487 | PubMed
23. **Campbell AE,** Chambers CD, Allen CPG, Hedge C, Sumner P (2017) Impairment of manual but not saccadic response inhibition following acute alcohol intoxication. *Drug Alcohol Depend* **181**:242-254 https://doi.org/10.1016/j.drugalcdep.2017.08.022 | PubMed

24. **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 https://doi.org/10.3758/bf03193779 | PubMed

25. **Bompas A,** Campbell AE, Sumner P (2020) Cognitive control and automatic interference in mind and brain: A unified model of saccadic inhibition and countermanding. *Psychol Rev* **127**:524-561 https://doi.org/10.1037/rev0000181 | PubMed

26. **Boag RJ,** et al. (2025) An Expert Guide to Planning Experimental Tasks For Evidence-Accumulation Modeling. *Adv Methods Pract Psychol Sci* **8**:25152459251336127 https://doi.org/10.1177/25152459251336127

27. **Bompas A,** Sumner P (2011) Saccadic Inhibition Reveals the Timing of Automatic and Voluntary Signals in the Human Brain. *J Neurosci* **31**:12501-12512 https://doi.org/10.1523/jneurosci.2234-11.2011 | PubMed

28. **Bompas A,** Hedge C, Sumner P (2017) Speeded saccadic and manual visuo-motor decisions: Distinct processes but same principles. *Cognit Psychol* **94**:26-52 https://doi.org/10.1016/j.cogpsych.2017.02.002 | PubMed

29. **Reingold EM,** Stampe DM (2004) Saccadic Inhibition in Reading. *J Exp Psychol Hum Percept Perform* **30**:194-211 https://doi.org/10.1037/0096-1523.30.1.194 | PubMed

30. **Stampe DM,** Reingold EM (2002) Influence of stimulus characteristics on the latency of saccadic inhibition. *Progress in Brain Research* **140**:73-87 https://doi.org/10.1016/s0079-6123(02)40043-x | PubMed

31. **Harrison JJ,** Freeman TCA, Sumner P (2014) Saccade-like behavior in the fast-phases of optokinetic nystagmus: An illustration of the emergence of volitional actions from automatic reflexes. *J Exp Psychol Gen* **143**:1923-1938 https://doi.org/10.1037/a0037021 | PubMed

32. **Rolfs M,** Kliegl R, Engbert R (2008) Toward a model of microsaccade generation: The case of microsaccadic inhibition. *J Vis* **8**:5 https://doi.org/10.1167/8.11.5 | PubMed

33. **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 https://doi.org/10.1037/0033-295x.114.2.376 | PubMed

34. **Salinas E,** Stanford TR (2013) The Countermanding Task Revisited: Fast Stimulus Detection Is a Key Determinant of Psychophysical Performance. *J Neurosci* **33**:5668-5685 https://doi.org/10.1523/jneurosci.3977-12.2013 | PubMed

35. **Logan GD,** Irwin DE (2000) Don't look! don't touch! inhibitory control of eye and hand movements. *Psychon Bull Rev* **7**:107-112 https://doi.org/10.3758/bf03210728 | PubMed

36. **Aron AR** (2011) From Reactive to Proactive and Selective Control: Developing a Richer Model for Stopping Inappropriate Responses. *Biol Psychiatry* **69**:e55-e68 https://doi.org/10.1016/j.biopsych.2010.07.024 | PubMed

37. **Bissett PG,** Poldrack RA (2022) Estimating the Time to Do Nothing: Toward Next-Generation Models of Response Inhibition. *Curr Dir Psychol Sci* **31**:556-563 https://doi.org/10.1177/09637214221121753

38. **McPhee MD,** Hendershot CS (2023) Meta-analysis of acute alcohol effects on response inhibition. *Neurosci Biobehav Rev* **152**:105274 https://doi.org/10.1016/j.neubiorev.2023.105274 | PubMed

39. **Kenemans JL,** Schutte I, Van Bijnen S, Logemann HNA (2023) How salience enhances inhibitory control: An analysis of electro-cortical mechanisms. *Biol Psychol* **177**:108505 https://doi.org/10.1016/j.biopsycho.2023.108505 | PubMed

40. **Pani P,** et al. (2018) Visual salience of the stop signal affects the neuronal dynamics of controlled inhibition. *Sci Rep* **8**:14265 https://doi.org/10.1038/s41598-018-32669-8 | PubMed

41. **Verbruggen F,** Stevens T, Chambers CD (2014) Proactive and reactive stopping when distracted: An attentional account. *J Exp Psychol Hum Percept Perform* **40**:1295-1300 https://doi.org/10.1037/a0036542 | PubMed

42. **Verbruggen F,** Aron AR, Stevens MA, Chambers CD (2010) Theta burst stimulation dissociates attention and action updating in human inferior frontal cortex. *Proc Natl Acad Sci* **107**:13966-13971 https://doi.org/10.1073/pnas.1001957107 | PubMed

43. **Logan GD,** Burkell J (1986) Dependence and Independence in Responding to Double Stimulation: A Comparison of Stop, Change, and Dual-Task Paradigms. *Journal of Experimental Psychology: Human Perception and Performance* **12**:549-563 https://doi.org/10.1037/0096-1523.12.4.549

