



# Tradeoffs of estimating reaction time with absolute and relative thresholds

Jarrold Blinch<sup>1</sup> · Coby Trovinger<sup>1</sup> · Callie R. DeWinne<sup>1</sup> · Guilherme de Cellio Martins<sup>1</sup> · Chelsea N. Ifediora<sup>1</sup> · Maryam Nourollahimoghadam<sup>1</sup> · John R. Harry<sup>1</sup> · Ty B. Palmer<sup>1</sup>

Accepted: 29 July 2023 / Published online: 25 August 2023  
© The Psychonomic Society, Inc. 2023

## Abstract

Measuring the duration of cognitive processing with reaction time is fundamental to several subfields of psychology. Many methods exist for estimating movement initiation when measuring reaction time, but there is an incomplete understanding of their relative performance. The purpose of the present study was to identify and compare the tradeoffs of 19 estimates of movement initiation across two experiments. We focused our investigation on estimating movement initiation on each trial with filtered kinematic and kinetic data. Nine of the estimates involved absolute thresholds (e.g., acceleration 1000 back to 200 mm/s<sup>2</sup>, micro push-button switch), and the remaining ten estimates used relative thresholds (e.g., force extrapolation, 5% of maximum velocity). The criteria were the duration of reaction time, immunity to the movement amplitude, responsiveness to visual feedback during movement execution, reliability, and the number of manually corrected trials (efficacy). The three best overall estimates, in descending order, were yank extrapolation, force extrapolation, and acceleration 1000 to 200 mm/s<sup>2</sup>. The sensitive micro push-button switch, which was the simplest estimate, had a decent overall score, but it was a late estimate of movement initiation. The relative thresholds based on kinematics had the six worst overall scores. An issue with the relative kinematic thresholds was that they were biased by the movement amplitude. In summary, we recommend measuring reaction time on each trial with one of the three best overall estimates of movement initiation. Future research should continue to refine existing estimates while also exploring new ones.

**Keywords** Reaction time · Movement initiation · Movement amplitude · Kinematics · Kinetics

## Introduction

Mental chronometry, which involves measuring the duration of cognitive processing, is fundamental to many subfields of psychology, including cognitive neuroscience (Fan et al., 2002; Posner, 2005), cognitive psychology (e.g., Telford, 1931; Stroop, 1935; Hick, 1952; Hyman, 1953, Fitts & Seeger, 1953; Henry & Rogers, 1960; Sternberg, 1966, 1969; Simon, 1969; Shepard & Metzler, 1971; Eriksen & Eriksen, 1974; Rosenbaum, 1980), and neuropsychology (Milner, 1986). One of the common ways to measure the duration of cognitive processing is with reaction time, which is the interval from the go signal to movement initiation. Reaction time was introduced to modern psychology by Hermann Helmholtz (Helmholtz, 1850), who used simple reaction time tasks

to measure neural conduction velocity. Wilhelm Wundt, the father of experimental psychology, was one of Helmholtz's students. Wundt used reaction time to measure the stages of information processing (Wundt, 1880). Reaction time has been important to the scientific study of mind and behavior, and reaction time continues to be used by psychologists today. There is no “gold standard” for how to measure reaction time, despite its use for over a hundred years in psychology. The go signal can be measured precisely with modern scientific equipment. The challenge, however, is measuring the exact moment of movement initiation. The purpose of the present study was to identify and compare the tradeoffs of 19 estimates of movement initiation.

When measuring reaction time, it is important to remember that it is only the first part of the total response. Reaction time is immediately followed by movement time, the interval from movement initiation to movement termination. Reaction time and movement time are collectively the total response time. Most early studies that measured reaction time involved straightforward responses, like pressing or releasing

✉ Jarrod Blinch  
jarrod.blinch@gmail.com

<sup>1</sup> Department of Kinesiology & Sport Management, Texas Tech University, Box 43011, Lubbock, TX 79409, USA

a push-button switch. Many modern studies still use the voltage change of a push-button switch to estimate movement initiation. Switches have remained popular because pressing or releasing a push-button switch produces a square wave voltage signal and it is very easy to detect the edge of a square wave (for example, when the voltage changes from 0 to 5 V). The simplicity of switches comes with a major drawback; they are unquestionably a late estimate of movement initiation. This causes an overestimation of reaction time and an underestimation of movement time. This occurs because push-button switches must travel a short distance before actuation occurs. This problem can be mitigated, but not eliminated, by using a micro push-button switch that has short travel and low force operating characteristics. In Experiment 1, we used a standard micro push-button switch, which was more sensitive than a typical mechanical keyboard key. We delved deeper into the operating characteristics of micro push-button switches in Experiment 2.

## Experiment 1

As psychologists began to appreciate the interactions between perception, cognition, and action, the responses got more complex, like pointing to a target or reaching and grasping an object. These complex responses are often recorded with motion capture systems and analyzed with trajectory analysis. The use of trajectory analysis has led to new methods of estimating movement initiation. Movement initiation is commonly estimated by finding the first time that velocity in the primary direction (or tangential velocity) of the movement exceeds a percentage (often 5%) of its maximum (early uses include Lacquaniti and Soechting [1982] for joint angular velocity, Kapoule and Robinson [1986] for eye velocity, Gracco and Abbs [1986] for lip velocity, and Corcos et al. [1988] and Darling et al. [1988] for arm and finger velocity). This is a relative threshold because it depends on the maximum velocity of each trial. The logic of a relative threshold is to normalize movements with different kinematic or kinetic profiles (Lacquaniti & Soechting, 1982; Teasdale et al., 1993). A relative threshold should normalize the different profiles of short- and long-amplitude movements or different participants with slower or faster movements. We will return to examine this assumption in the Discussion section, as we were surprised to find that relative thresholds were inappropriate for the range of movements within- and between-participants in the present experiment.

## Estimates of movement initiation

In the current experiment, we compared a total of 19 estimates of movement initiation and ten of these estimates had

relative thresholds. The first three relative thresholds were when the kinematic data (position, velocity, and acceleration) first exceeded 5% of its maximum value. The initial rise in the kinematics gets steeper and occurs earlier as you differentiate from position to velocity and from velocity to acceleration. The consequence is that, for example, 5% of maximum acceleration will occur before 5% of maximum velocity. In other words, the higher the derivative, the earlier the estimate of movement initiation and the shorter the estimate of reaction time. There are two downsides to using higher derivatives. First, a small change in position will cause a large change in acceleration. Second, differentiation degrades the signal-to-noise ratio. Both disadvantages make premature estimates of movement initiation more common with higher derivatives from small, spurious movements before the actual movement, from data degradation, or from both.

There is another type of relative threshold that has been advocated for by Brenner and Smeets (2019). Their extrapolation method was first intended to measure the onset of online corrections (Oostwoud Wijdenes et al., 2014). It involves finding 25 and 75% of maximum velocity, drawing a straight line through these points, and extrapolating it downward. A horizontal line is also drawn at the level of baseline velocity before movement initiation. The estimate of movement initiation is the point where these two lines intersect. An advantage of this extrapolation method is that it is far less likely for a spurious movement to reach 25% of maximum velocity compared to 5% of maximum velocity. The same advantage could be achieved by increasing the threshold of 5% of maximum velocity to 25%, but this would drastically increase the estimate of movement initiation. The extrapolation method avoids this by extrapolating downward from 25% to baseline velocity. In the present experiment, we compared the extrapolation method on position, velocity, and acceleration in the primary direction of the movement. We, again, expected that higher derivatives would have shorter estimates of reaction time.

Besides kinematics, we also recorded the kinetics around movement initiation by measuring vertical force on a micro push-button switch. It is less common to measure kinetics than kinematics in reaction time studies, but it is easy to add; all that is required is a force sensor, an amplifier, and an additional analog input channel (Englund & Patching, 2009). Movement initiation was estimated by applying 5% threshold and extrapolation methods to force and yank (the rate of change of force is called yank; Lin et al., 2019). Reaction time based on kinetics should be shorter than those based on kinematics because changes in force precede changes in position, changes in yank precede changes in velocity, etc.

The opposite of relative thresholds are absolute thresholds (e.g., Chua & Elliott, 1993; Binsted & Elliott, 1999;

Blinch et al., 2021). Nine of the estimates of movement initiation had absolute thresholds. The first and simplest absolute threshold was the actuation of a micro push-button switch. The remaining absolute thresholds involved the kinematic data. The next two absolute thresholds were when velocity or acceleration in the primary direction of the movement first exceeded 50 mm/s or 1000 mm/s<sup>2</sup>, respectively. These two absolute thresholds ended up being among the best estimates, and so we included three attempts to improve them. First, we reduced the velocity and acceleration thresholds by 80% to 10 mm/s and 200 mm/s<sup>2</sup>. These lower absolute thresholds should have shorter estimates of reaction time. The downside of a lower absolute threshold is an increase in premature estimates of movement initiation from small, spurious movements before the actual movement or from noise in the signal, especially with higher derivatives. We attempted to eliminate that downside with the second and third improvements to the absolute thresholds. These both involved adding a second criteria for movement initiation, which is like a double-threshold detector (Bonato et al., 1998). The second estimates required that the thresholds of 10 mm/s and 200 mm/s<sup>2</sup> were exceeded for 100 ms. The third estimates began by finding when the movements first exceeded 50 mm/s and 1000 mm/s<sup>2</sup> and then searched backwards in time to find when the movements last exceeded 10 mm/s and 200 mm/s<sup>2</sup>. These third estimates were similar to Algorithm B by Teasdale et al. (1993).

### Evaluation criteria

We used four criteria to evaluate the estimates of movement initiation: time, immunity/responsiveness, reliability, and efficacy. First, the time estimate of movement initiation, with shorter estimates likely being closer to the actual moment of movement initiation. We already detailed how the actuation of a push-button switch is a very late estimate of movement initiation. The other 18 estimates of movement initiation likely overestimate movement initiation to different degrees. Take velocity 50 mm/s for example; it overestimates movement initiation because the movement began at 1 mm/s (or 0.1 mm/s) and not at 50 mm/s. Thus, shorter estimates of movement initiation are likely to have smaller overestimates and be closer to the actual movement of movement initiation. It is possible for any estimate to underestimate movement initiation from either small, spurious movements before the actual movement or from data degradation. In the current study, this possibility was minimized by visually inspecting every estimate on every trial and making corrections when necessary.

Second, the immunity and responsiveness of the measures to the characteristics of the movements and the task. We compared movements to short- and long-distance targets, or, equivalently, short- and long-amplitude movements.

Long- and short-amplitude movements must have different response programming to reach different distance targets. However, we hypothesized that their response complexity, specifically their within-chunk complexity is comparable. It has been shown that increasing within-chunk complexity increases choice reaction time (reviewed by Klapp & Maslovat, 2020). An example of manipulating within-chunk complexity is a word with one or three syllables (e.g., Klapp, 2003). Another common type of manipulation is a key press-then-release (dit) or a key press-hold-release (dah; Klapp, 1995). Long- and short-amplitude movements likely have the same within-chunk complexity because they both have the same number of elements, with the long-amplitude movement being a scaled-up version of the short-amplitude movement (or vice versa). Thus, long- and short-amplitude movements should have comparable reaction time. In other words, they should be immune to the movement amplitude.

For the responsiveness of the estimates to the characteristics of the task, we compared reaction time with and without visual feedback during movement execution. Studies have shown that reaction time increases when vision is occluded during movement execution, especially when the visual feedback conditions are blocked (Carson et al., 1990; Hansen et al., 2006; Khan et al., 2002; Krigolson & Heath, 2004; Westwood & Goodale, 2003). Therefore, a valid estimate of reaction time should be longer without vision during movement execution. The reason that movements without visual feedback have longer reaction time is that the movement must be more carefully planned during the reaction time as online control is reduced without real-time visual feedback.

