Behavior Research Methods (2025) 57:15
https://doi.org/10.3758/s13428-024-02516-y

ORIGINAL MANUSCRIPT

The forced-response method: A new chronometric approach to measure conflict processing

Taraz G. Lee[1] · Jacob Sellers[1] · John Jonides[1] · Han Zhang[1]

Accepted: 14 October 2024

Abstract

Despite long-standing concerns about the use of free reaction times (RTs) in cognitive psychology, they remain a prevalent measure of conflict resolution. This report presents the forced-response method as a fresh approach to examine speed–accuracy trade-off functions (SATs) in conflict tasks. The method involves fixing the overall response time, varying the onset of stimuli, and observing response expression. We applied this method to an arrow flanker task. By systematically varying the time between stimulus onset and response, we reveal a comprehensive time course of the flanker interference effect that is rarely observed in previous literature. We further show that influential manipulations observed in free-RT paradigms similarly affect accuracy within the forced-response technique, suggesting that the forced-response method retains the core cognitive processing characteristics of traditional free-RT conflict tasks. As a behavioral method that examines the time course of cognitive processing, the forced-response method provides a novel and more nuanced look into the dynamics of conflict resolution.

Keywords Conflict resolution · Response conflict · Conflict processing · Speed–accuracy trade-off · Response signal

Imagine visiting a country where traffic flows on the opposite side of the road compared to the United States (e.g., the UK) Most American drivers who visit such a country struggle to adapt to this change. Or imagine that you are on a diet, but you smell the delicious flavor of fried chicken from a restaurant and feel compelled to stop in for a bite. These scenarios exemplify situations in which well-learned and nearly automatic responses conflict with goal-directed responses. The question arises: How do we overcome such habits to produce responses appropriate to a goal at hand?

For decades, cognitive psychologists have used so-called conflict tasks to model response conflicts of this sort. One widely used such task is the flanker task (Eriksen & Eriksen, 1974). In one popular version of the flanker task, participants are asked to indicate the direction of a middle arrow when flanking arrows could either be congruent (e.g., → → → → →) or incongruent (e.g., ← ← → ← ←) in direction compared to the middle arrow. Typically, participants can respond to the stimulus whenever they are ready to respond on each trial[1]. The existence and efficacy of conflict resolution is typically indexed by the difference in free response times (RTs) between congruent and incongruent conditions.

The flanker task is but one of many tasks in which a prepotent response conflicts with a goal-oriented response. Two other popular conflict tasks are the Stroop task (Stroop, 1935) and the Simon task (Simon, 1969). The congruency effect, as indexed by the difference in mean response time between congruent and incongruent trials in these tasks, has been the driving force behind the development of theories and models of conflict resolution over the years (Botvinick et al., 2001; Cohen et al., 1992; Egner, 2007; Gratton et al., 1992; Lu & Proctor, 1995).

However, there have been long-standing concerns about using free RT to index cognitive processes (Cronbach & Furby, 1970; Draheim et al., 2019; Edwards, 2001; Friedman & Miyake, 2004; Hedge et al., 2018, 2020; Miller & Ulrich, 2013). Despite these concerns, the practice of using free RTs to measure conflict resolution persists, with a recent survey finding that 84% of the difference scores in conflict tasks were based on mean RTs differences (von Bastian et al.,

[1] In some cases, a response deadline may be imposed, but it is rather lenient in order to time out extremely long responses. We still consider these tasks as "free RT" tasks because participants have substantial freedom to decide when to respond.

✉ Taraz G. Lee
tarazlee@umich.edu

[1] Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA

Published online: 12 December 2024

2020). The sheer reliance on mean RT differences, coupled with the issues associated with free RTs (discussed more below), calls for the development of alternative approaches to study conflict resolution.

In this paper, we revisit several issues having to do with using free RT to study conflict resolution. Then, we introduce a "forced-response" method as a new approach to studying conflict resolution that is complementary to the traditional free-RT-based approach. We will discuss how the forced-response method overcomes many issues with using free RT as a dependent measure of conflict resolution. We will also demonstrate that the forced-response method provides novel insights into the dynamics of conflict resolution beyond existing methods.

Free-RT-based approach

Since at least the 1960s, cognitive researchers have appreciated many of the shortcomings of free RT as a dependent measure, and several alternative methods have been adopted which allow for the explicit examination of the time course of cognitive processing. A complete review of these issues is beyond the scope of this paper and can be found elsewhere (Dosher, 1976; Draheim et al., 2019; Heitz, 2014; Wickelgren, 1977). However, we do point out several issues relevant to the current paper.

First, free-RT difference scores (and free RTs in general) do not consider the speed–accuracy trade-off, which differs from person to person and task to task (Heitz, 2014). Participants who are fast but error-prone may appear better than participants who are slower but more careful when only free RT is considered. One might think that this is resolvable by somehow combining RT and accuracy into a single composite measure of performance. Several composite speed–accuracy measures have been proposed in the literature with mixed success (Liesefeld & Janczyk, 2019; Stafford et al., 2020; Vandierendonck, 2017, 2021). Although these methods have been useful in some applications, they often make assumptions that are at odds with the trade-off between speed and accuracy (Heitz, 2014; Stafford et al., 2020). Most importantly, there is no theoretically justified basis for adjusting free RTs based on how accurate a person is (Pachella, 1974).

Second, RT difference scores are confounded with general processing speed in that they are proportionally larger for overall slower individuals than for overall faster individuals. Although applying a data transformation, such as using the logarithm of each RT, can mitigate this concern somewhat, this can reduce power and can eliminate significant effects that would be observable using the raw RT (Schramm & Rouder, 2019; Whelan, 2008). This makes it difficult to compare conflict resolution across populations that might differ in general processing speed.

Yet another problem has to do with whether RT itself is a pure measure of the cognitive processing required for a task. One study found that when participants were *forced* to respond earlier than they would normally, they could produce accurate responses with much less response time than suggested by their free RT (Haith et al., 2016). This indicates that RT is not an ideal measure of the time to complete processing of all component stages, as there is a substantial delay between this time point and the actual initiation of the response. Another recent study found that response times exhibit a use-dependent bias, meaning that people are biased to respond at times similar to their previous response times (Wong et al., 2017). This result again suggests that a person's response time reflects factors above and beyond the processing required to produce an accurate response. Overall, then, whereas RT is traditionally thought of as the duration of various information processing steps, an alternative view is that RT is itself an independent motor-control parameter. By this view, people decide *when* to respond as well as *what* to respond.

Speed–accuracy trade-off approaches

Many cognitive psychologists have argued for the superiority of methods that specifically examine speed–accuracy trade-off functions (SATs) as a tool to better understand cognitive processes of interest (Dosher, 1976; Heitz, 2014; Wickelgren, 1977). Though dwarfed by the volume of research examining mean RT, the explicit examination of SATs has allowed researchers to advance our understanding of the dynamics of conflict resolution since at least the 1990s (Gratton et al., 1992; Heitz & Engle, 2007; Ridderinkhof, 2002).

Conditional accuracy functions

One frequently used approach is to simply partition data from free RT tasks into a number of different time bins (e.g., 200–300 ms, 301–400 ms, 401–500 ms, etc.) and then calculate an accuracy rate in each time bin. These are sometimes called conditional accuracy functions (CAFs). With this binning method, researchers can examine the speed–accuracy trade off inherent in the responses participants make. This approach has been used to attempt to examine the time course of interference resolution (Gratton et al., 1992; Heitz & Engle, 2007; Manohar et al., 2015).