44. **Boehler CN,** et al. (2009) Sensory MEG Responses Predict Successful and Failed Inhibition in a Stop-Signal Task. *Cereb Cortex* **19**:134-145 https://doi.org/10.1093/cercor/bhn063 | PubMed

45. **Hervault M,** Zanone P.-G, Buisson J.-C, Huys R (2022) Hold your horses: Differences in EEG correlates of inhibition in cancelling and stopping an action. *Neuropsychologia* **172**:108255 https://doi.org/10.1016/j.neuropsychologia.2022.108255 | PubMed

46. **Huster RJ,** Messel MS, Thunberg C, Raud L (2020) The P300 as marker of inhibitory control - Fact or fiction?. *Cortex* **132**:334-348 https://doi.org/10.1016/j.cortex.2020.05.021 | PubMed

47. **Schaum M,** et al. (2021) Right inferior frontal gyrus implements motor inhibitory control via beta band oscillations in humans. *eLife* **10**:e61679 https://doi.org/10.7554/eLife.61679 | PubMed

48. **Swann N,** et al. (2009) Intracranial EEG Reveals a Time- and Frequency-Specific Role for the Right Inferior Frontal Gyrus and Primary Motor Cortex in Stopping Initiated Responses. *J Neurosci* **29**:12675-12685 https://doi.org/10.1523/jneurosci.3359-09.2009 | PubMed

49. **Wagner J,** Wessel JR, Ghahremani A, Aron AR (2018) Establishing a Right Frontal Beta Signature for Stopping Action in Scalp EEG: Implications for Testing Inhibitory Control in Other Task Contexts. *J Cogn Neurosci* **30**:107-118 https://doi.org/10.1162/jocn_a_01183 | PubMed

50. **Zhao W,** et al. (2018) Response Inhibition Deficits in Insomnia Disorder: An Event-Related Potential Study With the Stop-Signal Task. *Front Neurol* **9**:610 https://doi.org/10.3389/fneur.2018.00610 | PubMed

51. **Wessel JR,** Aron AR (2015) It's not too late: the onset of the frontocentral P3 indexes successful response inhibition in the stop-signal paradigm. *Psychophysiology* **52**:472-480 https://doi.org/10.1111/psyp.12374 | PubMed

52. **Statham H,** Bompas A, Schmid P, Singh K, Tait L (2025) Automatic and Selective Inhibition in the Brain: A MEG and Selective Stopping Study. *OSF* https://osf.io/a7s8r

53. **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 https://doi.org/10.1037/a0038893 | PubMed

54. **Congdon E,** et al. (2012) Measurement and Reliability of Response Inhibition. *Front Psychol* **3** https://doi.org/10.3389/fpsyg.2012.00037 | PubMed

55. **Krekelberg B** (2025) bayesFactor. Mathworks. version: 1.0.0 https://uk.mathworks.com/matlabcentral/fileexchange/69794-bayesfactor

56. **Rouder JN,** Morey RD (2012) Default Bayes Factors for Model Selection in Regression. *Multivar Behav Res* **47**:877-903 https://doi.org/10.1080/00273171.2012.734737 | PubMed

57. **Colonius H,** Diederich A (2018) Paradox resolved: Stop signal race model with negative dependence. *Psychol Rev* **125**:1051-1058 https://doi.org/10.1037/rev0000127 | PubMed

Peer reviews

Reviewer #1 (Public review):

This interesting manuscript challenges the current interpretation of the well-established stop signal reaction time task (SSRT), commonly used in many research areas. SSRT has been traditionally thought of as primarily a measure of inhibitory control. In this work, the authors argue that this is influenced significantly by sensory and motor transmission times,

and that these low-level processes may systematically confound SSRT estimates both in individual groups and also in many clinical populations.

Conceptually, this raises an important and significant question regarding the construct validity of one of the more widely used behavioral measures of response inhibition. The authors provide a clear theoretical framing and importantly address the overlooked assumptions in SSRT modelling, the underexplored source of variability.

Further, direct evidence separating peripheral sensory and motor contributions from central inhibitory processes is certainly needed, as well as a more balanced interpretation of prior methodological refinements in the field. Is a correlation between T0 and SSRT sufficient to conclude that SSSRT may be predominantly driven by peripheral delays? What proportion of SSRT variance remains unexplained after one accounts for T0? If T0 and inhibitory processes share neural processing speed, does that mean they covary? If one corrects T0, do the group differences become smaller or perhaps disappear?

Furthermore, how robust is the T0 estimate in noisy environments of all sorts and especially across modalities (visual and motor)? The authors argue that this relationship is clearer in "good quality data"; how do you define that, and what happens with noisy data? For instance, the authors refer to a range of clinical disorders such as ADHD and PD where noise is abundantly present, partly due to the disease itself or due to treatment.

In conclusion, this is certainly a thought-provoking and potentially influential contribution in the literature that raises important questions about the interpretation of the stop signal reaction time task.

https://doi.org/10.7554/eLife.111819.1.sa3

Reviewer #2 (Public review):

Continuing their work distinguishing sensory latencies of "cognitive" processes, the authors turn their attention to "response inhibition". The "square quotes" are being used to highlight how this manuscript aims to challenge previous descriptions of performance data and inferred computational processes. The authors assert that previous descriptions of the measure known as "stop signal reaction time" (SSRT) are flawed because they did not account for sensory latencies empirically or theoretically.