The third criterion was reliability; a more reliable measure would require fewer trials to achieve the conventional level of reliability. Variability is sometimes measured as a substitute for reliability but that is ill-advised because one is ultimately interested in reliability and there is a complex relationship between reliability and variability; for example, a more variable measure is not always less reliable (detailed in the Reliability and Generalizability section). Fourth, and finally, efficacy, which was the proportion of trials that were manually corrected after visual inspection. We reasoned that a worthwhile estimate of movement initiation should have 5% or less manually corrected trials. Estimates with more than 5% corrected trials were deemed inefficient and excluded from subsequent consideration.

### Overall assessment

A researcher could select the best estimate of movement initiation based on any individual criterion. However, we believe that all four criteria are relevant to most reaction time studies. We, therefore, calculated an overall score to determine the best and worst estimates of movement initiation.

First, individual scores were calculated for three criteria. These were reaction time, absolute difference in reaction time for long- and short-amplitude movements, and number of trials to achieve the conventional level of reliability for the time, immunity, and reliability criteria. Overall scores were only calculated for estimates with 5% or less manually corrected trials. The individual scores were converted to range from 0, for the best score, to 1. The overall score for each estimate was the mean of the individual scores for time, immunity, and reliability. Note that this procedure resulted in the same pattern of results as using Z-scores. Equal weights (i.e., the mean) were used because we thought that time, immunity, and reliability were equally important. It is likely that other researchers will weigh the criteria differently. This could even occur for the same researcher on different reaction time studies. Thus, an Excel spreadsheet, Overall Scores.xlsx, was uploaded to the project repository on osf.io so that anyone can change the weights of the criteria and see the effects on the overall scores. There were two reasons that efficacy was used as an inclusion criterion and not part of the overall score. First, every reaction time study should include visual inspection of every trial. This is a vital step regardless of whether 1% or 0.1% of the trials require manual corrections. Second, visual inspection is a subjective process that cannot be objectively quantified like time, immunity to the movement amplitude, and reliability.

### Additional assumptions

There are two additional assumptions about how we evaluated estimates of movement initiation. First, we only considered estimates that were calculated based on filtered data for each trial and not ones that were calculated after first averaging across trials in a condition. Individual trial data is needed for reliability analysis, which was one of the criteria we used to assess the estimates of movement initiation. It is also needed for calculating intertrial reaction time variability, distribution analysis, and trial-to-trial correlation analysis. However, individual trial data is not always necessary. In those cases, we recommend the reaction time methodology study by Brenner and Smeets (2019) that included estimates on condition averages.

The second assumption was that we were interested in the methodology of estimating the duration of cognitive processing with reaction time and not premotor reaction time. Reaction time can be divided into premotor time and motor time. Premotor time is the interval from the go signal until the onset of electromyography (EMG) activity in the prime mover muscle. Motor time is the interval from the onset of EMG activity until movement initiation. Premotor time is a shorter and more accurate measurement of cognitive processing than reaction time because cognitive processing of the open-loop portion of the motor command is likely

finished when EMG activity begins; the motor command has been initiated and has reached its destination. The reason we focused on the methodology of estimating reaction time is because it is more prevalent in psychology than premotor time. There are a few reasons why reaction time is more popular than premotor time, and we will return to these in the Discussion section.

In summary, the purpose of Experiment 1 was to identify and compare *the time, immunity/responsiveness, reliability, and efficacy* of 19 estimates of movement initiation. The goal was not to establish a “gold standard” for estimating reaction time, because it is unlikely that one measure would be the best on all four criteria in all reaction time tasks. Instead, the goal was to understand the tradeoffs of each measure so that researchers can choose the one that is the best for their experimental design.

## Method

### Participants

Forty-one volunteer participants (27 female participants and 14 male participants) were recruited from the university community. The age of participants ranged from 19 to 37 years old ( $M = 22.0$ ,  $Mdn = 21.0$ ,  $SD = 3.35$ ). Our goal was to test 40 participants, twice the minimum threshold advocated by Simmons et al. (2011). Hand-use preference was determined with a modified version of the Edinburgh (Oldfield, 1971) and Waterloo (Brown et al., 2006) handedness questionnaires (see Stone et al. [2013] for the questionnaire). Thirty-eight participants were right-handed (scores  $\geq 0.5$ ), two were left-handed (scores  $\leq -0.5$ ), and one had inconsistent handedness ( $-0.5 < \text{scores} < 0.5$ ). All participants reported normal or corrected-to-normal vision.

### Design

Participants first only completed the modified Edinburgh–Waterloo handedness questionnaire. This was followed by only four blocks of a two-choice reaction time task with pointing movements to either a short- or long-distance target. For half of the participants, visual information was available for the entire trial in the first two blocks, and, in the last two blocks, visual information was occluded during movement execution (from movement initiation, as estimated by release of the home switch, to movement termination). This order was reversed for the other half of the participants. Each block consisted of 64 trials, with 32 trials to the short-distance target and 32 trials to the long-distance target. The order of the short- and long-distance targets was randomized, and a different randomization was used for each block and participant.

## Apparatus

Participants were seated at a table with a button box (60.5 cm length, by 30.0 cm width, by 7.0 cm height) on the surface of the table. They were seated so that their midsagittal plane was centered with the button box. The button box consisted of a home switch, a short-distance target, and a long-distance target. The center of the home switch was 10 cm in from the front edge of the table and the button box. The short- and long-distance targets were 10 and 20 cm (center-to-center) in front of the home switch, respectively. The home switch was a black micro push-button switch with a diameter of 7.1 mm, a pretravel of 0.3 mm, and an operating force of 2.55 N (Adafruit Industries, product id 1119). The short- and long-distance targets were clear push-button switches with a diameter of 24.5 mm and a pretravel of 1.5 mm (Adafruit Industries, product id 492). The push-button switches were each illuminated by a diffused green light-emitting diode (LED). A 5-kg micro load cell (RobotShop Inc., RB-Phi-118) was positioned underneath the home switch, and the load cell was connected to a Wheatstone amplifier (AD8426, RobotShop Inc., RBC-Onl-38).

A jumbo craft stick (18 mm wide) was cut to the length of each participant's distal phalanx of their dominant index finger. This was placed on top of the distal phalanx (i.e., over the nail) and secured in place with Transpore surgical tape (3M). A  $\pm 3$  g 3-axis accelerometer (ADXL335, AdaFruit Industries, product id 163) was attached on top of the craft stick. An LED was attached on top of the accelerometer. A motion capture system with nine detectors (PhaseSpace, Impulse X2) recorded the position of the LED at 960 Hz. Participants wore visual occlusion spectacles (Translucent Technologies, PLATO) that controlled their access to visual information during the trials. A data acquisition device (National Instruments, USB-6012) recorded the voltage signals of the switches, accelerometer, and load cell at 1000 Hz. It also controlled the illumination of the target switches and the opening and closing of the spectacles.

## Procedures

All trials began with the spectacles open to allow visual information. The participants started each trial by pressing and holding down the home switch with the index finger of their dominant hand. There was a 1 to 2-s variable foreperiod before either the short- or long-distance target illuminated as the go signal. The participant was instructed to “react and press the illuminated target as quickly and as accurately as possible”. The participant held down the target switch at the end of their movement until the 2-s recording interval of the motion capture system ended and the target darkened. They could then begin the next trial by returning to the home switch. Trials in the condition without visual information

during movement execution were slightly different, in that the spectacles closed to occlude visual information from the release of the home switch until the target switch was pressed (i.e., from movement initiation to movement termination). The spectacles opened at movement termination to provide terminal feedback of the movement. Trials with more than 10 ms of missing motion capture data, anticipation (reaction time less than 100 ms), or inattention (reaction time or movement time greater than 1000 ms) were recycled to the end of the block, and the original trials were excluded from data analysis. There was a total of 87 recycled trials and 10,496 non-recycled trials. Fifty-nine recycled trials had missing motion capture data, nine were anticipation, and 19 were inattention. The number of recycled trials per participant ranged from 0 (for 11 participants) to 16 (for one participant).

## Data analysis<sup>1</sup>

Position and force data were filtered with a low-pass Butterworth filter (dual-pass, fourth order) with a cutoff frequency of 20 Hz (the order and cut-off frequency were after the two passes of the filter). Velocity, acceleration, and yank were calculated with a two-point central difference method.

### Estimates of movement initiation and reaction time

The estimates of movement initiation were calculated for each trial. An example of the estimates of movement initiation for one representative trial are shown in the [Appendix](#). The time of the go signal was determined by when the voltage signal to illuminate the target switch first exceeded 2.5 V. The first estimate of movement initiation was based on the release of the home switch. The actuation of the micro push-button switch caused a voltage change from 0 to 5 V. Movement initiation based on the home switch was determined by when the voltage first fell below 2.5 V.

The next three estimates of movement initiation were determined by when position, velocity, and acceleration in the primary direction of the movement first exceeded 5% of their maximum values (Fig. 7). Baseline position was the mean position during the 15 ms before the go signal. Baseline velocity and acceleration were 0 mm/s and 0 mm/s<sup>2</sup>, respectively. The next two estimates were when force and yank first exceeded 5% of their maximum values. The baselines for force and yank were maximum force and 0 N/s, respectively.

An extrapolation method was used on position, velocity, acceleration, force, and yank to determine another five

<sup>1</sup> The data and statistical analysis files for the present study are available at <https://osf.io/RBYNF/>.



estimates of movement initiation. For each trajectory, 25% and 75% of their maximum values were found and a straight line between these points was extrapolated back to the same baseline as the 5% method. The intersection of this line with the baseline was the estimate of movement initiation (Fig. 8).

The last eight estimates were absolute thresholds based on velocity or acceleration. For velocity, they were when velocity first exceeded 50 mm/s (Fig. 9), when velocity first exceeded 10 mm/s (Fig. 10), when velocity first exceeded 10 mm/s for at least 100 ms, and when velocity first exceeded 50 mm/s and then back in time to when it last exceeded 10 mm/s. For acceleration, the estimates were when acceleration first exceeded 1000 mm/s<sup>2</sup> (Fig. 9), when acceleration first exceeded 200 mm/s<sup>2</sup> (Fig. 10) when acceleration first exceeded 200 mm/s<sup>2</sup> for at least 100 ms, and when acceleration first exceeded 1000 mm/s<sup>2</sup> and then back in time to when it last exceeded 200 mm/s<sup>2</sup>. All 19 estimates of movement initiation were visually inspected and corrected when necessary.

The estimates of movement initiation were used to calculate 19 estimates of reaction time. Reaction time was defined as the duration from the go signal until movement initiation. The estimates of reaction time were analyzed with a 14 Estimate (position 5%, velocity 5%, acceleration 5%, position extrapolation, velocity extrapolation, acceleration extrapolation, force extrapolation, yank extrapolation, switch, velocity 50 mm/s, acceleration 1000 mm/s<sup>2</sup>, velocity 10 mm/s, velocity 10 mm/s for 100 ms and velocity 50 to 10 mm/s, acceleration 1000 to 200 mm/s<sup>2</sup>) by 2 Vision (vision, no vision) by 2 Movement Amplitude (short, long) repeated measures analysis of variance. Note that four estimates were excluded because they were ineffective (force 5%, yank 5%, acceleration 200 mm/s<sup>2</sup>, and acceleration 200 mm/s<sup>2</sup> for 100 ms), which is detailed in the Efficacy section of the Results. Furthermore, velocity 10 mm/s for 100 ms and velocity 50 to 10 mm/s were combined into one level because they had identical reaction time on every trial (after manual corrections based on visual inspection). The time of each estimate was analyzed with pairwise comparisons on the main effect of estimate. The immunity of reaction time to the movement amplitude was analyzed with the estimate by movement amplitude interaction. The responsiveness of reaction time to visual feedback was analyzed with the estimate by vision interaction. If an interaction was significant, then it was examined with pairwise comparisons on the mean differences between long- and short-amplitude movements or vision and no vision conditions, respectively.