There are several shortcomings of this approach, though. Binning RT leaves the sampling within each bin up to chance, which often leads to certain time bins being less well represented for a particular individual than others. A consequence of this is that some participants may provide a good deal of data for a particular bin whereas others may leave that bin underrepresented. Also, if the true SAT function varies randomly trial-to-trial, accuracy will be overestimated at fast RTs and underestimated at slow RTs. Unfortunately, it is the short reaction times that are often the most important when examining conflict resolution. This is because researchers are interested in investigating the latency at which fast, incorrect responses driven by prepotent or automatic cognitive processing give way to slower, more goal-directed responses. Additionally, as noted above, using free RT here still gives participants strategic control over when to respond and may not reflect the time course of the underlying cognitive processing required to emit a response.

Response deadlines and/or payoffs

Another class of methods that has been used to empirically obtain SATs has been to use response deadlines and/or payoffs (Fitts, 1966; Pachella & Pew, 1968). Participants can be instructed to respond before some time deadline following stimulus presentation. This time deadline can be varied from one block of trials to the next and participants can be given feedback to ensure that they produce enough responses that fall below each deadline. An experimental manipulation that is somewhat similar is to give participants monetary payoffs for speed vs. accuracy. Correct responses can be rewarded, errors can be punished by taking away accumulated rewards, and these values can be proportional to RT. For example, participants can be paid $P - k \times \text{RT}$ for each correct response and punished $-k \times \text{RT}$ for incorrect responses. P and k can be varied in different blocks of trials to induce participants to respond at different speeds and with different accuracy rates (Swensson & Edwards, 1971). Although these methods allow researchers to plot SATs, they also leave a substantial opportunity for participants to exert strategic control over how they are responding. Especially when payoffs and deadlines are blocked, participants may choose to adopt different strategies that vary systematically over the different conditions.

Response signals

To deal with some of the drawbacks in these SAT methods, some researchers have adopted what is often referred to as "the response-signal paradigm" (Reed, 1973). In a typical response-signal task, a stimulus is presented and is then followed at some short time lag by a signal that tells the research participant to initiate a response. This procedure allows the researcher to use the timing of the response signal as an independent variable and allows an examination of accuracy as the dependent measure. The timing of this response signal is usually varied randomly from trial to trial from very short times at which performance is at chance to long times at which performance reaches an asymptote.

Although this procedure may seem similar to tasks which employ a time deadline with a fixed lag following the stimulus, it is much more difficult for subjects to adjust their strategy as a function of the given processing time in that they are not informed of the timing before the trial begins. It is often assumed that a subject's strategy does not vary in any controlled way. By varying the timing of the response signal relative to the timing of stimulus presentation from trial to trial, researchers can be more certain that a subject is in the same state following stimulus presentation at any given time across all trials up until the timing of the response cue.

The response-signal paradigm has been fruitful in examining a variety of cognitive processes in several domains including psycholinguistics, memory, and decision-making in perceptual tasks (e.g., McElree & Dosher, 1989; Ratcliff, 2006; Ratcliff & McKoon, 1982). This has been especially true in investigations of component processes, like those in conflict tasks, that differ with respect to their underlying time courses and may lead to competing responses. For example, long-term memory researchers have suggested that performance on recognition memory tasks is driven by two processes with different latencies: familiarity and recollection (see Yonelinas, 2002 for a comprehensive review). Using a response-signal paradigm in concert with an item recognition task, McElree and Dosher (1989) showed that lures that were members of a previous, no longer relevant study list led to increased false alarm rates early in the time course of processing. When more ample time was given to make a response, these false alarms were eliminated. This finding was interpreted as familiarity initially driving responses, which is later corrected by the retrieval of items on the current study list. McElree (1998) later extended this finding to show that both episodic familiarity and semantic similarity can drive early false alarm rates using the response signal procedure. However, to our knowledge, the response-signal paradigm has seldom been used to examine the conflict resolution process by applying it in conjunction with conflict tasks commonly used in the cognitive control literature (but see Hilchey et al., 2011 and Teichert et al., 2014).

A path forward

We believe a path forward in the study of conflict resolution is to move from RT as a dependent variable and instead treat it as an independent variable by adopting a significant variant of the response-signal paradigm which we call the "forced-response method." A version of this method was one of the first response-signal paradigms ever reported (Schouten & Bekker, 1967), but it has never been widely adopted, likely due to its difficulty in administration prior to the use of modern computer-based stimulus presentation.

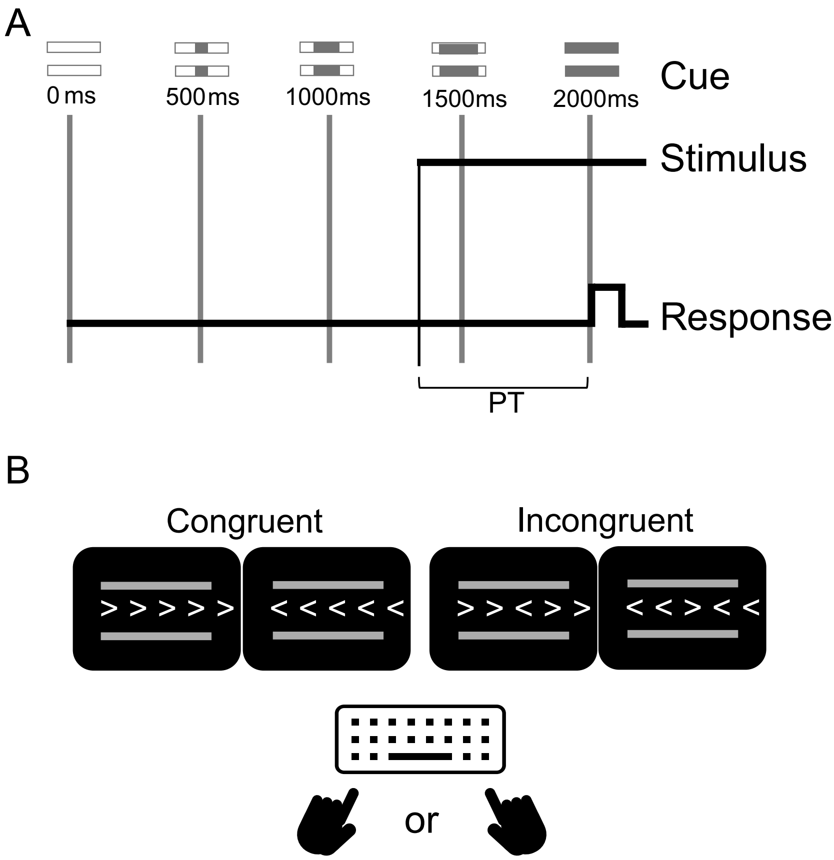

Fig. 1 A schematic illustration of the forced-response flanker task. **A** shows the forced-response procedure. As a trial began, white ink filled 25% of each rectangle's inner space every 500 ms, resulting in a fully filled rectangle at 2000 ms. Participants were asked to respond exactly when the two rectangles were completely filled, namely at 2000 ms. The imperative stimulus appears unpredictably during the last 1000 ms of the countdown. **B** shows the task display at 2000 ms, at which participants must press the left or the right key to respond to the direction of the middle arrow. Note that although this figure shows an example where a response is made exactly at the time of the "go" signal, in all analyses PT is defined as the time from stimulus presentation to when a response was actually made (i.e., usually not at the exact time of the "go" signal)

In the balance of this article, we outline the forced-response method and note its specific advantages for studying conflict resolution. To allay fears that adopting a forced-response method might alter the nature of the processing in conflict tasks, we additionally present evidence showing that using this method to study conflict preserves several of the classic effects observed when using the free-RT methodology.