Enthusiasm for the manuscript cannot be high in light of many weaknesses countering the possible strengths. Strengths include offering an opportunity to more carefully characterize the quantity SSRT and a specific empirical approach offered to the research community. However, these strengths are countered by the following structural, theoretical, and empirical weaknesses:

As announced by the elephant in the title, the writing could be described as excessively polemical. However, the characterization and interpretation of previous empirical and theoretical work is disputable.

The major theoretical claim regarding sensory delays inherent in SSRT is not novel. The authors assert, "...this corpus of work may have been misinterpreted because the SSRT is systematically influenced by low level sensory and motor transmission times, arguably more so than by inhibition or cognitive processes." This was certainly recognized by Logan and Cowan in their original work. They wrote, "An act of control, like any other act, must take time. The theory provides methods for measuring the latency of control even when the act of control is not directly observable." (page 298) Also, "... the estimate of stop-signal reaction time includes the latency of the internal response to the stop signal and the duration of the ballistic process." (page 316-317). Moreover, subsequent computational and empirical work,

some noted by the authors, has distinguished the sensory encoding interval from the interval during which the STOP process interrupts the GO process.

The theoretical suggestion that an accounting for sensory delays undermines the functional interpretation of SSRT mischaracterizes the original literature. For example, in the Abstract the authors write "Sensory and motor contributions must be ruled out before linking SSRT results to inhibition or cognition". The original Logan and Cowan theory was about what happens at the end of SSRT, and that was described only as an "act of control", in perfectly positivist fashion. For example, Logan and Cowan wrote, "Estimates of stop-signal reaction time provide a measure of the latency of control." (page 315). Thus, the authors are misstating what was meant originally by SSRT. In addition, the authors offer no specific or formal definition to specify what they mean by "inhibition or cognitive processes".

Confidence in the new empirical conclusions of the manuscript must be low because the new performance data are of questionable quality. The first issue is that the stopping accuracy (or inhibition functions in original terminology) shown in Figure S3 is very problematic for the interpretation of the authors' empirical work in this manuscript. There are two problems. First, these plots should span from nearly 0% to nearly 100%. It is not possible to resolve the span of each individual in the figure, but it is clear that many, if not most, in both the Manual and Saccadic data span just 20-30%. Second, the plots should span the 50% success value. It is clear that the maximum or minimum values for many participants do not reach the 50% value. These two problems indicate that many (most?) participants were not really sensitive to the stop signal.

The second issue concerns the pattern of response times (RTs) on "ignore" trials. The authors portray performance as exemplifying a "pause-then-go" strategy. This is not uncommon, but it is not the only way participants perform. Many participants across multiple studies of selective stimulus stopping produce RTs on "Ignore" trials essentially indistinguishable from RTs on no-stop trials. The authors must acknowledge and account for such individual variability. In fact, the "T_s" value is measured by the difference in distributions of RT on no-signal and ignore trials. If these distributions are not different, then the measurement and interpretation of this quantity is questionable.

Related, the distributions of RT on stop trials, particularly for saccade responses, are portrayed with a second mode in the schematic illustrations and clearly peaking at SSRT in Figure S1. This second mode is not observed in other saccade stop signal studies. This indicates that the participants in this study were in a peculiar mode of performance.

Finally, given the pivotal role of measures of differences of RT distributions and the pronounced variation of stopping accuracy (Figure S3), the authors must show the distributions for all of their new participants. The authors' claim to higher resolution obliges them to reveal every step of analysis.

In its current form, this manuscript is unlikely to change the thinking of modelers or practitioners of the stop signal task.

https://doi.org/10.7554/eLife.111819.1.sa2

Reviewer #3 (Public review):

Summary:

Statham and colleagues test an assumption underpinning a very large literature: that the stop-signal reaction time (SSRT) indexes the speed or efficacy of top-down inhibitory control. They argue instead, and support their claims with a total of eight datasets, that SSRT is substantially occupied by visuomotor deadtime (i.e., incompressible sensory and motor

delays common to all visually guided responses), which varies across individuals, conditions and populations in ways that mimic effects usually attributed to inhibitory control. They propose two remedies: subtracting an independent estimate of visuomotor deadtime (T_0) from SSRT, and a new index, the selective stopping delay (ΔT), from the stimulus-selective stopping task.

Strengths:

The paper's principal strength is the combination of these components. That SSRT must contain peripheral delays is not itself new, as the authors point out (Boucher et al., 2007; Salinas and Stanford, 2013; Bompas et al., 2020). What is new is the quantification of the problem at scale, across seven archival datasets and a preregistered replication, together with the demonstration that T_0 can be recovered from existing stop-task data. That is important, as it provides a diagnostic that can be applied to data already collected. The authors' offer to assist others in doing so is exemplary. The supplementary analyses of trial numbers and participant pooling are very useful, and the paper provides important sanity checks, notably confirming that stop and ignore signals produce indistinguishable initial interference before pooling them.