When local sphericity was violated (as indicated by Mauchly's test,  $\alpha \leq .10$ ), the Huynh–Feldt correction was used when  $\epsilon$  was greater than or equal to .75 and the Greenhouse–Geisser correction was used otherwise (Huynh & Feldt, 1976). Uncorrected degrees of freedom and  $\epsilon$  values were reported (Huynh–Feldt  $\epsilon_{HF}$ , Greenhouse–Geisser  $\epsilon_{GG}$ ). The Bonferroni correction was used to control the

familywise error rate. Reported values were means and 95% correlation- and difference-adjusted confidence intervals (calculated with the Cousineau–Morey method; Cousineau, 2017) or 95% confidence intervals of the difference. Effect sizes were reported using partial eta-squared ( $\eta_p^2$ ).

### Reliability and generalizability theory

According to classical test theory, reliability is the degree to which a test score is consistent across repeated observations, with the assumption that the true score is the average of observed test scores obtained over an infinite number of observations. Generalizability theory (G theory) is an extension of classical test theory that uses analysis of variance methods to decompose the variance of observed test scores into multiple sources of variability and the interactions between sources (reviewed by Brennan, 2001; Cronbach et al., 1972; Vispoel et al., 2018). This is similar to main effects and interactions in an analysis of variance. In the current experiment, reaction time was measured for a group of participants on several trials, which is a person crossed with trial experiment in G theory terminology.

A generalizability study (G study) first estimated the variance contributed by person, trial, and the person-by-trial interaction. Person variance was true variance that was caused by interindividual differences in reaction time. Trial variance was systematic error associated with inter-trial variability in reaction time. Person-by-trial variance was random error caused by intra-individual variability in reaction time. These latter two components of variability (trial and person-by-trial) are measurement error. The estimated variance components of the G study were interpreted as the percentage they each contribute to the total estimated variance. G theory quantifies reliability (the generalizability coefficient) as the ratio of person variance to person variance plus person-by-trial variance. An increase in person variance will increase reliability, an increase in person-by-trial variance will decrease reliability, and trial variance has no effect on reliability. This demonstrates the complex relationship between reliability and variability; an increase in variability, depending on the source, can either increase, decrease, or have no effect on reliability.

Based on the estimated variance components from the G study, a follow-up decision study (D study) then estimated the change in reliability by increasing or decreasing the number of trials per participant. We determined the minimum number of trials to achieve the conventional level of reliability (generalizability coefficient [ $g$ ]  $\geq .80$ ). Note that the number of trials determined by the D study can be more than the actual number of trials collected in the experiment.

The estimates of reaction time in the four conditions (2 Vision [vision, no vision] by 2 Movement Amplitude [short, long] by) were analyzed with a MATLAB G theory script

written by Mushquash and O'Connor (G1.m<sup>2</sup>; 2006). The type variable was set to 1 to indicate a single-facet, fully crossed design. The nfacet1 variable, the number of trials, was set to 64.

### Overall scores

Overall scores were calculated for the 15 effective estimates of movement initiation. These calculations are shown in the Overall Scores.xlsx Excel file inside the project repository on osf.io. The overall scores were based on three criteria: time, immunity to the movement amplitude, and reliability. As an example, the grand mean reaction time of the switch was 282.2 ms. Position extrapolation had the longest reaction time (314.2), and yank extrapolation had the shortest reaction time (199.5 ms). The normalized time score for the switch was calculated by taking its reaction time, subtracting the minimum reaction time, and then dividing it by the reaction time range ( $[282.2 - 199.5] / [314.2 - 199.5] = .72$ ). This process converted the scores to range from 0 to 1. Yank extrapolation had a normalized time score of 0 ( $[199.5 - 199.5] / [314.2 - 199.5] = 0$ ), and position extrapolation had a normalized time score of 1 ( $[314.2 - 199.5] / [314.2 - 199.5] = 1$ ). This means that they had the best and worst reaction time out of the 15 effective estimates. Back to the switch, it had normalized scores of .72, 0, and 0.32 for time, immunity to the movement amplitude, and reliability. Its overall score was calculated by taking the mean of these three scores ( $[(.72 + 0 + .32) / 3 = .35]$ ).

## Results

### Efficacy

The efficacy of each estimate was determined by the percentage of trials that had to be manually corrected after visual inspection. Estimates of movement initiation were classified as either effective (Table 1) or ineffective. Ineffective estimates had more than 5% manually corrected trials. There were four ineffective estimates of movement initiation: acceleration 200 mm/s<sup>2</sup>, acceleration 200 mm/s<sup>2</sup> for 100 ms, force 5%, and yank 5%. These ineffective estimates were excluded from subsequent analysis. For acceleration 200 mm/s<sup>2</sup>, the proportion of manually corrected trials was so high, 21.29%, after visual inspection of the first four participants that we prematurely ended the visual processing. Acceleration 200 mm/s<sup>2</sup> for 100 ms was more efficient at

**Table 1** Percentage of manually corrected trials in descending order

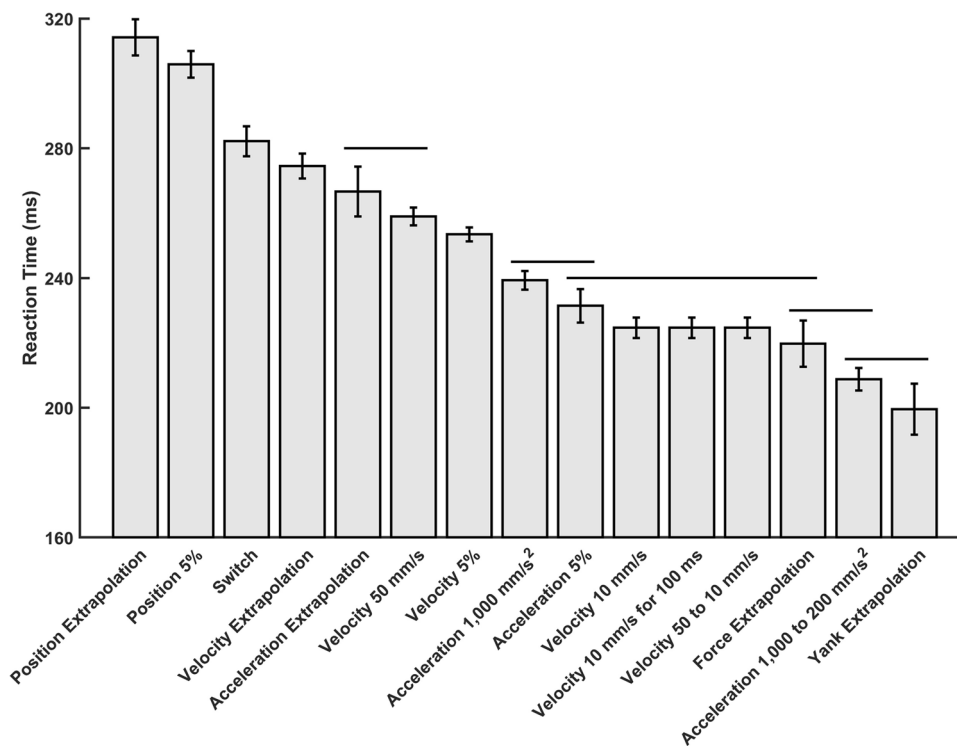
Estimate	Percentage of manually corrected trials (%)
Velocity 10 mm/s	3.24
Yank extrapolation	1.93
Acceleration 5%	1.66
Force extrapolation	1.06
Acceleration 1000 mm/s <sup>2</sup>	0.40
Acceleration 1000 to 200 mm/s <sup>2</sup>	0.40
Velocity 10 mm/s for 100 ms	0.29
Velocity 50 to 10 mm/s	0.10
Velocity 5%	0.09
Acceleration extrapolation	0.08
Velocity 50 mm/s	0.08
Position 5%	0.00
Position extrapolation	0.00
Velocity extrapolation	0.00
Switch	0.00

10.61%, but still above the 5% threshold. Force and yank exceeded their 5% threshold between the start of data collection and 100 ms after the go signal on 41.3% and 83.2% of trials, respectively. In contrast, it was rare for fluctuations in the kinematic data to prematurely exceed the 5% thresholds (0, 0.06, and 1.39% of trials for position, velocity, and acceleration).

### Reaction time

Mean reaction time for the 15 effective estimates of movement initiation are shown in Fig. 1. There was a large range in the estimates of reaction time; the difference between the longest mean reaction time (position extrapolation) and the shortest (yank extrapolation) was 115 ms, [100, 130]. Higher derivatives (e.g., acceleration and yank) had shorter estimates of reaction time compared to lower derivatives (e.g. position and force); for example, the 5% relative threshold based on position, velocity, and acceleration yielded reaction time of 306 ms, [302, 310], 253 ms, [251, 256], and 231 ms, [226, 237], respectively. This is because the initial rise in the data gets steeper and occurs earlier with higher derivatives, which exceeds the relative threshold sooner. Reaction time from the extrapolation method based on kinetics (force and yank) were shorter than those based on kinematics (position and velocity). This occurred because changes in force and yank precede changes in position and velocity, respectively. As for the statistical analysis, there was a significant main effect of estimate,  $F(13, 520) = 422.93, p < 0.01, \epsilon_{GG} = .24, \eta_p^2 = .91$ , which was analyzed with pairwise comparisons (Fig. 1).

<sup>2</sup> G1.m is available for download from <https://people.ok.ubc.ca/briocomm/gtheory/gtheory.html>. There are also versions for SAS, SPSS, and R.



**Fig. 1** Mean reaction time in descending order. *Note.* Error bars are 95% difference- and correlation-adjusted confidence intervals (Cousineau, 2017). Horizontal lines are drawn overestimates that are not significantly different ( $p > 0.05$ )

If we consider just movement initiation (collapsed across vision and movement amplitude), then the three shortest estimates of movement initiation are yank extrapolation, acceleration 1000 to 200 mm/s<sup>2</sup>, and force extrapolation. As for the longest estimates of movement initiation, we do not recommend position extrapolation or position 5% because they were longer than the switch, which is a late estimate of movement initiation.

### Immunity to the movement amplitude

Mean reaction time differences between long- and short-amplitude movements are shown in Fig. 2. Positive values indicate that long-amplitude movements had longer reaction time compared to short-amplitude movements. The increase in reaction time for long-amplitude movements compared to short-amplitude movements ranged from 18.8 ms, [17.1, 20.5] for position 5% to  $-1.1$  ms,  $[-2.3, 0.0]$  for acceleration 1000 mm/s<sup>2</sup>. There was a significant interaction between estimate and movement amplitude,  $F(13, 520) = 161.01$ ,  $p < 0.01$ ,  $\epsilon_{GG} = .32$ ,  $\eta_p^2 = .80$ . The difference in reaction time between long- and short-amplitude movements for the 15 effective estimates were analyzed with pairwise comparisons (Fig. 2). If we consider just this characteristic of the movements, then nine of the estimates were comparable,  $ps \geq 0.27$ , and not significantly different from zero,  $ps \geq 0.45$ .