Methods

The forced-response method involves fixing overall response time from trial to trial, varying the onset of stimuli prior to a demanded response, and observing response expression. In this way, the forced-response method differs importantly from the response-signal paradigm. As we comment above, with a typical response-signal paradigm there is constant uncertainty in when a response will be required. With the forced-response method, there is complete certainty as illustrated with the procedure shown in Fig. 1. Participants receive *N* equally spaced signals (e.g., an auditory beep or a visual cue presented every 500 ms) and are instructed to respond exactly at the onset of the final "go" signal. Processing time (PT) is manipulated by varying the time from stimulus onset to the "go" signal, but response time is fixed on each trial by virtue of participants having to respond at the time of the "go" signal. So, the timing of the demanded response is always predictable on each trial and from trial to trial; however, which particular response is required is unknown until the stimulus appears. Response accuracy at each PT is the dependent measure. Next, we describe in detail an application of the forced-response method to an arrow flanker task.

Forced-response flanker task

Participants We recruited 137 participants from Prolific.co to participate in this study. Participants were compensated at a rate of $10 per hour for participating. Participants completed the forced-response flanker task (described in detail below) plus several questionnaires which are not relevant to the presentation of the method that is the topic of this paper. The duration for completing the forced-response task alone was about 35 minutes. Data from 11 participants were discarded due to poor timing accuracy in producing a response at the "go" signal (< 30% of trials ±100 ms), leaving a total of 126 participants for analysis (mean age = 36.9, SD age = 12.4, 64.3% female).

Task and experimental timeline All experiments were programmed using PsychoPy and were run online using Pavlovia (Peirce et al., 2019). A demo of the experiment is available online: https://run.pavlovia.org/hanzh/forced-response-flanker-demo.

The task consisted of two practice phases and one experimental phase. In the first practice phase, participants trained for 40 trials on a free-response flanker task. On each trial, five arrows appeared on the screen and participants were asked to indicate the direction of the center arrow while ignoring the direction of the flanking arrows. They were instructed to press the "W" key (for left) or the "P" key (for right). There were two types of trials that appeared equally often: congruent trials (e.g., <<<<<<) and incongruent trials (e.g., <<><<). Each arrow had a size of (.06, .06), with the unit being a percentage of screen height. On each trial, two outlined rectangles that had the same width as the row of arrows were also shown on the screen, one above and one below the arrows, as illustrated in the example presented in Fig. 1B. These outlined rectangles served as response signals in later phases of the task. The arrows appeared in between the two outlined rectangles as shown in the figure. The arrows remained on the screen until the response expression. Once a response was made, participants received feedback about their response accuracy. The feedback message was displayed for 400 ms.

In the second practice phase, participants were given 30 trials of training to learn to produce a response at the time of a "go" signal. In this phase, there were no arrows shown on the screen. Instead, on each trial the two outlined rectangles were incrementally filled with white ink (see Fig. 1A). As a trial began, white ink filled 25% of each rectangle's inner space every 500 ms, resulting in a fully filled rectangle at 2000 ms. The filling began at the center of each rectangle and expanded to the edges, and the rectangles were filled synchronously. The rectangles were removed from the screen after 2100 ms (i.e., 100 ms after the "go" signal). Participants were instructed to respond exactly when both rectangles were filled with white ink, (i.e., 2000 ms). Participants were told that they could respond with either the W key or the P key, but they were encouraged to practice timing with both keys. After each trial, participants were given feedback for 1000 ms about whether they responded too quickly (< 1900 ms from the start of the trial), too slowly (> 2100 ms), or with perfect timing.

The experimental phase was a combination of the two practice phases (see Fig. 1A). Participants performed 10 blocks of 48 trials of the flanker task with forced-response timing and with stimulus presentation time that was uniformly varied prior to the "go" signal. As in the first practice phase, participants were asked to respond to the direction of the center arrow using the "W" and "P" keys. As in the second practice phase, participants were asked to respond exactly when the two rectangles were filled with white ink (i.e., 2000 ms). The onset time of the arrows was selected with replacement from a uniform distribution between 1000 ms and 2000 ms in increments of 20 ms, with the value on each trial unpredictable. In other words, the arrows would appear at a random time within one second before the "go" signal. This approach allowed us to measure the accuracy of responses when the exact amount of time allowed for stimulus processing and response preparation was controlled. The entire display (rectangles and arrows) was removed 100 ms after the "go" signal. After each trial, participants received feedback about their timing accuracy, as in the second practice phase. After the feedback, there was an intertrial interval of 1000 ms before the next trial began. Participants were not given feedback on their response accuracy in an effort to reduce sequential effects and changes in strategy due to feedback. For example, it is well known that people tend to slow down responses following errors. Delays in emitting a response are particularly problematic for our forced-response studies as there is a timing criterion (±100 ms) and we are specifically interested in examining the time course of cognitive processing (see Adkins et al., 2024 for a discussion of post-error effects in forced response paradigms).

Analysis We define processing time (PT) as PT = RT − t, where RT is the time of the response relative to the start of the trial and t is the time of stimulus onset. For example, if on a given trial the stimulus appeared at 1500 ms from the start of the trial, and the response was made at 2050 ms, PT is then calculated as 2050 − 1500 = 550 ms. This calculation method accounts for the fact that participants are not always accurate in responding exactly at the "go" signal, and it reflects the actual time it took for participants to generate a response. Note that a negative PT occurred when participants responded before the stimulus appeared. We removed trials if PT was smaller than 0 ms or greater than 1000 ms (3.99% of trials).

The data produced by the forced-response method consist of response accuracies and their corresponding processing times. This type of data can be analyzed in many different ways at either the group or individual level, including binning by PT, sliding window analyses, and/or fitting computational models (e.g., Adkins et al., 2024; Adkins & Lee, 2023; Hardwick et al., 2019). Here we present analyses using SMART, a state-of-the-art technique for reconstructing the time course from one-sample-per-trial data, which also permits statistical analysis of the accuracies (van Leeuwen et al., 2019). Simply put, the method involves first smoothing the data of each participant and then using a weighted average to construct a group-average time course. Furthermore, a cluster-based permutation test can be used to examine differences between time courses or from a baseline. Python scripts for implementing the SMART method can be found at the project's OSF website: https://osf.io/qa2uc/.

Results

The relationship between response accuracy and PT for congruent and incongruent trials, as analyzed using the SMART method, is shown in Fig. 2. Several features in these data are worth highlighting. First, when PT is short (before approximately 200 ms), response curves are no different from chance (50%). This suggests that on these trials, participants were simply guessing as there was little time for any meaningful processing of the stimulus to take place.

The most intriguing feature of this plot appears at intermediate levels of PT. On congruent trials, response accuracy rose monotonically from chance as PT became longer. By contrast, on incongruent trials, response accuracy initially *decreased* from chance as PT increased, followed by a recovery in response accuracy at longer PTs. A cluster-based permutation test shows that response accuracy for incongruent trials was significantly below 50% from 236 to 415 ms, with a nadir at 367 ms. In essence, for incongruent trials, there exists a period when accuracy worsened despite longer processing times. This time window reflects the influence of the prepotent, automatic response that eventually gives way to the goal-directed response when more time is allowed for processing.