Weaknesses:

The evidence for the central claim is strong but presented in a way that overstates it. Figure 2 reports 85% and 80% shared variance between SSRT and T_0, but these pool across datasets and, more critically, across response modality: manual and saccadic estimates from the same participants are plotted together with a single regression line through both. Because manual and saccadic deadtimes differ by roughly 130 ms, the resulting correlation largely reflects a between-condition difference rather than covariation among individuals. The numbers that speak to individual differences are more modest (40% for manual responses; 7% for saccades). The manual result is convincing and consequential; the saccadic result is not, and the explanation in terms of restricted range, while plausible, is offered after the fact and is directly testable by reporting the reliability of saccadic T_0 or correcting the correlation for attenuation. This limitation is arguably good news for the paper's practical message, since it implies saccadic measures are relatively protected, but the manuscript should make clear (including in the abstract) that the strong individual-differences case rests on the manual data, where motor execution delay is the main driver.

A related point concerns interpretation rather than analysis. Since SSRT is, on the authors' own account, approximately the sum of T_0 and a decision-related component, covariation between the two is expected on structural grounds; the preregistered correlation with reaction times from separate speeded blocks mitigates this, but the finding is less surprising than its current framing implies. What would determine whether past conclusions must be revised is not whether SSRT correlates with T_0 across individuals, but whether the decision-related component tracks the independent variable in any given study. The alcohol reanalysis could be a test case for this: the authors show that alcohol raises T_0 commensurately with SSRT and conclude the effects are "consistent with these effects being fully driven by visuomotor delays," yet (unless I missed something) they do not report the corrected measure for these data, while they do so for signal contrast and response modality. Running that analysis, and stating plainly what Campbell et al. (2017) would have concluded under the proposed treatment, would be an important demonstration.

The case for ΔT is the least developed part of the paper. ΔT is a difference between two independently estimated, individually noisy quantities, extracted by a non-trivial procedure (see also below), and no reliability estimates are reported for T_0, TS or ΔT. This would be possible based on the two-session design (and the group has prior work on the reliability of cognitive control measures). This matters because the argument that ΔT is superior rests, to some extent, on null findings: ΔT does not correlate with SSRT, with stopping accuracy, or

with the differential response to stop and ignore trials. These null correlations are interpreted as freedom from confounds, but an unreliable measure would produce the same pattern, and the seven participants with implausible negative ΔT values indicate that noise is not negligible.

In addition, the subjective correction of dip onsets ("Departure points were visually inspected and adjusted if it was deemed that the algorithm had placed them in inappropriate places"), which is critical to the paper's central measurement, should be blinded to condition or show inter-rater agreement. Since T_0 and TS are compared across conditions and ΔT is their difference, this introduces researcher degrees of freedom.

This is particularly critical when a dip is not easy to extract. Figure 3 depicts an idealized ignore-trial distribution with a clean, deep dip. Real distributions are unlikely to look like this, and dip depth should depend on the behavioral relevance and salience of the ignored event; published work on rapid manual inhibition indicates that dips to behaviorally irrelevant events can be very shallow. Since ΔT is extractable only where the dip is resolvable, the generality of the method can be questioned. Ideally, the empirical distributions underlying every dataset analyzed should be shown to alleviate this concern.

One uncontrolled procedural difference also deserves comment. Corrective feedback about stopping too often or stopping too rarely was given after manual blocks only; saccadic blocks received none, and fixed rather than staircased delays were used throughout. Since the manual-saccadic contrast carries much of the argument, and saccadic blocks yielded both lower stopping accuracy (36% vs 52%) and far more exclusions (8 vs 1 of 37), this asymmetry offers an alternative to the interpretation that saccades are simply harder to inhibit.

A final point concerns the comparison between response modalities. Raw SSRT suggests that saccadic inhibition is faster than manual (174 vs 266 ms), while both proposed corrections reverse this, with SSRT−T_0 and ΔT each indicating that saccadic inhibition is slower (the latter consistently across nearly every participant). This is one of the clearest illustrations of the paper's thesis, but it is not taken up in the discussion, which returns to modality only to note that saccadic T_0 varies little (the one reference to variation across action modalities appears in the modeling section, without stating its direction). It would also benefit from a caveat. Both corrected measures subtract the same T_0, and manual and saccadic T_0 differ by roughly 130 ms, so the two do not corroborate one another independently (TS is itself longer for manual responses, and yields a shorter ΔT only once the larger manual T_0 is removed). The accuracy of the subtraction therefore matters here: if the manual regression slope of 0.75 reflects sub-additivity rather than attenuation, subtracting the full T_0 would overcorrect manual responses more than saccadic ones, and could produce the reversal on its own.

These concerns qualify rather than undermine the contribution. The core observation is robust, the diagnostic is practical and immediately applicable, and the case that a large body of work requires re-examination is well made. If the corrected analyses are carried through on the datasets already in hand, this will be an important paper for anyone who uses the stop-signal task.

https://doi.org/10.7554/eLife.111819.1.sa1

Author response:

We are grateful to the editor and reviewers for providing their time and expertise in the assessment of this article. We are glad that the overall evaluation is supportive.