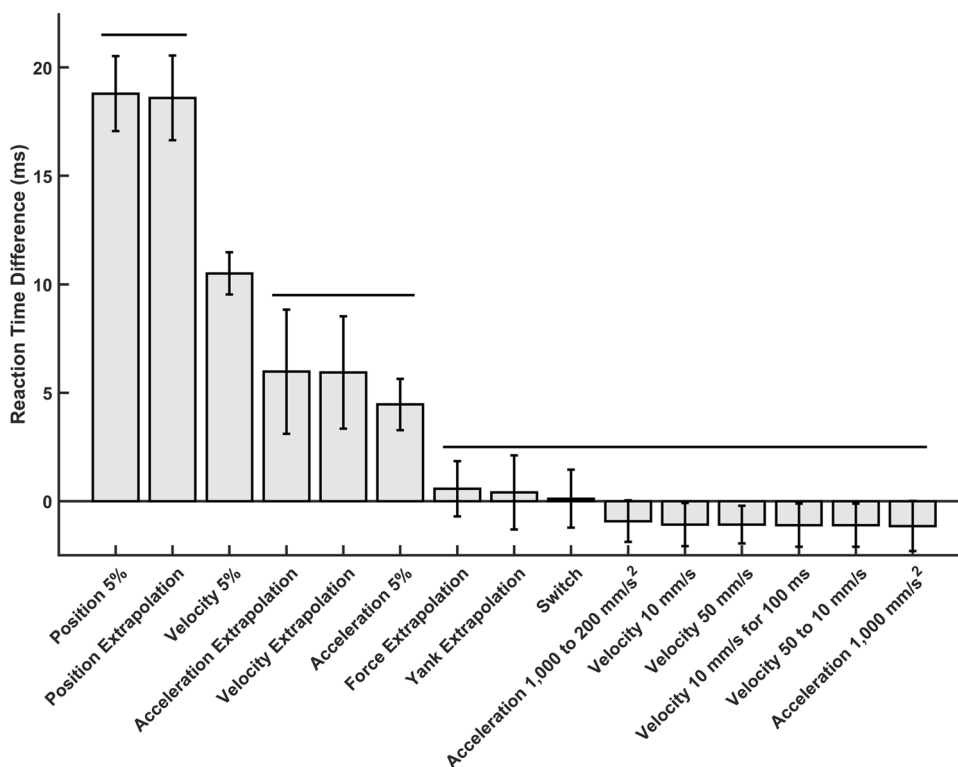
These results suggest that these nine estimates had comparable reaction time for short- and long-amplitude movements. Five of the reaction time differences were significantly larger than zero,  $ps < 0.01$ , which indicated longer reaction time for long-amplitude movements compare to short-amplitude movements. The largest reaction time differences were for position 5% and position extrapolation.

Recall that we predicted that relative thresholds (5% and extrapolation) would be immune to the effects of movement amplitude, but this was not the case. For the relative estimates, there was a positive linear relationship between the size of the reaction time difference (Fig. 2) and reaction time (Fig. 1),  $r = .91$ ,  $[.58, .98]$ ,  $p < 0.01$ . This casts doubt on the assumption that relative thresholds should normalize the different profiles of short- and long-amplitude movements and yield comparable reaction time. We will return to this topic in the Discussion. Interestingly, all the absolute estimates had comparable reaction time for short- and long-amplitude movements.

### Responsiveness to visual feedback

For the responsiveness of reaction time to visual feedback, reaction time was 17.7 ms, [11.3, 24.0] longer without vision than with vision,  $F(1, 40) = 31.45$ ,  $p < 0.01$ ,  $\eta_p^2 = .44$ . The estimate by vision interaction was not significant,  $F(13, 520)$





**Fig. 2** Mean reaction time differences between long- and short-amplitude movements in descending order. *Note.* Error bars are 95% difference- and correlation-adjusted confidence intervals (Cousineau,

2017). Horizontal lines are drawn over estimates that are not significantly different ( $p > 0.05$ )

$= 0.38, p = 0.75, \epsilon_{GG} = .22, \eta_p^2 = .01$ . These two results suggested that the effect of vision on reaction time was consistent across the 15 effective estimates. If we consider just this characteristic of the tasks, then all 15 estimates of movement initiation are comparable.

**Reliability**

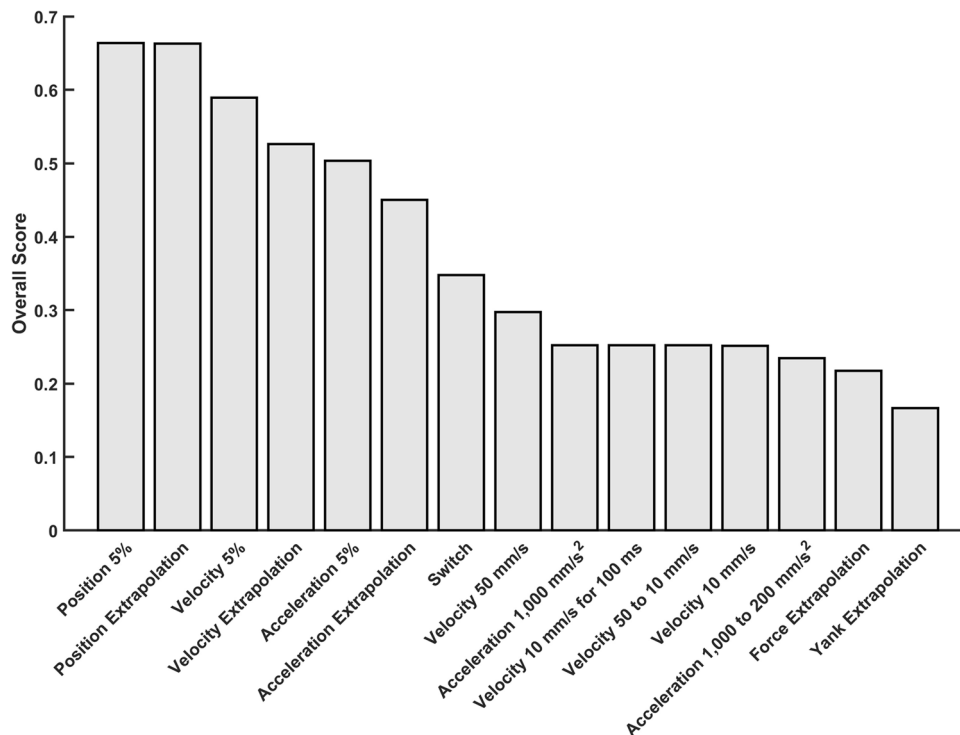
The reliability of the 15 effective estimates of reaction time were calculated in the four conditions (2 Vision [vision, no vision] by 2 Movement Amplitude [short, long]). The most relevant results are from the D studies, which determined the minimum number of trials to achieve the conventional level of reliability (generalizability coefficient [ $g \geq .80$ ]). For the 15 estimates, the number of trials for .80 reliability ranged from 17 to 9 (Table 2). Movement amplitude had a small effect on reliability, with 12 or 13 trials required for the conventional level of reliability for long- or short-amplitude movements, respectively. The availability of visual feedback during movement execution also had a small effect on reliability, with 13 or 12 trials required with vision or no vision, respectively. If we consider just reliability, then the best estimates are position extrapolation and position 5%, which both required nine trials. The worst estimate is acceleration 5%, which required 17 trials.

**Overall scores**

Figure 3 shows the overall scores for the 15 effective estimates of movement initiation, which are interpreted in the next section.

**Table 2** The minimum number of trials to achieve the conventional level of reliability in descending order

Estimate	D study result
Acceleration 5%	17
Velocity 5%	15
Velocity extrapolation	14
Acceleration 1000 to 200 mm/s <sup>2</sup>	13
Yank extrapolation	13
Velocity 10 mm/s	13
Velocity 10 mm/s for 100 ms	13
Velocity 50 to 10 mm/s	13
Acceleration extrapolation	12
Force extrapolation	12
Acceleration 1000 mm/s <sup>2</sup>	12
Switch	11
Velocity 50 mm/s	11
Position 5%	9
Position extrapolation	9



**Fig. 3** Overall scores in descending order. *Note.* The lower the overall score, the better the estimate of movement initiation

## Discussion

Before we delve into interpretations of the overall scores, it is important to keep in mind how they were calculated. We began with 19 estimates of movement initiation that were calculated for each trial. The proportion of manually corrected trials after visual inspection was used to classify the estimates as either effective or ineffective. Four estimates (force 5%, yank 5%, acceleration 200 mm/s<sup>2</sup>, acceleration 200 mm/s<sup>2</sup> for 100 ms) were classified as ineffective because more than 5% of the trials required manual corrections. Overall scores were calculated for the remaining 15 effective estimates. The overall scores were based on an equal weighting of three criteria: time, immunity to the movement amplitude, and reliability. Responsiveness to visual feedback was excluded because all estimates were comparable on that criterion. Other researchers might weigh the three criteria differently or include or exclude other criteria. Therefore, the Excel spreadsheet that generated the overall scores is available on Open Science Framework. The weights of the criteria can be changed, and the overall scores will be automatically recalculated. We calculated overall scores based on an equal weighting of three criteria because we believe that an argument could be made for their relevance to most reaction time studies.

## Overall scores: Winners

The two best overall estimates of movement initiation were yank and force extrapolation. They both had excellent time and immunity to the movement amplitude and fair reliability. Yank extrapolation had a better overall score than force extrapolation because it had the shortest estimate of reaction time (Fig. 1). The reliability of force and yank extrapolation could be improved by a better force sensor and amplifier with a higher signal-to-noise ratio. It would be interesting to compare the inexpensive micro load cell and amplifier in the present experiment to a more expensive load cell and input conditioner. We will pursue that possibility in a future experiment. The extrapolation method (Brenner & Smeets, 2019; Oostwoud Wijdenes et al., 2014) is the newest method we investigated. As a new method, it may have the most potential for improvement. Different thresholds could be tested; for example, the 25 and 75% thresholds could be compared to 20 and 80% or even asymmetric thresholds of 25 and 80%.

The third best overall estimate was acceleration 1000 to 200 mm/s<sup>2</sup>. It had excellent time and immunity to the movement amplitude and fair reliability, similar to force and yank extrapolation. Acceleration 1000 mm/s<sup>2</sup> had a similar overall score, but it achieved that score with trade-off between time

and reliability; acceleration 1000 to 200 mm/s<sup>2</sup> had shorter reaction time but acceleration 1000 mm/s<sup>2</sup> had better reliability. The reliability of the estimates based on acceleration could be improved by measuring acceleration with an accelerometer rather than the second derivative of position (Hansen et al., 2007). We did place a three-axis accelerometer on the participant's index finger, but we did not properly track the orientation of the accelerometer with at least two motion capture LEDs. This prevented us from being able to align the data from the accelerometer to the coordinate system of the motion capture system.

There was a four-way tie for the fourth best overall estimate. Velocity 10 mm/s, velocity 50 to 10 mm/s, velocity 10 mm/s for 100 ms, and acceleration 1000 mm/s<sup>2</sup> had good time, excellent immunity to the movement amplitude, and fair reliability. The three absolute estimates based on velocity 10 mm/s had similar scores on the three criteria. If we consider efficacy as a tiebreaker, then we recommend either velocity 50 to 10 mm/s or velocity 10 mm/s for 100 ms over velocity 10 mm/s because they required less manual corrections (0.10% and 0.29% vs. 3.24%). The fifth best overall estimate was velocity 50 mm/s, but we recommend either velocity 50 to 10 mm/s or velocity 10 mm/s for 100 ms with their slightly better overall scores.