Using a cluster-based permutation test, we also determined that the two curves diverged significantly starting at 227 ms. If we define the flanker interference effect as the difference in accuracy between congruent and incongruent trials, then this point of divergence represents the earliest time point when the flanker interference effect emerges. At 391 ms, the two curves reached their maximum divergence with a difference in accuracy rate of 44.9%. Following the same definition, this point represents when the flanker interference effect was the strongest.

To illustrate the advantage of the forced-response method relative to the free-response method, let us compare the

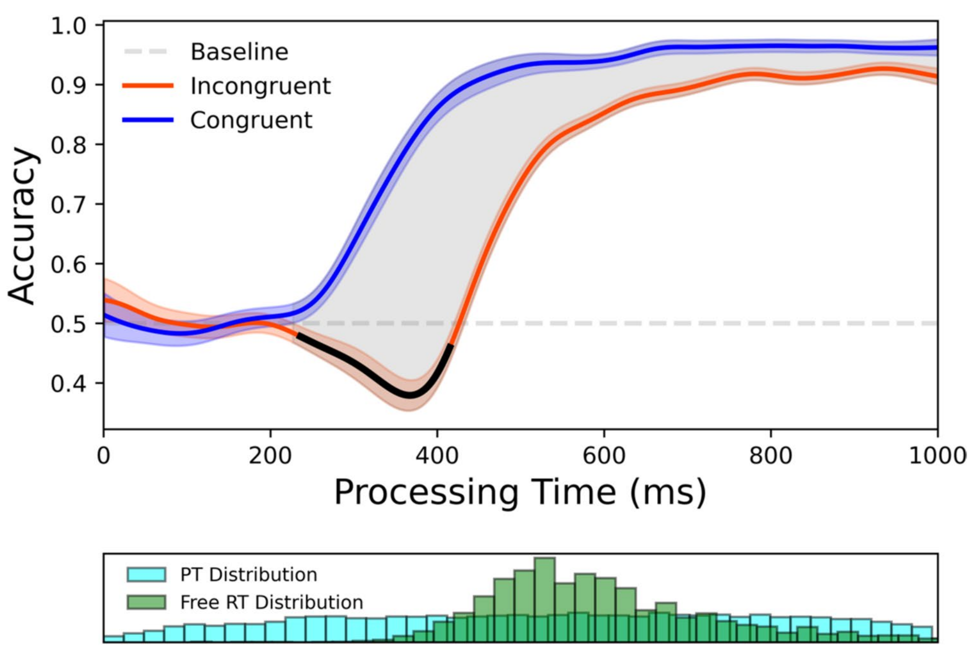

Fig. 2 Response accuracy as a function of PT. The relationship between processing time (PT) and response accuracy in the forced-response flanker task. In the top panel, the shaded area between the two curves indicates the period of significant divergence in accuracies for congruent and incongruent trials. The black segment on the incongruent curve represents the period when accuracy for incongruent trials was significantly below 50%. Confidence bands indicate a 95% confidence level. In the bottom panel, the teal distribution represents PTs. For comparison, the green distribution shows RTs in the incongruent condition for the same participants completing a free-RT practice block

distributions of PTs and free RTs. In the bottom panel, the teal distribution represents the distribution of PTs in this task, which is roughly uniformly distributed by design. For comparison, the green distribution shows free RT in the incongruent condition for the same participants completing a free-RT practice block. The bulk of the distribution of free RT is located after the largest interference effect, as revealed by the forced-response analysis. Indeed, the maximally divergent point (391 ms) has a percentile rank of only 2.28% in this free-RT distribution. In other words, the free-RT distribution missed the bulk of the time period when the flanker effect was the strongest according to our forced response paradigm. But the forced-response method by its very design captures this effect and samples this time point relatively densely.

Finally, response accuracy on incongruent trials never recovered to the same level as that on congruent trials. For incongruent trials, response accuracy appears to have plateaued after roughly 750 ms of PT, even though there is still "room for improvement" (accuracy at PT = 750 ms was 91.2 % for incongruent trials and 96.4% for congruent trials). This suggests a lingering effect of response conflict despite sufficient time to process the stimuli. Of course, if we had extended PT beyond 1000 ms, this difference between congruent and incongruent trials may have been reduced. Indeed, consider the thought experiment of giving participants as long as 10 sec of processing time. By that time, it is likely that accuracy for the congruent and incongruent trials would be nearly if not entirely identical.

To summarize, the forced response method offers a more nuanced understanding of response conflict dynamics than the traditional free-RT method. By systematically varying the time between stimulus onset and response, this approach uncovers the complete time course of the flanker interference effect, from its early emergence to its peak, and its lingering presence even with extended processing time.

Next, we assess several aspects of the forced-response paradigm that speak to the validity of the method.

Timing accuracy

One might wonder to what extent participants can be forced to respond at a specific time. Figure 3 shows the distribution of when participants emit their response relative to the start of each trial in this task. These responses were generally centered around the "go" signal (2000 ms), with a mean of 2017 (Fig. 3A). Of course, it is virtually impossible to force a response at precisely the "go" signal. There was variance around this value with a standard deviation of 128 ms. Nevertheless, we consider this standard deviation to be quite small relative to that of a canonical free-RT distribution. Overall, 67.68% of the trials were judged

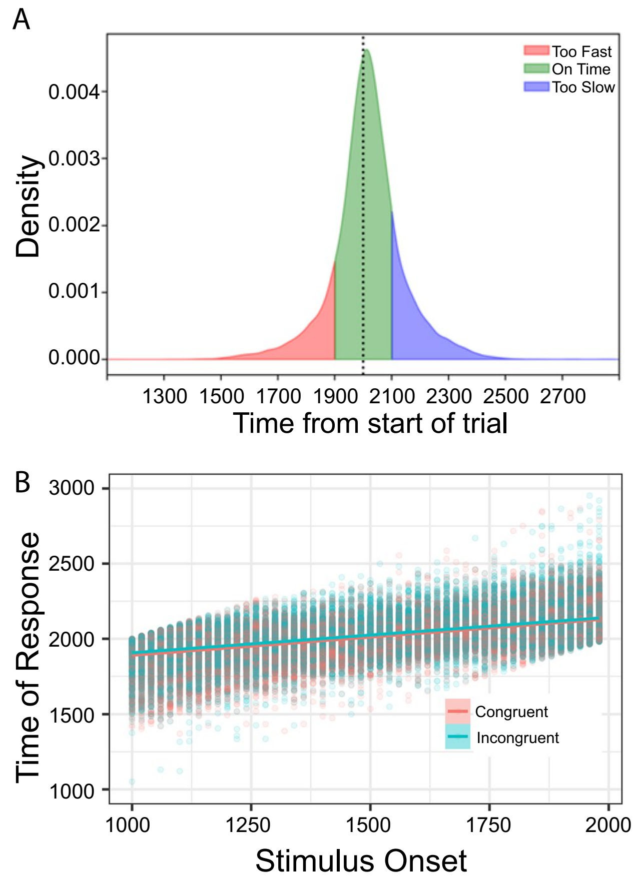

Fig. 3 Distribution of the timing of responses in the forced-response flanker task and the relationship with stimulus onset. **A** The green region indicates responses judged as "on time" (±100 ms). The red and blue regions indicate responses judged as "too fast" and "too slow" respectively. The dotted line at 2000 ms from the start of the trial represents the "go" signal when the rectangles are completely filled on the screen. **B** Earlier stimulus onset times corresponded with slightly faster responses. Although there is a small effect of congruency on the timing of responses, there was no significant difference between congruent and incongruent trials with respect to the influence of the time of stimulus onset on the timing of responses

as "on time," meaning that participants responded within an arbitrarily defined 100-ms margin of error. In contrast, 12.70% of the trials were "too fast," meaning that participants responded more than 100 ms prior to the "go" signal, whereas 19.62% of trials were "too slow," meaning that participants responded more than 100 ms after the "go" signal.