> ***Reviewer #1 (Public review):***

> *This interesting manuscript challenges the current interpretation of the well-established stop signal reaction time task (SSRT), commonly used in many research areas. SSRT has been traditionally thought of as primarily a measure of inhibitory control. In this work, the authors argue that this is influenced significantly by sensory and motor transmission times, and that these low-level processes may systematically confound SSRT estimates both in individual groups and also in many clinical populations.*
>
> *Conceptually, this raises an important and significant question regarding the construct validity of one of the more widely used behavioral measures of response inhibition. The authors provide a clear theoretical framing and importantly address the overlooked assumptions in SSRT modelling, the underexplored source of variability.*
>
> *Further, direct evidence separating peripheral sensory and motor contributions from central inhibitory processes is certainly needed, as well as a more balanced interpretation of prior methodological refinements in the field. Is a correlation between T0 and SSRT sufficient to conclude that SSSRT may be predominantly driven by peripheral delays? What proportion of SSRT variance remains unexplained after one accounts for T0? If T0 and inhibitory processes share neural processing speed, does that mean they covary? If one corrects T0, do the group differences become smaller or perhaps disappear?*

We thank the reviewer for their positive assessment of our work. We agree all these questions are important. Our revision will provide repeatability estimates for all our indices, and use the repeatability of SSRT and T0 to estimate the proportion of SSRT variance that remains unexplained. We will mention that it is theoretically possible that shared neural processing speed between T0 and inhibitory processes contributes to the reported effect but, since the neural pathways are anatomically largely distinct, the available cognitive neuroscience literature suggests such contribution is unlikely to be strong enough to explain the observed relationship. The impact of correcting group differences in SSRT for T0 will entirely depend on where differences in SSRT between these groups come from. If it comes exclusively from peripheral delays, then the effect may indeed disappear. If it is central, then stronger differences may be revealed after correction. We appreciate these are key questions we would like answers to already, but identifying and reanalysing suitable datasets to answer it will be the purpose of future papers.

> *Furthermore, how robust is the T0 estimate in noisy environments of all sorts and especially across modalities (visual and motor)? The authors argue that this relationship is clearer in "good quality data"; how do you define that, and what happens with noisy data? For instance, the authors refer to a range of clinical disorders such as ADHD and PD where noise is abundantly present, partly due to the disease itself or due to treatment.*

By good quality data, we mean enough trials at the optimal RT and SOA combinations, and an adequate preprocessing pipeline that excludes non-standard trials (poor fixation, pre-emptive responses, large undershoot …). Participants with increased intra-individual variance or low compliance will need more trials overall to get enough trials around divergence times for these to be accurately estimated. Supplementary figure 5 illustrates how low trial numbers lead to an overestimation of T0, which can be mitigated by pooling across participants. We expect noisy data to have a similar effect, although its impact may differ across clinical conditions based on where the additional variability comes from. We shall have more clarity on the reliability of T0 in clinical populations once relevant datasets have been reanalysed, and use this to define constraints for future data collection. Again, this will need to wait for future papers.

> *In conclusion, this is certainly a thought-provoking and potentially influential contribution in the literature that raises important questions about the interpretation of the stop signal reaction time task.*

We thank the reviewer for their in-depth and thoughtful comments and suggestions

> ***Reviewer #2 (Public review):***
>
> *Continuing their work distinguishing sensory latencies of "cognitive" processes, the authors turn their attention to "response inhibition". The "square quotes" are being used to highlight how this manuscript aims to challenge previous descriptions of performance data and inferred computational processes. The authors assert that previous descriptions of the measure known as "stop signal reaction time" (SSRT) are flawed because they did not account for sensory latencies empirically or theoretically.*
>
> *Enthusiasm for the manuscript cannot be high in light of many weaknesses countering the possible strengths. Strengths include offering an opportunity to more carefully characterize the quantity SSRT and a specific empirical approach offered to the research community. However, these strengths are countered by the following structural, theoretical, and empirical weaknesses:*
>
> *As announced by the elephant in the title, the writing could be described as excessively polemical. However, the characterization and interpretation of previous empirical and theoretical work is disputable.*
>
> *The major theoretical claim regarding sensory delays inherent in SSRT is not novel. The authors assert, "...this corpus of work may have been misinterpreted because the SSRT is systematically influenced by low level sensory and motor transmission times, arguably more so than by inhibition or cognitive processes." This was certainly recognized by Logan and Cowan in their original work. They wrote, "An act of control, like any other act, must take time. The theory provides methods for measuring the latency of control even when the act of control is not directly observable." (page 298) Also, "... the estimate of stop-signal reaction time includes the latency of the internal response to the stop signal and the duration of the ballistic process." (page 316-317). Moreover, subsequent computational and empirical work, some noted by the authors, has distinguished the sensory encoding interval from the interval during which the STOP process interrupts the GO process.*

We agree that our manuscript should acknowledge that Logan and Cowan (1984) explicitly stated that internal and ballistic delays contribute to SSRT and thank the reviewer for highlighting the need to clarify the relationship between our work and the original Logan and Cowan framework. We will clarify that the novelty of our claim is not that peripheral delays contribute to SSRT, which is a logical necessity, but that differences in peripheral delays (across conditions or people) contribute to differences in SSRT, sometimes to a large extent. Such differences in SSRT are very widely assumed to reflect inhibitory control in the large corpus of work that followed this initial literature. This corpus has essentially ignored the message about stimulus processing and ballistic delays and their implications for individual differences or changes across conditions. Therefore, we maintain that SSRT differences may have been widely misinterpreted. We reference the articles where these implications were clearly spelled out: these empirical and modelling studies were based on a few monkeys or human participants, and therefore unable to provide the large-scale demonstration we provide here.