The sixth best overall estimate of movement initiation was the switch. It had poor time, excellent immunity to the movement amplitude, and good reliability. Estimating movement initiation with a push-button switch is simple, neither motion capture nor force measurement equipment is necessary. However, the switch had the third longest estimate of reaction time, besting only position 5% and position extrapolation. The reaction time of the switch could be improved by using a switch with more sensitive operating characteristics. Recall that the switch in the present experiment had an operating force of 2.55 N and a pretravel of 0.3 mm. We are curious whether a switch could be sourced with lower operating force, shorter travel, or both. This could improve the switch's estimate of movement initiation while retaining excellent immunity and good reliability. We will test this hypothesis in Experiment 2.

In summary, our top recommendation is yank extrapolation, which had the best overall score. Most studies do not collect kinetic data, and so in the more common situation with kinematic data, our second recommendation is acceleration 1000 to 200 mm/s<sup>2</sup>.

### Overall scores: Losers

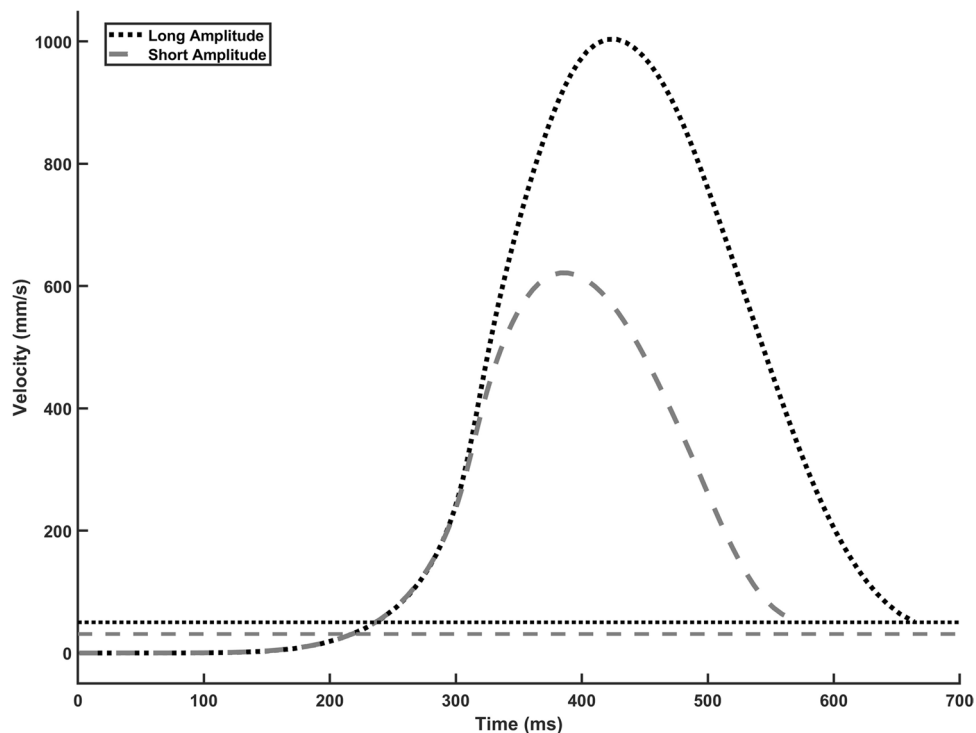
The relative estimates (5% and extrapolation) based on kinematics (position, velocity, and acceleration) had the six worst overall scores. The overall scores were better for the higher derivatives and extrapolation outperformed the 5% method. However, we do not recommend any of the relative

estimates because of their poor overall scores. Velocity 5% is, unfortunately, a common estimate of movement initiation in research. Our results suggested that it had fair time and immunity to the movement amplitude and poor reliability. The 5% method could be improved by decreasing the relative threshold (e.g., velocity 1%, velocity 1% for 100 ms, velocity 5 to 1%). We suspect that, even with improvements, estimates based on 5% of kinematics are unlikely to rival the best overall estimates. For the estimates based on extrapolation of kinematics, it is, perhaps, unsurprising that they had poor overall scores, as this technique was originally developed to detect online corrections (Oostwoud Wijdenes et al., 2014). However, as we previously mentioned, this new method may have the most potential for improvement.

### Immunity to the movement amplitude

Two of the criteria require further discussion: immunity to the movement amplitude and reliability. For immunity, all the absolute estimates had comparable reaction time for long- and short-amplitude movements (Fig. 2). For the relative estimates (5% and extrapolation), there was a relationship between reaction time and the size of the reaction time difference between long- and short-amplitude movements; specifically, the later the estimate of movement initiation, the larger the reaction time difference. In fact, there was 83.3% shared variance between movement initiation and reaction time difference of the relative estimates. This relationship is logical because the later the estimate of movement initiation, the more that estimate falls into the movement time. And the more it falls into movement time, the more it will be affected by the movement amplitude.

It is important to note that the two relative estimates with the shortest reaction time, force and yank extrapolation, had comparable reaction time for short- and long-amplitude movements, and so they were immune to the movement amplitude. The remaining six relative estimates were biased by the movement amplitude. This bias is surprising for at least two reasons. First, relative estimates are supposed to normalize the different profiles of short- and long-amplitude movements (Lacquaniti & Soechting, 1982; Teasdale et al., 1993). Second, the absolute thresholds, which do not attempt to normalize different profiles, had comparable reaction time for short- and long-amplitude movements. We investigated the role of movement amplitude on absolute and relative thresholds by plotting grand mean velocity profiles for long- and short-amplitude movements (Fig. 4). Note that the velocity profiles are shown from the go signal to movement termination. Importantly, the initial rise in velocity is comparable for long- and short-amplitude movements. This was also the case for the initial change in position, acceleration, force, and yank. The consequence of this is that long- and short-amplitude movements will cross an absolute



**Fig. 4** Velocity threshold (5%) applied to the spatiotemporal averages. *Note.* The dotted line (50.2 mm/s) and dashed line (31.1 mm/s) are 5% of maximum velocity from the long- and short-amplitude movements, respectively

threshold at about the same time. In other words, absolute thresholds will have comparable reaction time for long- and short-amplitude movements. Figure 4 shows the bias in the velocity 5% estimate, a relative threshold. The dotted and dashed lines are 5% of maximum velocity for the long- and short-amplitude movements, respectively. The 5% threshold of the long-amplitude movement is exceeded after the 5% threshold for the short-amplitude movement, which results in longer reaction time for long-amplitude movements.

The take home message is that the absolute thresholds were immune to the movement amplitude whereas relative thresholds were biased by it. This was surprising because relative threshold should normalize movements with different kinematic or kinetic profiles. However, the initial change in kinetics and kinematics was comparable for long- and short-amplitude movements and so a normalization was not needed and did more harm than good. Relative thresholds would be important when the kinematic or kinetic profiles are different around movement initiation. This could occur when there is a wider range of movement characteristics or a larger difference between participants; for example, when comparing the movements of young adults and older adults. We recommend checking whether there are differences around movement initiation that bias the estimates with the best overall scores, like yank extrapolation and acceleration 1000 to 200 mm/s<sup>2</sup>. If there are differences, then a relative threshold might be useful.

A quick note on the responsiveness to visual feedback, which was excluded from the overall scores because all estimates of movement initiation were comparably longer (17.7 ms, [11.3, 24.0]) without visual feedback during movement execution. We expected reaction time to be longer without visual feedback. The challenge with this criterion is that we did not have an exact prediction of how much longer reaction time should be without visual feedback. Imagine that one estimate of movement initiation was 10 ms longer without visual feedback and another estimate was 20 ms longer. In this hypothetical situation, we would be unable to say which estimate was closer to the true increase in reaction time without visual feedback. For this reason, we discourage the use of responsiveness to visual feedback as a criterion in future research unless the increase in reaction time can be estimated more accurately.

## Reliability

The other criterion that requires further discussion is reliability. It is possible to compare the results from the present study to one of our previous studies. In Blinch et al. (2018), we reported that a similar two-choice reaction time task to long-distance (20 cm) targets required six trials to achieve the conventional level of reliability. If we isolate the similar condition in the current study (switch, two-choice,

vision, long-distance [20 cm] target), then 13 trials were required to achieve conventional reliability. One difference in our previous study is that participants made reaching and pointing movements with a stylus. The push-button switch that estimated movement initiation was inside the stylus and actuated by pressing the tip. It is possible that the different hardware caused the different results of the D studies. It is also possible that six or 13 trials is normal inter-experiment variability of D studies, even with similar tasks. We, therefore, recommend collecting more than the minimum number of trials suggested by the D studies in case the data are less reliable than previous datasets. It is far better to overestimate the number of trials and exceed the conventional level of reliability than it is to underestimate it. Also, some reaction time analyses require more trials than suggested by D studies; for example, it is suggested that at least 100 trials per participant and condition are needed to fit a distribution function to the reaction time distribution of each participant and condition (Lacouture & Cousineau, 2008).

### Comparisons to Brenner and Smeets (2019)

Let us now compare the results of the current experiment to a reaction time methodology experiment by Brenner and Smeets (2019). They examined an auditory simple reaction time task where the movement was lifting the right index finger off a force sensor. There were two conditions: one where the movement of the finger was unconstrained and one where the movement was constrained by a straw lying loosely above the finger that should not be hit. These conditions tested the immunity of the estimates to the movement constraints, which shares some overlap with our criterion of immunity to the movement amplitude. Brenner and Smeets compared 182 estimates of movement initiation. Based on three unique criteria, they concluded that the best overall estimate was force extrapolation on the average raw data for each participant and condition.

Both the present experiment and the study by Brenner and Smeets (2019) included a switch (real or emulated), kinematic data (position, velocity, and acceleration), and kinetic data (force and yank). Both studies also used 5%, extrapolation, and double-threshold detectors. We included nine absolute thresholds, whereas their only absolute threshold was an emulated switch (based on a 0.5 mm vertical finger displacement threshold). Brenner and Smeets included baseline thresholds, which was the kinematic or kinetic data exceeded a baseline mean plus three standard deviations. We did not include baseline thresholds because we forgot to record kinematic and kinetic data for at least 100 ms before the go signal. The scope of the study by Brenner and Smeets was larger because they also considered three methods of smoothing (Butterworth filtered, polynomial smoothed, and raw data) and three methods of averaging reaction time

(mean, median, and averaging the raw data across trials for each participant and condition). We only considered Butterworth-filtered data and using the mean. Despite these differences, there were some common conclusions; specifically, that higher derivatives have shorter reaction time, kinetic estimates have shorter reaction time than their kinematic counterparts, and reaction time increases from double-threshold detectors, to 5%, to extrapolation estimates.

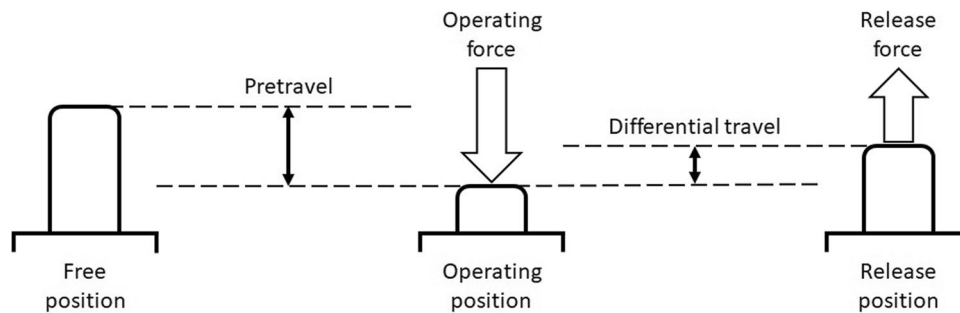
Two important differences between these studies were the criteria used to evaluate the estimates of movement initiation and the tasks. Recall that our criteria were time, immunity to the movement amplitude, and reliability. Brenner and Smeets' (2019) criteria were time, immunity to the movement constraint, and reliability. Immunity to the movement amplitude and constraint are both related to response complexity, but they are not identical. Both experiments evaluated reliability but with different operational definitions. We quantified reliability as the minimum number of trials to achieve the conventional level of reliability. Brenner and Smeets measured reliability with bootstrapping and the sensitivity to smoothing techniques. All three are valid ways to quantify reliability, but they are unique aspects of reliability that are unlikely to share much overlap.