Recall that the calculation of PT is based on the time from stimulus onset to the actual time the response was emitted rather than when the "go" signal actually occurred (i.e., 2000 ms), which accounts for the variation in the distribution of responses around the timing goal. It is also important to emphasize that the decision to categorize responses as "fast," "slow," and "on time" was arbitrary. The purpose of this was simply to display a feedback message after each trial so that participants could adjust their response timing accordingly. Therefore, it may not be appropriate to simply discard "fast" and "slow" trials during data analysis. We recommend a sensitivity analysis to examine the potential impact of the timing criterion on the results. We provide an example of such an analysis in the next section.

We also examined the relationship between the time of stimulus onset and the timing of responses with respect to the "go" cue (see Fig. 3B). We fit a linear mixed model with the time of stimulus onset, the time of response, and their interaction term as fixed effects. Additionally, the model included random intercepts and random slopes for each combination of participant and congruency. We found that there was a significant positive relationship between stimulus onset and the timing of responses ($\beta = 0.239$, $t = 18.471$, $p < 0.001$). That is, when the stimulus was presented earlier in a trial, participants responded slightly faster relative to the "go" cue. However, we did not find a significant effect of congruency on the timing of responses ($\beta = 0.021$, $t = 0.791$, $p = 0.43$) nor did we find a significant interaction between congruency and stimulus onset time ($\beta = -0.005$, $t = -0.298$, $p = 0.77$). Given that we did not observe any difference between congruent and incongruent trials with respect to the influence of stimulus onset time, it is unlikely that participants were strategically delaying their responses in a way that would impact our measure of conflict resolution.

Sensitivity analysis of the timing criterion

We varied the timing criterion from ±25 to ±400 ms in increments of 5 ms. These criteria encompass a range from 22.8% to 98.8% of the entire distribution of responses depicted in Fig. 3. For each cutoff point, we constructed group-average time courses using the SMART method and calculated the nadir of the incongruent condition, referring to the point at which accuracy was lowest. The results are displayed in Fig. 4. As illustrated in the figure, the time courses based on different timing criteria are strikingly similar. The nadir distribution centers around 367 ms, which is the original value calculated based on "on-time" trials. These results reinforce the argument against simply discarding "too fast" or "too slow" trials: What truly matters is not whether participants were actually on time, but rather whether they were attempting to be on time.

Does the forced-response method change the nature of the task?

The forced-response method requires participants to resolve response conflict while at the same time monitoring when to emit their responses at the proper time. One might wonder whether this dual-task situation yields performance in

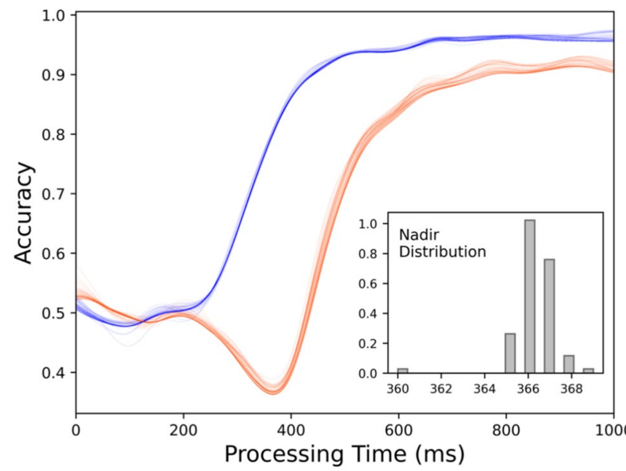

Fig. 4 The time course of response accuracies for different timing criteria. Each line represents a distinct timing criterion (ranging from ±25 to ±400 ms in increments of 5 ms). The inset panel displays the distribution of the nadir across various timing criteria

conflict tasks that is fundamentally different from performance that would be seen with free responding. To examine this issue, we selected three variables that are well known to modulate the congruency effect in free-response tasks, and we examined whether qualitatively similar effects can be observed using the forced-response task. The three variables are distractor salience, conflict frequency, and previous-trial congruency.

Distractor salience In the flanker task, a larger interference effect (i.e., poorer performance on incongruent vs. congruent trials) can be obtained by increasing the salience of the flanking arrows relative to the center arrow (Zeischka et al., 2011). We recruited 45 participants from Prolific.co to complete a forced-response flanker task with a distractor salience manipulation. Data from six participants were discarded due to poor performance (accuracy less than 50%, less than 30% of trials on time, or more than 15 trials in a row not being on time). Participants were randomly assigned to one of two conditions: low salience ($N = 20$) or high salience ($N = 19$). In the low distractor-salience condition, the flanking arrows were smaller than the center arrow (flanking arrows: (.04, .04), center arrow: (.08, .08), unit: percentage of screen height). In the high distractor salience condition, the flanking arrows were larger than the center arrow (flanking arrows: (.08, .08), center arrow: (.04, .04), unit: percentage of screen height). In addition, there were two minor differences from the task described above. First, the feedback message was displayed only when participants were too slow or too fast. Second, the intertrial interval randomly varied between 0 to 1 second. All other aspects of the task were the same as the original task.

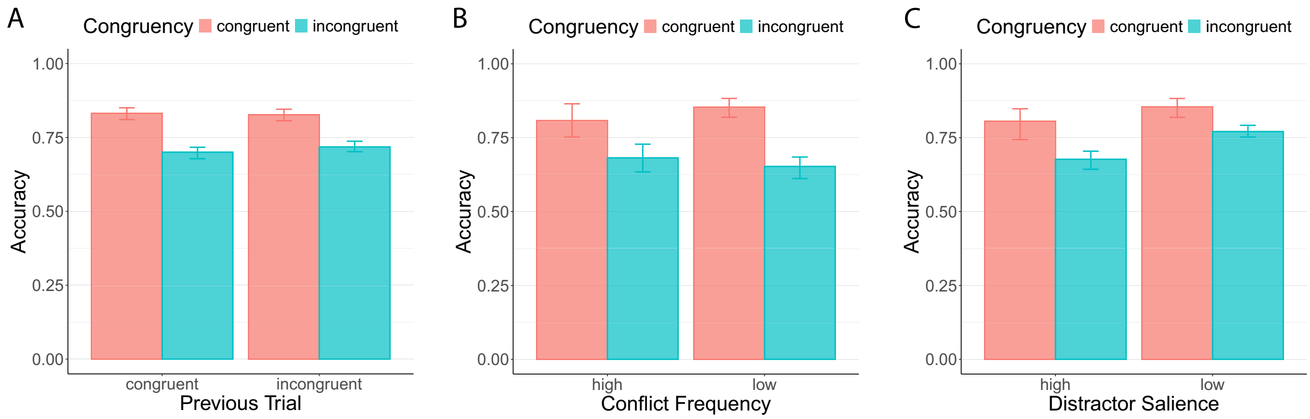

Fig. 5 Effects of three variables on performance under conflict. **A** Congruency sequence effect. The congruency effect is diminished following incongruent trials. **B** Conflict frequency effect. Low: 30% incongruent and 70% congruent. High: 70% incongruent and 30% congruent. **C** Distractor salience effect. High: flanking arrows bigger than the center arrow. Low: flanking arrows smaller than the center arrow. Columns and error bars reflect means and bootstrapped 95% confidence intervals, respectively