> *The theoretical suggestion that an accounting for sensory delays undermines the functional interpretation of SSRT mischaracterizes the original literature. For example, in the Abstract the authors write "Sensory and motor contributions must be ruled out*

> *before linking SSRT results to inhibition or cognition". The original Logan and Cowan theory was about what happens at the end of SSRT, and that was described only as an "act of control", in perfectly positivist fashion. For example, Logan and Cowan wrote, "Estimates of stop-signal reaction time provide a measure of the latency of control." (page 315). Thus, the authors are misstating what was meant originally by SSRT. In addition, the authors offer no specific or formal definition to specify what they mean by "inhibition or cognitive processes".*

We thank the reviewer for pointing out that their original approach was mechanistically agnostic, which we will explicitly clarify in revision. We will remove the words “top-down” from our 4th sentence and reword the quoted sentence into "Changes in peripheral delays must be ruled out before linking changes in SSRT to inhibition or cognition". It remains the case that many hundreds of studies have since interpreted “ability to inhibit” and “latency of control” as specific to inhibition and control, and therefore have assumed that changes in SSRT directly reflect an inhibitory cognitive process. We agree that we do not currently offer a definition of what we mean by cognitive processes, except that they do not include incompressible sensory and motor delays. Based on the reviewer’s clarification, this common shortcut in the literature appears inconsistent with the spirit of the initial work, which we seek to rectify.

> *Confidence in the new empirical conclusions of the manuscript must be low because the new performance data are of questionable quality. The first issue is that the stopping accuracy (or inhibition functions in original terminology) shown in Figure S3 is very problematic for the interpretation of the authors' empirical work in this manuscript. There are two problems. First, these plots should span from nearly 0% to nearly 100%. It is not possible to resolve the span of each individual in the figure, but it is clear that many, if not most, in both the Manual and Saccadic data span just 20-30%. Second, the plots should span the 50% success value. It is clear that the maximum or minimum values for many participants do not reach the 50% value. These two problems indicate that many (most?) participants were not really sensitive to the stop signal.*

We acknowledge our inhibition functions are narrow, but we do not believe this undermines the main conclusions, for several reasons. On the question of sensitivity to the stop signal, average spans for inhibition functions after participant exclusions were 37% for manual and 29% for saccades. This limited span is mainly attributable to our fixed SOA design, which was a necessary feature of a direct comparison between manual and saccadic behaviours. Figure S3 covers only 80 ms spread of SOA. The slopes are commensurate with most other studies, which cover much wider differences in SOA. If one selects the central 100 ms (where the slopes are steepest) from the figures in most previous papers, one will find stopping accuracy changes of around 30%. Therefore, sensitivity is similar.

On the question of some functions not crossing 50%, the correlations between SSRT and T0 remain the same if we only keep those participants who crossed the 50% point ($R(26)=0.64$ for manual, $R(11)=0.4$ for saccades, same statistical significance levels). We will additionally rerun our SSRT and T0 correlation with stopping accuracy as a covariate. As stopping accuracy affects SSRT but not T0, it is unlikely to drive our results.

> *The second issue concerns the pattern of response times (RTs) on "ignore" trials. The authors portray performance as exemplifying a "pause-then-go" strategy. This is not uncommon, but it is not the only way participants perform. Many participants across multiple studies of selective stimulus stopping produce RTs on "Ignore" trials essentially indistinguishable from RTs on no-stop trials. The authors must acknowledge and account for such individual variability. In fact, the "T_s" value is measured by the difference in distributions of RT on no-signal and ignore trials. If these distributions are not different, then the measurement and interpretation of this quantity is questionable.*

We agree that the issue raised by the reviewer would be important if no difference were present between the distributions. In our dataset, however, all participants showed a measurable distributional difference. We believe the difference in perspective that 'many participants…produce RTs on ignore trials essentially indistinguishable from RTs on no-stop trials' may be attributed to differences in the way we analyse data (RT distributions versus mean RT).

Nearly all participants in our final sample had mean RTignore – RTgo > 10 ms (significant at the individual level), except for 2 in the manual condition (and none for saccades). Following Bisset & Logan (2014), a lack of clear mean RTignore - RTgo difference in these 2 participants might have been interpreted as reflecting a different strategy. However, all our participants showed clear dips between go and ignore RT distributions when locked on signal onset, and these two manual participants were no exception. Therefore, accounting for response probability and RT at each SOA in our distributional analysis revealed clear ignore versus go differences, masked when relying on mean RT. The lack of mean RT difference for these two participants in the manual modality had no impact on our main hypothesis testing because T0 was extracted by comparing signal-absent and signal-present trials (pooling ignore and stop), while TS was extracted by comparing ignore and stop (not ignore and go).

In terms of interpretation and whether participants employ a pause-then-go strategy, we understand performance in the selective stopping task as reflecting a combination of automatic activation and interference, and endogenous activation and inhibition. Our interpretation is that the pause component primarily reflects automatic interference triggered by stimulus onset, although strategic factors may also contribute in some circumstances. Individual differences in mean RTignore - RTgo could reflect both automatic interference and endogenous pausing (both of which would increase the difference between ignore and go trials), as well as subsequent failing to go on ignore trials (omissions, which would decrease differences by removing longer latency responses from ignore distributions just as in stop signal distributions). Although some of this can be described as strategic, some won't be, and we therefore refrain from inferring strategy based on mean RTignore – RTgo.