Finally, the tasks were different; we tested a two-choice, visual reaction time task with a movement to a short- or long-distance target and they tested a simple, auditory reaction time task. Reaction time is shorter and less variable in simple than two-choice reaction time because the movement can be selected and prepared before the go signal. We predict that the best overall estimates of movement initiation in two-choice would also perform well in simple reaction time, but that should be tested in a single study. In summary, the current experiment and the study by Brenner and Smeets (2019) are valid and unique investigations of reaction time methodology with some overlap and shared conclusions. Future research could focus on finding principles of reaction time methodology that generalize across tasks.

### Reaction time vs. premotor reaction time

The duration of cognitive processing can be measured by reaction time or premotor reaction time. Premotor time is a shorter estimate of cognitive processing because it excludes motor time, and yet, reaction time is more commonly used in psychology. There are at least three reasons for this. First, it is simple and inexpensive to measure reaction time with a switch. Measuring reaction time with a load cell requires a few more pieces of equipment. The cost and complexity of EMG recording equipment is similar to a motion capture system. Second, and perhaps most importantly, motor time is typically constant for similar movements (Botwinick & Thompson, 1966; Christina & Rose, 1985); for example, motor time is unaffected by movement amplitude (Lagasse &





**Fig. 5** Push-button switch operating characteristics

Hayes, 1973). As motor time is typically constant, reaction time overestimates premotor time by a constant amount. In other words, reaction time and premotor time, despite different magnitudes, should yield the same experimental effects. Although motor time is typically comparable for similar movements, it is not always constant; one example is that it increases with the inertia of the body segment being moved (Anson, 1989). If a difference in motor time between conditions is suspected, then reaction time should be fractionated into premotor and motor time.

The third reason that reaction time is more commonly measured than premotor time is that estimating the onset of EMG activity is far more complex than estimating movement initiation. That is because EMG signals are far more complex than movement kinematics and kinetics. EMG signals require sophisticated signal conditioning, like an adaptive pre-whitening filter (e.g., Staude et al., 2001) or a Teager-Kaiser energy operator (e.g., Solnik et al., 2010). Detecting EMG onset then often involves advanced statistical methods like the approximated generalized likelihood ratio (e.g., Staude et al., 2001), wavelet transform (e.g., Vannozzi et al., 2010), or sample entropy (e.g., Zhang & Zhou, 2012). Furthermore, EMG signals must be monitored “to exclude premature triggering on background EMG, especially in tense normal subjects or in patients with hypertonina or Parkinson’s disease” (Tomberg et al., 1991; p. 277). In this regard, switching from reaction time to premotor time is out of the frying pan and into the fire.

## Conclusions

There are many ways to estimate movement initiation on each trial, 19 of which were investigated in the current experiment. The 15 effective estimates had a wide range to their reaction time, immunity to the movement amplitude, and reliability. None of the estimates were the best on all three of these criteria. We recommend the three estimates with the best overall scores: yank extrapolation, force extrapolation, and acceleration 1000 to 200 mm/s<sup>2</sup>. We do not recommend

relative estimates based on kinematics (5% and extrapolation) because they had the six worst overall scores. The switch had a mediocre overall score. It did have excellent immunity to the movement amplitude and good reliability, but its Achilles’ heel was being a late estimate of movement initiation (poor time). We thought it was worthwhile to attempt to improve the time of the switch because it is likely the most common way to measure the duration of cognitive processing. We attempted to improve the overall score of the switch in Experiment 2 by using a micro push-button switch with better, more sensitive operating characteristics.

## Experiment 2

The goal of the present experiment was to determine whether a micro push-button switch with better, more sensitive operating characteristics could improve the estimate of reaction time. An understanding of the operating characteristics of push-button switches is vital for the current experiment. The starting position of a switch is called the free position (Fig. 5). A switch can be depressed downward to the operating position, which is where the switch actuates from an open state to a closed state. Pretravel is the distance between the free position and the operating position, and operating force is the force required to move the switch from the free position to the operating position. The switches in Experiments 1 and 2 could not be depressed past the operating position, but that is not the case for all switches.

When a switch is released, it moves up from the operating position to the release position, which is where the switch actuates from closed to open. Differential travel is the distance between the operating position and the release position, and release force is the force required to move the switch from the operating position to the release position. Differential travel and release force are typically unreported in manufacturer datasheets. However, differential travel and release force are smaller than pretravel and operating force, respectively.

The home switch in Experiment 1 had an operating force of 2.55 N and a pretravel of 0.3 mm. The release force was 0.49 N, and the differential travel was unreported by the manufacturer. In the current experiment, we compared switches with the same operating characteristics as Experiment 1 to the most sensitive switches we could find. These new micro push-button switches had an operating force of 0.5 N and a pretravel of 0.2 mm. The release force and differential travel were unreported by the manufacturer, but they were smaller than the operating force and pretravel. The new switches required less force and had shorter travel than the switches from Experiment 1. We, therefore, predicted that the new switches would have shorter estimates of reaction time compared to the switches from Experiment 1 because they were more sensitive.

## Method

### Participants

Thirty volunteer participants (17 female participants and 13 male participants) were recruited from the university community. The age of participants ranged from 19 to 39 years old ( $M = 23.3$ ,  $Mdn = 21.5$ ,  $SD = 5.13$ ). Our goal was to test at least 30 participants and as many as possible during the fall 2021 semester. A sensitivity analysis suggested that data from 30 participants would allow us to detect differences (two tailed) between switches with at least medium effect sizes ( $d_z = 0.53$ ; paired samples  $t$  test,  $\alpha = .05$ ,  $1 - \beta = .80$ ).

### Design

Participants completed a two-choice reaction time task in only two conditions. In one condition, responses were made with standard micro push-buttons switches. In the other condition, responses were made with sensitive micro push-button switches. The order of the two conditions was counterbalanced across participants. Each condition consisted of 16 practice trials and 108 test trials. More than 100 test trials were included because at least 100 test trials per participant and condition are necessary to perform a reaction time distribution analysis by fitting ex-Gaussian distribution functions (Lacouture & Cousineau, 2008). The correct response for half of the trials was a left response; the correct response for the other half of trials was a right response. The order of left and right responses was randomized, and a different randomization was used for each condition and participant. At the end of the experiment, the participant was asked whether they preferred using the standard switches or the sensitive switches.

## Apparatus

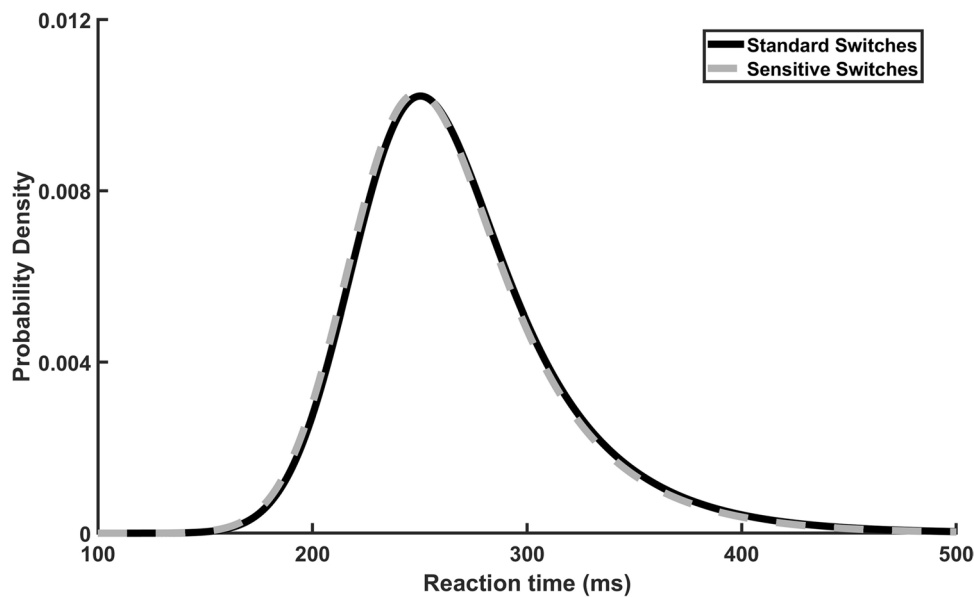
The participant was seated at a table with a button box (20.4 cm length, by 10.2 cm width, by 4.3 cm height) and a stimulus box (20.4 cm length, by 7.8 cm width, by 7.8 cm height) on the surface of the table. The top of the stimulus box was at 45° to face the participant. The participant was seated so that their midsagittal plane was centered with the button and stimulus boxes. The distance from the front edge of the table to the middle of the button box was 15 cm. The distance from the middle of the button box to the middle of the stimulus box was 28 cm. There were two button boxes, one for each condition. Both button boxes consisted of two micro push-button switches that were 10 cm apart.

One button box had standard micro push-button switches with an operating force of 2.55 N, a pretravel of 0.3 mm, a release force of 0.49 N, and a diameter of 7.1 mm (Omron B3F-4005). These switches had the same operating characteristics as the home switch in Experiment 1. We decided to use the Omron switches in this experiment because the home switch in Experiment 1 was a low-cost imitation of the Omron switches. The other button box had sensitive micro push-button switches with an operating force of 0.50 N, a pretravel of 0.2 mm, and a diameter of 3.0 mm (Panasonic EVQQ2D03W). These Panasonic switches had the lowest operating force and the shortest pretravel that we could find.

The stimulus box consisted of three diffused LEDs. The left and right LEDs were red and 10 cm apart. In the middle of these LEDs was a blue fixation LED. A data acquisition device (National Instruments, USB-6002) recorded the voltage signals of the switches and the left and right LEDs at 1000 Hz; it also controlled the illumination of the LEDs.

## Procedures

The participants started each trial by pressing and holding down the left and right micro push-button switches with their left and right index fingers, respectively. The middle, blue LED then illuminated for 500 ms as a warning signal and as a fixation. This was followed by a 1–2 s variable foreperiod before either the left or right, red LED illuminated as the go signal. The participant was instructed to lift the finger on the same side as the go signal as quickly as possible. A finger lift in response to the go signal was used instead of a finger press to be consistent with the initial finger movement in Experiment 1. The target LED was turned off after the 1.1-s recording interval of the data acquisition device. The participant could then begin the next trial by pressing and holding down the left and right switches. Test trials with anticipation (reaction time less than 100 ms), inattention (reaction time greater than 1000 ms), or incorrec-



**Fig. 6** Mean ex-Gaussian distribution functions for the standard and sensitive switches

responses were recycled to the end of the block, and the original trials were excluded from data analysis. Out of all the test trials, 1.89% were excluded; the overwhelming majority (74.4%) of these trials were when both switches were released.