To facilitate comparison, we computed the average response accuracy for each condition, aggregating across all PTs (see Fig. 5A). The data were then analyzed using a 2 (trial type: congruent/incongruent) × 2 (distractor salience: high/low) analysis of variance (ANOVA). The results of the ANOVA show a statistically significant interaction between trial type and distractor salience ($F(1, 37) = 5.90$, $p = 0.020$, $\eta_p^2 = 0.14$). This was driven by a larger interference effect when distractor salience was high (0.045 difference in accuracy rate), a result that is similar to what the literature has shown for free response situations (Zeischka et al., 2011).

Conflict frequency Previous research has shown that increasing the proportion of incongruent trials leads to smaller interference effects (Logan, 1980). We recruited 61 participants from Prolific.co to complete a forced-response flanker task with a conflict frequency manipulation. Data from seven participants were removed due to poor performance (accuracy less than 50%, less than 30% of trials on time, or more than 15 trials in a row not being on time). Participants were randomly assigned to one of two conditions: low conflict frequency ($N = 36$) or high conflict frequency ($N = 18$). In the low conflict-frequency condition, congruent trials had a 70% chance of appearing on each trial and incongruent trials had a 30% chance. In the high conflict-frequency condition, congruent trials had a 30% chance of appearing, and incongruent trials had a 70% chance of appearing on each trial. All arrows had an equal size of (.06, .06). All other aspects of the task were the same as those of the distractor salience task.

We again computed the average response accuracy for each condition, aggregating across all PTs (see Fig. 5B). A 2 (trial type: congruent/incongruent) × 2 (conflict frequency: high/low) ANOVA show a statistically significant interaction between trial type and conflict frequency ($F(1,52) = 6.48$, $p = 0.014$, $\eta_p^2 = 0.11$). Again, this pattern of results mimics what the literature has documented for free response tasks (Logan, 1980): interference effects are smaller when conflict frequency is high (0.074 difference in accuracy rate).

Previous-trial congruency The effect of previous-trial congruency on the congruency effect, or the "congruency sequence effect" (CSE), is the observation that congruency effects are typically smaller following incongruent trials than following congruent trials (Gratton et al., 1992). For this variable, we simply reanalyzed the forced-response flanker data reported in the previous section with the addition of previous-trial congruency ($N = 126$).

We observed a CSE effect using the forced-response flanker task (see Fig. 5C). A 2 (current trial congruency: congruent/incongruent) × 2 (previous-trial congruency: congruent/incongruent) show a statistically significant interaction between current and previous-trial congruency ($F(1, 125) = 7.34$, $p = 0.008$, $\eta_p^2 = 0.06$). Once again, this mimics the results from free response experiments (Gratton et al., 1992): interference effects are smaller following incongruent trials (0.023 difference in accuracy rate).

In short, we have compared results from the forced-response method to previously documented results using free-response methods, and we found qualitatively similar outcomes in all cases: Greater salience of the distractors resulted in more interference; lower distractor frequency resulted in more interference; and the CSE effect was demonstrated. We take these data to indicate that in spite of the forced-response method requiring a "dual-task" on the part of participants, this method does not do injustice to the underlying interference effect in the flanker task.

Discussion

The forced-response method we advance here can be fruitfully used to explicitly examine speed–accuracy trade-off functions in a variety of conflict resolution tasks to gain greater insight into the mechanisms of cognitive control. We show that the effects of classic manipulations that affect behavior in free-RT paradigms are preserved when accuracy is examined using the forced-response technique. We found that increasing the salience of the distractor led to worse accuracy, especially on incongruent trials. We also observed the classic conflict-frequency effect wherein a high proportion of incongruent trials leads to a reduction in interference and relatively improved performance on incongruent trials. Finally, we observed a congruency-sequence effect whereby an incongruent trial leads to improvement in performance if the subsequent trial is also incongruent, but a worsening in performance if the subsequent trial is congruent. These findings give us confidence that the forced-response method does not fundamentally change the nature of cognitive processing seen in conflict tasks with free RT and allows researchers to empirically observe how the speed–accuracy trade-off function shifts with respect to independent variables of interest. Compared with free-RT tasks, the forced-response method is RT-irrelevant, accuracy-based, and provides a more nuanced look into the dynamics of conflict resolution.

The advantage of speed–accuracy trade-off methods

Arguments for the superiority of explicitly measuring speed–accuracy trade-offs relative to free-RT experiments are not new (e.g., Wickelgren, 1977). However, it seems that the force behind these arguments has diminished over time, or perhaps they have just been forgotten as the vast majority of experimental psychology experiments over the last several years have employed free-RT measures. In addition to some of the issues we raised in the introduction, the basic issue is that participants can trade their accuracy for speed such that *any* mean value of reaction time is possible for *any* task depending on what level of accuracy participants set as their criterion. Unless accuracy is carefully measured, mean RT is not a good indicator of task difficulty or the time it takes for the cognitive processes of interest. Often researchers seek to have participants maintain an extremely high level of accuracy so that RT differences are more interpretable. Unfortunately, if one simply observes the form of a speed–accuracy trade-off function, it is exactly at near ceiling levels of accuracy where one would expect to observe large variation in RT with negligible changes in accuracy (Pachella, 1974; Reed, 1973; Wickelgren, 1977). That is, very small changes in accuracy may lead to RT differences that might be larger than RT differences due to independent variables of interest. The forced-response method we advance here wrests control of the speed–accuracy criterion away from our participants back to the experimenter and obviates these concerns.

Forced-response versus other response signal paradigms

As we outline in the introduction, there are several different methods that can be used to empirically obtain speed–accuracy trade-off functions (e.g., response deadlines, response signals, etc.). While the forced-response method is quite similar to response signal methods used in previous literature, there are particular advantages and disadvantages to each approach. The response signal method used most often in previous work presents a stimulus at the start of each trial and varies the timing of a response signal (usually an auditory tone) that cues the participant to respond immediately. From trial to trial, there is uncertainty about *when* a response is to be emitted, but full certainty about when a stimulus will appear. Although lags in responding to the response signal in these paradigms are usually quite short (~ 200 ms), this may make it difficult to effectively sample the earliest state of cognitive processing following a stimulus. For example, one of the only prior efforts that has attempted to combine the response signal paradigm with a conflict task was unable to get participants to reliably respond within ~300 ms of stimulus presentation (Hilchey et al., 2011).