> *Related, the distributions of RT on stop trials, particularly for saccade responses, are portrayed with a second mode in the schematic illustrations and clearly peaking at SSRT in Figure S1. This second mode is not observed in other saccade stop signal studies. This indicates that the participants in this study were in a peculiar mode of performance.*

The second mode indicates that participants occasionally ignore the stop signal, which is why it peaks at the same latency as the rebound for the ignore distribution. These are not unusual, in particular in selective stopping designs, but are not as easily seen on cumulative functions, which are the standard way of plotting the results in this field.

> *Finally, given the pivotal role of measures of differences of RT distributions and the pronounced variation of stopping accuracy (Figure S3), the authors must show the distributions for all of their new participants. The authors' claim to higher resolution obliges them to reveal every step of analysis.*

We will save figures showing the individual distributions in the OSF folder. Note that these figures can be produced by running the code we shared, so each step is already fully transparent, but we will create tidy versions that also highlight manually corrected indices.

> *In its current form, this manuscript is unlikely to change the thinking of modelers or practitioners of the stop signal task.*

We thank the reviewer for their in-depth and thoughtful comments and suggestions, so that the paper can be revised to address the concerns.

Reviewer #3 (Public review):

Summary:

Statham and colleagues test an assumption underpinning a very large literature: that the stop-signal reaction time (SSRT) indexes the speed or efficacy of top-down inhibitory control. They argue instead, and support their claims with a total of eight datasets, that SSRT is substantially occupied by visuomotor deadtime (i.e., incompressible sensory and motor delays common to all visually guided responses), which varies across individuals, conditions and populations in ways that mimic effects usually attributed to inhibitory control. They propose two remedies: subtracting an independent estimate of visuomotor deadtime (T_0) from SSRT, and a new index, the selective stopping delay (ΔT), from the stimulus-selective stopping task.

Strengths:

The paper's principal strength is the combination of these components. That SSRT must contain peripheral delays is not itself new, as the authors point out (Boucher et al., 2007; Salinas and Stanford, 2013; Bompas et al., 2020). What is new is the quantification of the problem at scale, across seven archival datasets and a preregistered replication, together with the demonstration that T_0 can be recovered from existing stop-task data. That is important, as it provides a diagnostic that can be applied to data already collected. The authors' offer to assist others in doing so is exemplary. The supplementary analyses of trial numbers and participant pooling are very useful, and the paper provides important sanity checks, notably confirming that stop and ignore signals produce indistinguishable initial interference before pooling them.

Weaknesses:

The evidence for the central claim is strong but presented in a way that overstates it. Figure 2 reports 85% and 80% shared variance between SSRT and T_0, but these pool across datasets and, more critically, across response modality: manual and saccadic estimates from the same participants are plotted together with a single regression line through both. Because manual and saccadic deadtimes differ by roughly 130 ms, the resulting correlation largely reflects a between-condition difference rather than covariation among individuals. The numbers that speak to individual differences are more modest (40% for manual responses; 7% for saccades). The manual result is convincing and consequential; the saccadic result is not, and the explanation in terms of restricted range, while plausible, is offered after the fact and is directly testable by reporting the reliability of saccadic T_0 or correcting the correlation for attenuation. This limitation is arguably good news for the paper's practical message, since it implies saccadic measures are relatively protected, but the manuscript should make clear (including in the abstract) that the strong individual-differences case rests on the manual data, where motor execution delay is the main driver.

We will add separate regression lines and R-values for manual and saccadic modalities on Fig.2A and an inset showing the variance only driven by individual differences across all archival data (i.e. z-scored per condition and datasets, $R(215)=0.43$, $p<0.001$). We will also state more explicitly that the overall correlation may not be the relevant one for researchers specifically interested in individual differences.

While a large portion of SSRT literature is about individual differences, there are also many studies about differences between conditions, including comparing different response modalities. Therefore, it is a general question whether differences of any kind in SSRT reflect differences in inhibitory control or differences in sensory-motor delays. At a conceptual level, most users of the SSRT are intending to measure control, and have a conceptual model in

which control is separate from the modality of response or the exact characteristics of stimulus delivery. Thus, it is important to point out that their measure of control is very much dependent on these things, and in fact to a much larger degree than the more subtle individual differences, group differences or conditions of interest.

We agree that repeatability is critical for interpreting null results and will provide split-half repeatability for all our indices, including T0 and SSRT, and use these to correct their correlations. We thank the reviewer for this suggestion.

> *A related point concerns interpretation rather than analysis. Since SSRT is, on the authors' own account, approximately the sum of T_0 and a decision-related component, covariation between the two is expected on structural grounds; the preregistered correlation with reaction times from separate speeded blocks mitigates this, but the finding is less surprising than its current framing implies. What would determine whether past conclusions must be revised is not whether SSRT correlates with T_0 across individuals, but whether the decision-related component tracks the independent variable in any given study. The alcohol reanalysis could be a test case for this: the authors show that alcohol raises T_0 commensurately with SSRT and conclude the effects are "consistent with these effects being fully driven by visuomotor delays," yet (unless I missed something) they do not report the corrected measure for these data, while they do so for signal contrast and response modality. Running that analysis, and stating plainly what Campbell et al. (2017) would have concluded under the proposed treatment, would be an important demonstration.*

We fully agree that the presence of a correlation is indeed entirely expected and obvious in our own account, as conveyed early on in the manuscript (“From Fig. 1D, it seems clear that SSRT and T_0 are inevitably connected”). We agree the main question is what is left for SSRT to explain. We will update our analysis of the Campbell et al. (2017) alcohol study as suggested. Future work can then focus on other “independent variables”.