### Data analysis

The time of the go signal was determined by when the voltage signal to illuminate the target LED first exceeded 2.5 V. Similarly, movement initiation was estimated by when the voltage signal from the switch release first fell below 2.5 V. Reaction time was the interval between the go signal and movement initiation. A reaction time distribution analysis was performed with the method described by Lacouture and Cousineau (2008). In short, an ex-Gaussian distribution function was fitted to the reaction time data of each participant and condition (standard and sensitive switches). An ex-Gaussian distribution function is described by three parameters:  $\mu$  (mu), the mean of the Gaussian component,  $\sigma$  (sigma), the standard deviation of the Gaussian component, and  $\tau$  (tau), the mean and standard deviation of the exponential component. The parameters of the ex-Gaussian distribution functions in the conditions with standard and sensitive switches were compared with paired samples *t* tests.

Besides reaction time, we also compared the efficacy and reliability of the standard and sensitive switches. The same overall score as Experiment 1 could not be calculated because the task was a switch release without a subsequent short- or long-amplitude movement; thus, immunity to the movement amplitude was not applicable. Reported values

were means and 95% difference- and correlation-adjusted confidence intervals (Cousineau, 2017). Effect sizes were reported using Cohen's  $d_z$  (Lakens, 2013).

### Results

The mean ex-Gaussian distribution functions for the standard and sensitive switches were visibly similar (Fig. 6). This was supported by paired samples *t* tests that suggested that the  $\mu$ ,  $\sigma$ , and  $\tau$  parameters were comparable in both conditions (Table 3). When the participants were asked which switches they preferred, 14 participants preferred the standard switches, 15 preferred the sensitive switches, and one participant had no preference. Both types of switches required no corrections after visual inspection (0%), which is perfect efficacy. As for reliability, the standard switches required ten trials to achieve the conventional level of reliability. The sensitive switch had better reliability with only seven trials.

### Discussion

Our hypothesis that the sensitive micro push-button switches would have shorter estimates of reaction time compared to the standard switches was not supported. Even though the sensitive switches had lower operating force and shorter travel, the reaction time distribution functions were comparable. There might not have been a difference in the reaction time of the switches because the differences in the operating

**Table 3** Parameters of the ex-Gaussian distribution functions for the standard and sensitive switches

Parameter	Switches		Statistics
	Standard	Sensitive	
$\mu$	226 ms, [215, 237]	225 ms, [214, 236]	$t(29) = 0.19, p = 0.85, d_z = 0.03$
$\sigma$	26.0 ms, [21.9, 30.1]	26.4 ms, [22.3, 30.5]	$t(29) = -0.18, p = 0.86, d_z = -0.03$
$\tau$	40.7 ms, [33.6, 47.9]	39.3 ms, [32.2, 46.4]	$t(29) = 0.41, p = 0.68, d_z = 0.08$

Error bars are 95% difference- and correlation-adjusted confidence intervals (Cousineau, 2017)

characteristics were small: the difference in operating force was 2.05 N and the difference in pretravel was 0.1 mm. However, the relevant characteristics for releasing a switch are release force and differential travel. These values were unreported by the manufacturers, but the values would be smaller than the operating force and pretravel. This would result in smaller differences in release force and differential travel than operating force and pretravel, respectively.

Overall scores for the standard and sensitive switches could not be calculated in the present experiment because immunity to the movement amplitude was not applicable to the current task. However, it was possible to compare the efficacy and reliability of the standard and sensitive switches. Both types of switches had perfect reliability, which was also seen for the switch in Experiment 1. As for reliability, the standard switch required ten trials to achieve the conventional level of reliability. This was similar to the reliability of 11 trials in Experiment 1 for a comparable switch. The sensitive switch had better reliability with only seven trials. This was even better than the most reliable estimates in Experiment 1, which were position extrapolation and position 5% with nine trials. Thus, the sensitive switches are slightly better overall estimate compared to the standard switches because of their better reliability. Recall that the comparable standard switch in Experiment 1 was immune to the movement amplitude. We predict that this would also be the case for the sensitive switches as reaction time was comparable for the standard and sensitive switches. Therefore, we recommend the use of sensitive micro push-button switches as a slightly better overall estimate than standard micro push-button switches.

Returning to reaction time, the results from the current experiment suggest that reaction time was comparable for the standard and sensitive micro push-button switches. It is possible that reaction time could be longer for switches with higher force or longer travel than the standard switches; for example, Omron model B3F-1006 has an operating force of 4.9 N and a release force of 0.7 N. Similarly, reaction time could be shorter for switches with lower operating force or shorter travel than the sensitive switches. Although, we have yet to find a switch that is more sensitive than the ones in the present experiment. We did not test whether there is a difference in reaction time

for pressing the standard and sensitive switches. It is possible that reaction time could be longer for the standard switches when pressing but not releasing because operating force and pretravel (press characteristics) are larger than release force and differential travel (release characteristics). If a researcher is planning to estimate reaction time based on the press of a micro push-button switch, then this methodological issue should be addressed.

Another common way to estimate reaction time is with a keyboard switch. A typical mechanical keyboard switch has an operating force of 0.55 N and a pretravel of 2.0 mm (Cherry MX Brown). This mechanical keyboard switch also has 2.0 mm of overtravel, which means that the switch can be depressed 2.0 mm past the operating position. When comparing the release of the sensitive micro push-button switch to this typical mechanical keyboard switch, the release force is similar (0.50 vs. 0.55 N). The sensitive switch would move the differential travel (< 0.2 mm) before it actuates, but the keyboard switch would move the overtravel (2.0 mm) and then the differential travel (< 2.0 mm). We predict that this difference of at least 2.0 mm would cause a longer reaction time to release a keyboard switch compared to the sensitive switch. The reaction time to press this keyboard switch would probably also be longer than to press the sensitive switch. In this case, the operating force is, again, similar but the pretravel is 1.8 mm longer for the keyboard switch (0.2 vs. 2.0 mm).

## General discussion

The purpose of the present study was to identify and compare the tradeoffs of 19 common estimates of movement initiation. These estimates were applied to individual trial data. The criteria were the duration of reaction time, immunity to the movement amplitude, reliability, and the number of manually corrected trials (efficacy). In Experiment 1, the three effective estimates with the best overall scores were yank extrapolation, force extrapolation, and acceleration 1000 to 200 mm/s<sup>2</sup>. Our highest recommendation is for yank extrapolation because it had the best overall score. If one does not collect kinetic data, then we recommend acceleration 1000

to 200 mm/s<sup>2</sup>. Velocity 10 mm/s and velocity 50 to 10 mm/s are also good options that had slightly lower overall scores compared to acceleration 1000 to 200 mm/s<sup>2</sup>.

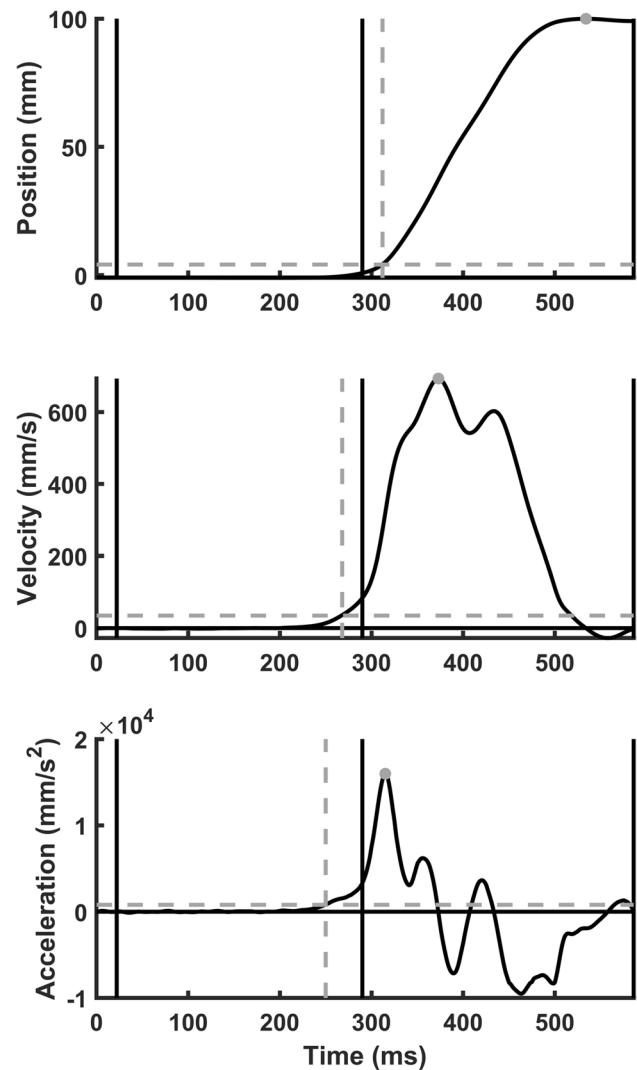
The switch had a mediocre overall score that we attempted to improve in Experiment 2. The biggest issue with the switch was that it was a late estimate of movement initiation. We tried to improve that by testing a micro push-button switch with better, more sensitive operating characteristics. The standard and sensitive switches had comparable reaction time despite differences in operating characteristics. Importantly, the sensitive switch had better reliability than the standard switch. Thus, we recommend the use of a sensitive micro push-button switch as a slightly better overall estimate than a standard micro push-button switch. The sensitive switch is unlikely to rival the best overall estimates because of its poor time, but it does have three advantages. First, a micro push-button switch is simple and inexpensive, especially compared to other estimates that rely on trajectory analysis and a motion capture system. Second and third, it had the best efficacy and reliability of all the estimates. Therefore, a sensitive micro push-button switch is a decent way to estimate the duration of cognitive processing. Importantly, we did not test push-button switches (not micro) or keyboard switches, but we predict they would be worse estimates of movement initiation because of their longer pretravel and higher operating force.

To summarize, we recommend measuring reaction time on each trial with one of the three best overall estimates of movement initiation: yank extrapolation, force extrapolation, or acceleration 1000 to 200 mm/s<sup>2</sup>. Future research could attempt to improve these estimates with better or new algorithms. These estimates could also be improved with time series data with higher signal-to-noise ratios. This could be achieved with better data acquisition technology or offline processing (e.g., Franks et al., 1990; Hansen et al., 2007). It might also be worthwhile to investigate other methods of estimating movement initiation like force-sensing resistors, capacitive touch sensors, infrared break-beam sensors, or opening a circuit between a conductive plate and a stylus (e.g., Fitts, 1954). One final note on practicality; a researcher may not have the luxury of choosing one of the three best overall estimates. This could occur when the experiment involves an established protocol. In this case, we suggest measuring and analyzing reaction time twice: first with the established method and then with one of the best overall estimates. This two-step approach could allow established protocols to transition to better estimates of movement initiation.