In the forced-response method, there is full certainty about when to respond on each trial, but there is uncertainty about when the stimulus will appear and what response will be required. This approach could make sampling short PTs more effective, but also has its own drawbacks. There could be a concern that stimuli presented early on in a trial (long PT) could be processed differently than those presented later on in a trial (short PT) as participants have more certainty about how much time they will have to process the stimulus and make a response as the trial unfolds. The forced-response method might also not be ideal for studying tasks that require relatively long amounts of processing time to reach asymptotic levels of accuracy. To keep the response cues (filling rectangles here) consistent across the experiment, each trial needs to be as long as the longest processing time the researcher would like to sample from. This could make a complete session take an infeasible amount of time. Nevertheless, both the standard response signal approach and the forced-response method yield the same type of data (the underlying time course of processing), and researchers should carefully consider whether one approach would be more appropriate to the tasks used in studying the cognitive processes of interest.

Application to other conflict and dual-process tasks

While the results presented here combine the forced-response task with classic interference resolution tasks in cognitive psychology, we believe that this and other response-signal methods can be fruitfully extended to any task that seeks to examine dual processes that have different latencies and competing responses. For example, the most frequent use of the response signal paradigm in the extant literature has been to examine memory processes (Benjamin & Bjork, 2000; Hintzman & Curran, 1994; McElree & Dosher, 1989; Öztekin & McElree, 2007, 2010; Wickelgren et al., 1980). Specifically, these studies have been interested in dissociating the contribution of familiarity (a fast, automatic process) and recollection (a slower, controlled process) to performance on recognition memory tasks. By ensuring that familiarity and recollection correspond with competing responses, researchers have been able to reveal the time course of each process and their relative contribution to performance in a variety of situations. The forced-response method could similarly be used to examine memory processes such as these.

Many visual search tasks involve visually scanning an array of stimuli for a target stimulus with salient competing distractor stimuli. Researchers who are interested in the dynamics of distractor and target processing could adopt the forced-response methodology to fruitfully examine the time course of both automatic attention capture and more endogenous, goal-oriented attention (e.g., Zhang et al., 2024). Additionally, a recent study applied the forced-response method to look at how trained stimulus–response mappings affect goal-directed processing when stimulus–response relationships are remapped (Hardwick et al., 2019). There are also many applications of this methodology outside of the strict domain of cognitive psychology. Dual-process models in social psychology have been used to explain stereotyping using tasks such as the implicit-association task and dual-process models in behavioral economics are often invoked to explain reasoning and decision-making in a variety of contexts such as delayed discounting. The application of the forced response method might provide insights that would not be possible using standard tasks in these domains.

Compared with the traditional free-RT approach, the forced-response method is capable of revealing a comprehensive time course of conflict resolution. The data this method returns allows for several different ways in which one could conceptualize and measure conflict resolution. For example, the nadir of accuracy on incongruent trials (the "dip"), the total area enclosed between the congruent and incongruent curves, or the earliest timepoint in which the two curves diverge could all reflect unique aspects of the conflict resolution process. The richness of this information gives us a more in-depth view of conflict resolution than a simple RT difference score.

The current analyses all report effects of interest at the group level. However, it is certainly possible to analyze data generated by the forced-response method for a single participant if individual differences are of interest. Although beyond the scope of the current manuscript, future work could examine whether data obtained via this method are more reliable within an individual relative to free-RT methods.

Limitations

Like any method used in cognitive psychology, there are still some potential limitations for the forced-response method. The method requires a large number of trials to produce the sort of speed–accuracy trade-off curves shown in Fig. 2 relative to free-RT experiments that simply seek to measure a mean RT. This is because the entire time course of processing must be sampled to allow for inference. Trials that are not "on time" either need to be thrown out or accounted for when drawing inferences about the nature of cognitive processing at a particular processing time (although see our analysis above showing that what criterion one uses has very modest effects on the outcome). Additionally, there might still be concern that the forced-response method changes the nature of the task. Since participants are required to resolve conflict while monitoring their timing, one might wonder if this yields a dual task that differs in performance observed using free response. However, this concern should be mitigated somewhat by our experiments above showing classic findings in conflict resolution using the forced-response method. Although the effect sizes we report for these effects (i.e., distractor salience, conflict frequency, previous-trial congruency) are quite robust, they are somewhat weaker than the effect sizes reported when using free RT as the dependent variable as in prior work (e.g., Blais & Bunge, 2010). This attenuation is at least partially driven by the fact that there is a limited overall congruency effect to be modulated at the short PTs where participants are simply guessing and at the longer PTs where accuracy has reached its asymptote. Researchers who are interested in using the forced-response method to investigate these effects could consider more densely sampling PTs from the critical window where the congruency effect is the largest to magnify the overall effects of interest.

At very short processing times at which the stimulus appears just before a response must be emitted, responding accurately or responding on time is a special challenge. This is illustrated in the bottom panel of Fig. 4, which shows a drop in the proportion of PTs at early time points. Although these responses should essentially leave participants at chance performance, a critical number of trials is required to establish this with a high level of confidence. There also may be concerns that when the error rate across the experiment is high, participants might change their response strategy due to a high number of errors. Although this concern is again somewhat allayed by the fact that our method can reproduce classic effects in the cognitive control literature, researchers should bear in mind that the error rate in a forced response experiment may be much higher than tasks they typically use and may alter participants' strategies. If this is of concern, it can be mitigated by oversampling longer processing times where accuracy is expected to be higher. Finally, as is the case with response-signal methods, participants have to withhold their response until the cue at longer processing times. Researchers should again interrogate whether this situation is too distinct from the real-world situations in which they are interested, and whether using the forced-response method would fundamentally change the nature of the relevant cognitive processing.

Conclusion

As we reviewed, the response-signal method is a true advance in the study of cognitive control in that it provides a view of the entire chronometric course of processing in a task. However, its limitation is that there is uncertainty on the part of participants about when a response has to be emitted. What we have proposed here is a significant variant of the response-signal method that eliminates this uncertainty in response production. Participants learn through training to emit a response when it is demanded using this method, and by varying the timing of the onset of the stimulus that they must process, we preserve the value of the response-signal method in that we can trace out the entire course of processing from start to finish as opposed to free-RT methods that provide information at just one slice of time. We recommend application of the forced-response method to a variety of tasks for which it is appropriate, ones in which a prepotent response competes with a goal-directed response. Applying it to these tasks should enrich our understanding of the underlying cognitive processes that are engaged.

Funding This research was supported in part by the National Institute of Mental Health (NIMH) grant R21MH129909 to the University of Michigan (Principal Investigator: JJ), in part by the National Science Foundation (NSF) grant 2342319 to the University of Michigan (Principal Investigator: TL) and in part by a grant from the College of Literature, Science, and the Arts to TL.

Data and Code availability The datasets and code used for the current study are available in the OSF repository, https://osf.io/qa2uc/.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethics approval Approval was obtained from the ethics committee of the University of Michigan. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent to publish Not applicable.

References

Adkins, T. J., & Lee, T. G. (2023). Reward accelerates the preparation of goal-directed actions under Conflict. *Journal of Cognitive Neuroscience*, 1–16. https://doi.org/10.1162/jocn_a_02072

Adkins, T. J., Zhang, H., & Lee, T. G. (2024). People are more error-prone after committing an error. *Nature Communications, 15*(1), 6422. https://doi.org/10.1038/s41467-024-50547-y

Benjamin, A. S., & Bjork, R. A. (2000). On the relationship between recognition speed and accuracy for words rehearsed via rote versus elaborative rehearsal. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 638–648. https://doi.org/10.1037/0278-7393.26.3.638

Blais, C., & Bunge, S. (2010). Behavioral and neural evidence for item-specific performance monitoring. *Journal of Cognitive Neuroscience, 22*(12), 2758–2767. https://doi.org/10.1162/jocn.2009.21365

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Cohen, J. D., Servan-Schreiber, D., & McClelland, J. L. (1992). A parallel distributed processing approach to automaticity. *The American Journal of Psychology, 105*(2), 239–269. https://doi.org/10.2307/1423029

Cronbach, L. J., & Furby, L. (1970). How we should measure" change": Or should we? *Psychological Bulletin, 74*(1), 68.