> *The case for ΔT is the least developed part of the paper. ΔT is a difference between two independently estimated, individually noisy quantities, extracted by a non-trivial procedure (see also below), and no reliability estimates are reported for T_0, TS or ΔT. This would be possible based on the two-session design (and the group has prior work on the reliability of cognitive control measures). This matters because the argument that ΔT is superior rests, to some extent, on null findings: ΔT does not correlate with SSRT, with stopping accuracy, or with the differential response to stop and ignore trials. These null correlations are interpreted as freedom from confounds, but an unreliable measure would produce the same pattern, and the seven participants with implausible negative ΔT values indicate that noise is not negligible.*

We fully agree with all this and will update the wording surrounding the lack of correlation between ΔT and the other measures in light of its repeatability

> *In addition, the subjective correction of dip onsets ("Departure points were visually inspected and adjusted if it was deemed that the algorithm had placed them in inappropriate places"), which is critical to the paper's central measurement, should be blinded to condition or show inter-rater agreement. Since T_0 and TS are compared across conditions and ΔT is their difference, this introduces researcher degrees of freedom.*

All divergence times (T0, T0,stop, T0,ignore and TS) were confirmed by two of the authors. Each index for each modality is plotted on a separate figure (showing all individuals). It is technically easy to compare, say, T0 and TS, for one individual, but we refrained from doing this (and indeed ended up with many T0 > TS). We agree that blinding and inter-rater reliability are important safeguards and that our current analysis fell short of this. We will

explore whether this can be done retrospectively, and report on this exercise alongside guidance on criteria used for manual corrections.

> *This is particularly critical when a dip is not easy to extract. Figure 3 depicts an idealized ignore-trial distribution with a clean, deep dip. Real distributions are unlikely to look like this, and dip depth should depend on the behavioral relevance and salience of the ignored event; published work on rapid manual inhibition indicates that dips to behaviorally irrelevant events can be very shallow. Since ΔT is extractable only where the dip is resolvable, the generality of the method can be questioned. Ideally, the empirical distributions underlying every dataset analyzed should be shown to alleviate this concern.*

We will make figures available in the OSF folder with individual distributions that supported the extraction of each index, flagging those that got manually corrected. This will make apparent that the vast majority of dips were very clear, for both T0 and TS. Unclear dips led to missing indices, and were therefore excluded from our hypothesis testing.

> *One uncontrolled procedural difference also deserves comment. Corrective feedback about stopping too often or stopping too rarely was given after manual blocks only; saccadic blocks received none, and fixed rather than staircased delays were used throughout. Since the manual-saccadic contrast carries much of the argument, and saccadic blocks yielded both lower stopping accuracy (36% vs 52%) and far more exclusions (8 vs 1 of 37), this asymmetry offers an alternative to the interpretation that saccades are simply harder to inhibit.*

Indeed, blockwise feedback would have been hard to implement reliably for saccades. As manual and saccadic blocks were interleaved, our hope was that participants could use the feedback received for manual to adjust their strategy for both modalities. We will check how often the feedback was triggered for manual blocks and, if more than negligible, we will note the reviewer's suggestion as a possible driver for modality differences.

> *A final point concerns the comparison between response modalities. Raw SSRT suggests that saccadic inhibition is faster than manual (174 vs 266 ms), while both proposed corrections reverse this, with SSRT−T_0 and ΔT each indicating that saccadic inhibition is slower (the latter consistently across nearly every participant). This is one of the clearest illustrations of the paper's thesis, but it is not taken up in the discussion, which returns to modality only to note that saccadic T_0 varies little (the one reference to variation across action modalities appears in the modeling section, without stating its direction). It would also benefit from a caveat. Both corrected measures subtract the same T_0, and manual and saccadic T_0 differ by roughly 130 ms, so the two do not corroborate one another independently (TS is itself longer for manual responses, and yields a shorter ΔT only once the larger manual T_0 is removed). The accuracy of the subtraction therefore matters here: if the manual regression slope of 0.75 reflects sub-additivity rather than attenuation, subtracting the full T_0 would overcorrect manual responses more than saccadic ones, and could produce the reversal on its own.*

We agree with all this. We will use the repeatability of SSRT and T0 to disattenuate the slopes and consider alternatives to subtraction to correct for visuo-motor deadtime. Before we can elaborate on the modality effect on the speed of inhibition, we need to simulate the effect of motor variability on T0, TS and SSRT. If motor noise affects TS or SSRT more or less than it affects T0, this will affect our conclusions. We will explore this issue in the revision and report any analyses that bear on the robustness of the modality effect.

> *These concerns qualify rather than undermine the contribution. The core observation is robust, the diagnostic is practical and immediately applicable, and the case that a large body of work requires re-examination is well made. If the corrected analyses are carried*

> *through on the datasets already in hand, this will be an important paper for anyone who uses the stop-signal task.*

We thank the reviewer for their in-depth and thoughtful comments and suggestions

https://doi.org/10.7554/eLife.111819.1.sa0