Dosher, B. A. (1976). The retrieval of sentences from memory: A speed-accuracy study. *Cognitive Psychology, 8*(3), 291–310.

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin, 145*(5), 508–535. https://doi.org/10.1037/bul0000192

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*(3), 265–287.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience, 7*(4), 380–390. https://doi.org/10.3758/CABN.7.4.380

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149.

Fitts, P. M. (1966). Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of Experimental Psychology, 71*(6), 849.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480.

Haith, A. M., Pakpoor, J., & Krakauer, J. W. (2016). Independence of movement preparation and movement initiation. *Journal of Neuroscience, 36*(10), 3007–3015. https://doi.org/10.1523/JNEUROSCI.3245-15.2016

Hardwick, R. M., Forrence, A. D., Krakauer, J. W., & Haith, A. M. (2019). Time-dependent competition between goal-directed and habitual response preparation. *Nature Human Behaviour, 3*(12), 1252–1262. https://doi.org/10.1038/s41562-019-0725-0

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*(3), 1166–1186.

Hedge, C., Powell, G., Bompas, A., & Sumner, P. (2020). Strategy and processing speed eclipse individual differences in control ability in conflict tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.31234/osf.io/vgpxq

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00150

Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General, 136*(2), 217–240. https://doi.org/10.1037/0096-3445.136.2.217

Hilchey, M. D., Ivanoff, J., Taylor, T. L., & Klein, R. M. (2011). Visualizing the temporal dynamics of spatial information processing responsible for the Simon effect and its amplification by inhibition of return. *Acta Psychologica, 136*(2), 235–244. https://doi.org/10.1016/j.actpsy.2010.09.003

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language, 33*(1), 1–18. https://doi.org/10.1006/jmla.1994.1001

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology, 12*(4), 523–553.

Lu, C., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review, 2*(2), 174–207. https://doi.org/10.3758/BF03210959

Manohar, S. G., Chong, T.T.-J., Apps, M. A., Batla, A., Stamelou, M., Jarman, P. R., Bhatia, K. P., & Husain, M. (2015). Reward pays the cost of noise reduction in motor and cognitive control. *Current Biology, 25*(13), 1707–1716.

McElree, B. (1998). Attended and non-attended states in working memory: Accessing categorized structures. *Journal of Memory and Language, 38*(2), 225–252. https://doi.org/10.1006/jmla.1997.2545

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General, 118*(4), 346–373. https://doi.org/10.1037/0096-3445.118.4.346

Miller, J., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin & Review, 20*(5), 819–858. https://doi.org/10.3758/s13423-013-0404-5

Öztekin, I., & McElree, B. (2007). Proactive interference slows recognition by eliminating fast assessments of familiarity. *Journal of Memory and Language, 57*(1), 126–149. https://doi.org/10.1016/j.jml.2006.08.011

Öztekin, I., & McElree, B. (2010). Relationship between measures of working memory capacity and the time course of short-term memory retrieval and interference resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(2), 383–397. https://doi.org/10.1037/a0018029

Pachella, R. G. (1974). The Interpretation of reaction time in information-processing research. In B. Kantowitz (Ed.), *Human information processing* (1st ed., pp. 41–82). Routledge.

Pachella, R. G., & Pew, R. W. (1968). Speed-accuracy tradeoff in reaction time: Effect of discrete criterion times. *Journal of Experimental Psychology, 76*(1p1), 19.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Ratcliff, R. (2006). Modeling response signal and response time data. *Cognitive Psychology, 53*(3), 195–237. https://doi.org/10.1016/j.cogpsych.2005.10.002

Ratcliff, R., & McKoon, G. (1982). Speed and accuracy in the processing of false statements about semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(1), 16–36. https://doi.org/10.1037/0278-7393.8.1.16

Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science, 181*(4099), 574–576.

Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action* (1st ed., pp. 494–519). Oxford University Press.

Schouten, J., & Bekker, J. (1967). Reaction time and accuracy. *Acta Psychologica, 27*, 143–153.

Schramm, P., & Rouder, J. N. (2019). *Are Reaction Time Transformations Really Beneficial?* OSF. https://doi.org/10.31234/osf.io/9ksa6

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81*(1), 174.

Stafford, T., Pirrone, A., Croucher, M., & Krystalli, A. (2020). Quantifying the benefits of using decision models with response time and accuracy data. *Behavior Research Methods, 52*(5), 2142–2155. https://doi.org/10.3758/s13428-020-01372-w

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643.

Swensson, R. G., & Edwards, W. (1971). Response strategies in a two-choice reaction task with a continuous cost for time. *Journal of Experimental Psychology, 88*(1), 67.

Teichert, T., Ferrera, V. P., & Grinband, J. (2014). Humans optimize decision-making by delaying decision onset. *PLoS ONE, 9*(3), e89638. https://doi.org/10.1371/journal.pone.0089638

van Leeuwen, J., Smeets, J. B. J., & Belopolsky, A. V. (2019). Forget binning and get SMART: Getting more out of the time-course of response data. *Attention, Perception, & Psychophysics, 81*(8), 2956–2967. https://doi.org/10.3758/s13414-019-01788-3

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods, 49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

Vandierendonck, A. (2021). On the utility of integrated speed-accuracy measures when speed-accuracy trade-off is present. *Journal of Cognition, 4*(1), 22. https://doi.org/10.5334/joc.154

von Bastian, C. C., Blais, C., Brewer, G. A., Gyurkovics, M., Hedge, C., Kałamała, P., Meier, M. E., Oberauer, K., Rey-Mermet, A., Rouder, J. N., Souza, A. S., Bartsch, L. M., Conway, A. R. A., Draheim, C., Engle, R. W., Friedman, N. P., Frischkorn, G. T., Gustavson, D. E., Koch, I., …., & Wiemers, E. A. (2020). *Advancing the understanding of individual differences in attentional control: Theoretical, methodological, and analytical considerations* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/x3b9k

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58*(3), 475–482. https://doi.org/10.1007/BF03395630

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41*(1), 67–85.

Wickelgren, W. A., Corbett, A. T., & Dosher, B. A. (1980). Priming and retrieval from short-term memory: A speed accuracy trade-off analysis. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 387–404. https://doi.org/10.1016/S0022-5371(80)90276-5

Wong, A. L., Goldsmith, J., Forrence, A. D., Haith, A. M., & Krakauer, J. W. (2017). Reaction times can reflect habits rather than computations. *Elife, 6*, e28075.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Zeischka, P., Coomans, D., Deroost, N., Vandenbossche, J., & Soetens, E. (2011). Target–flanker discriminability affects conflict size but not sustained suppression. *Acta Psychologica, 136*(1), 148–156.

Zhang, H., Sellers, J., Lee, T., & Jonides, J. (2024). The Temporal Dynamics of Visual Attention. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001661