## Appendix

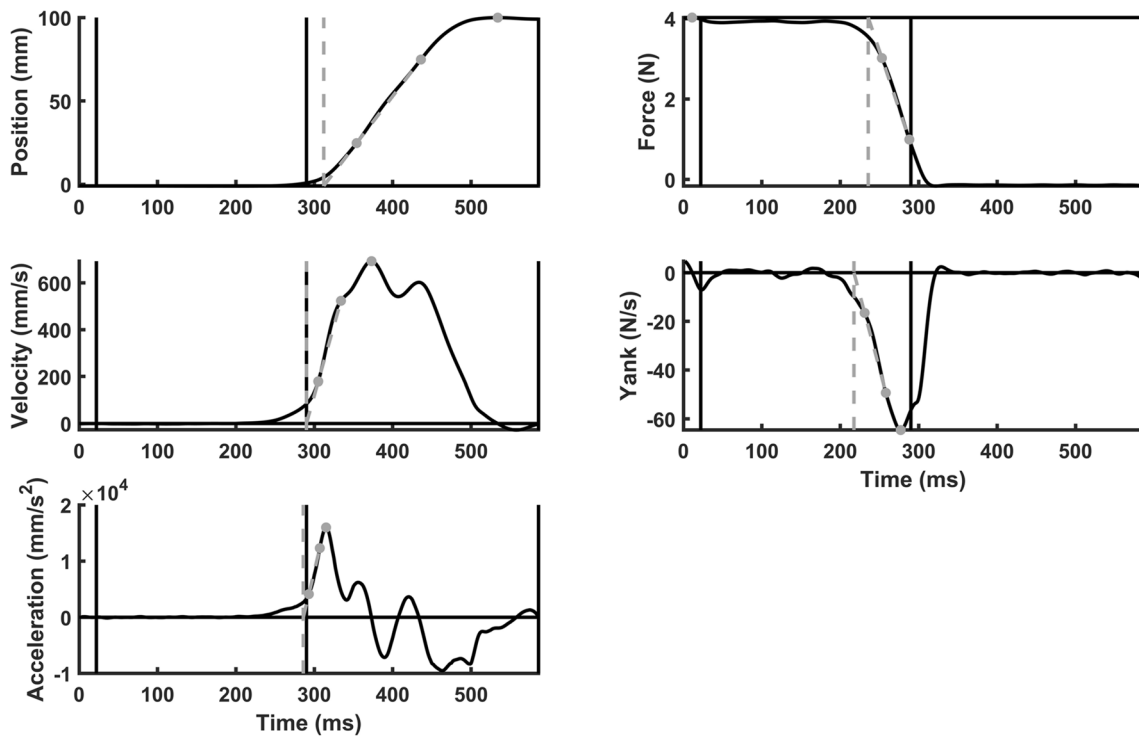
### Examples of the effective estimates of movement initiation

The figures below illustrate several estimates of movement initiation for one representative trial (participant 36, vision, block 1, trial 28). Each graph has three black vertical lines, which represent, in order, the go signal, movement initiation estimated by the switch, and movement termination.



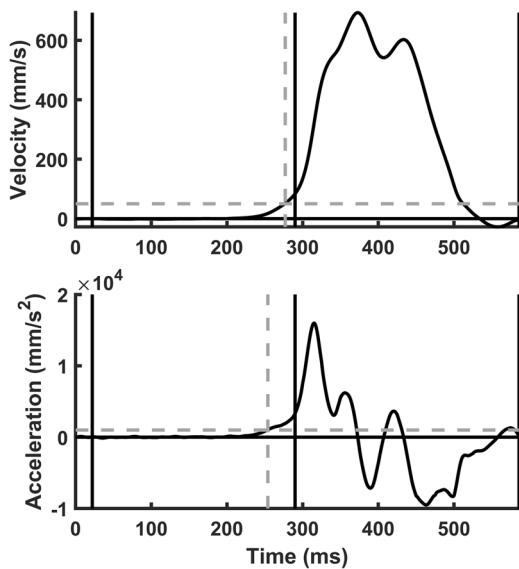
**Fig. 7** Relative estimates of movement initiation based on 5% of maximum kinematics. *Note.* The grey dot shows the maximum of each kinematic, and the horizontal dashed line is 5% of that maximum. The vertical dashed line is when the kinematic data first exceeds the 5% thresholds.



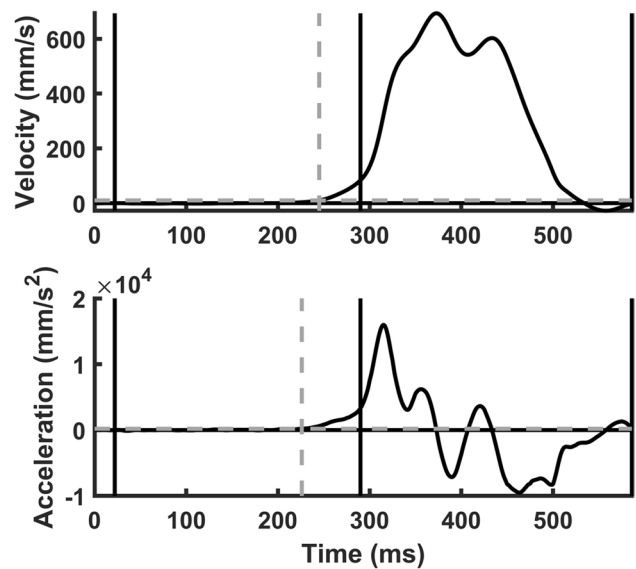


**Fig. 8** Relative estimates of movement initiation based on the extrapolation of kinematics or kinetics. *Note.* The *grey dots* show 25% of maximum, 75% of maximum, and maximum kinematics or kinetics. The *sloped dashed line* goes from 75% to 25% and then extrapolates

back to the baseline. The *dashed horizontal line* is the intersection of the *sloped dashed line* and *baseline*, which is the estimate of movement initiation.



**Fig. 9** Absolute estimates of movement initiation based on 50 mm/s and 1000 mm/s<sup>2</sup>. *Note.* The *horizontal dashed line* in the top and bottom graphs is 50 mm/s and 1000 mm/s<sup>2</sup>, respectively. The *vertical dashed line* is when the kinematic data first exceeds that threshold.



**Fig. 10** Absolute estimates of movement initiation based on 10 mm/s and 200 mm/s<sup>2</sup>. *Note.* The *horizontal dashed line* in the top and bottom graphs is 10 mm/s and 200 mm/s<sup>2</sup>, respectively. The *vertical dashed line* is when the kinematic data first exceeds that threshold.

**Acknowledgments** Our thanks to two anonymous reviewers for their insightful critiques.

**Data availability** The data and statistical analysis files for the present study are available at <https://osf.io/RBYNF/>. This experiment was not preregistered.

**Code availability** G1.m is available for download from <https://people.ok.ubc.ca/briocomm/gtheory/gtheory.html>. There are also versions for SAS, SPSS, and R.

## Declarations

**Ethics approval** The Human Research Protection Program at Texas Tech University approved the study. 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 for publication** Not applicable.

**Conflict of interest/Competing interests** The authors have no relevant financial or non-financial interests to disclose.

## References

- Anson, J. G. (1989). Effects of moment of inertia on simple reaction time. *Journal of Motor Behavior*, 21(1), 60–71.
- Binsted, G., & Elliott, D. (1999). Ocular perturbations and retinal/extraretinal information: The coordination of saccadic and manual movements. *Experimental Brain Research*, 127(2), 193–206.
- Blinch, J., Kim, Y., & Chua, R. (2018). Trajectory analysis of discrete goal-directed pointing movements: How many trials are needed for reliable data? *Behavior Research Methods*, 50(5), 2162–2172.
- Blinch, J., Homes, J., Cameron, B. D., & Chua, R. (2021). Investigating information processing of the bimanual asymmetric cost with the response priming technique. *Journal of Experimental Psychology: Human Perception and Performance*, 47(5), 673–688.
- Bonato, P., D'Alessio, T., & Knafitz, M. (1998). A statistical method for the measurement of muscle activation intervals from surface myoelectric signal during gait. *IEEE Transactions on Biomedical Engineering*, 45(3), 287–299.
- Botwinick, J., & Thompson, L. W. (1966). Premotor and motor components of reaction time. *Journal of Experimental Psychology*, 71(1), 9.
- Brennan, R. L. (2001). *Generalizability theory*. Springer.
- Brenner, E., & Smeets, J. B. J. (2019). How can you best measure reaction times? *Journal of Motor Behavior*, 51(5), 486–495.
- Brown, S. G., Roy, E. A., Rohr, L. E., & Bryden, P. J. (2006). Using hand performance measures to predict handedness. *Laterality*, 11(1), 1–14.
- Carson, R. G., Chua, R., Elliott, D., & Goodman, D. (1990). The contribution of vision to asymmetries in manual aiming. *Neuropsychologia*, 28(11), 1215–1220.
- Christina, R. W., & Rose, D. J. (1985). Premotor and motor reaction time as a function of response complexity. *Research Quarterly for Exercise and Sport*, 56(4), 306–315.
- Chua, R., & Elliott, D. (1993). Visual regulation of manual aiming. *Human Movement Science*, 12(4), 365–401.
- Corcus, D. M., Gottlieb, G. L., & Agarwal, G. C. (1988). Accuracy constraints upon rapid elbow movements. *Journal of Motor Behavior*, 20(3), 255–272.
- Cousineau, D. (2017). Varieties of confidence intervals. *Advances in cognitive psychology*, 13(2), 140–155.
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability of scores and profiles*. Wiley.
- Darling, W. G., Cole, K. J., & Abbs, J. H. (1988). Kinematic variability of grasp movements as a function of practice and movement speed. *Experimental Brain Research*, 73(2), 225–235.
- Englund, M. P., & Patching, G. R. (2009). An inexpensive and accurate method of measuring the force of responses in reaction time research. *Behavior Research Methods*, 41(4), 1254–1261.
- 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.
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 14(3), 340–347.
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381–391.
- Fitts, P. M., & Seeger, C. M. (1953). SR compatibility: spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 46(3), 199.
- Franks, I. M., Sanderson, D. J., & Van Donkelaar, P. (1990). A comparison of directly recorded and derived acceleration data in movement control research. *Human Movement Science*, 9(6), 573–582.
- Gracco, V. L., & Abbs, J. H. (1986). Variant and invariant characteristics of speech movements. *Experimental Brain Research*, 65(1), 156–166.
- Hansen, S., Elliott, D., & Khan, M. A. (2007). Comparing derived and acquired acceleration profiles: 3-D optical electronic data analyses. *Behavior Research Methods*, 39(4), 748–754.
- Hansen, S., Glazebrook, C. M., Anson, J. G., Weeks, D. J., & Elliott, D. (2006). The influence of advance information about target location and visual feedback on movement planning and execution. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 60(3), 200.
- Helmholtz, H. L. F. (1850). Über die Methoden, kleinste Zeittheile zu messen, und ihre Anwendung für physiologische Zwecke. *Philos. Mag.*, 6, 313–325.
- Henry, F. M., & Rogers, D. E. (1960). Increased response latency for complicated movements and a “memory drum” theory of neuro-motor reaction. *Research Quarterly*, 31(3), 448–458.
- Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4(1), 11–26.
- Huynh, H., & Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, 1(1), 69–82.
- Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, 45(3), 188–196.
- Kapoula, Z., & Robinson, D. A. (1986). Saccadic undershoot is not inevitable: Saccades can be accurate. *Vision Research*, 26(5), 735–743.
- Khan, M. A., Elliott, D., Coull, J., Chua, R., & Lyons, J. (2002). Optimal control strategies under different feedback schedules: Kinematic evidence. *Journal of Motor Behavior*, 34(1), 45–57.
- Klapp, S. T. (1995). Motor response programming during simple choice reaction time: The role of practice. *Journal of Experimental Psychology: Human perception and performance*, 21(5), 1015.
- Klapp, S. T. (2003). Reaction time analysis of two types of motor preparation for speech articulation: Action as a sequence of chunks. *Journal of Motor Behavior*, 35(2), 135–150.
- Klapp, S. T., & Maslovat, D. (2020). Programming of action timing cannot be completed until immediately prior to initiation of the response to be controlled. *Psychonomic Bulletin & Review*, 27(5), 821–832.

- Krigolson, O., & Heath, M. (2004). Background visual cues and memory-guided reaching. *Human Movement Science*, 23(6), 861–877.
- Lacouture, Y., & Cousineau, D. (2008). How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. *Tutorials in Quantitative Methods for Psychology*, 4(1), 35–45.
- Lacquaniti, F., & Soechting, J. F. (1982). Coordination of arm and wrist motion during a reaching task. *Journal of Neuroscience*, 2(4), 399–408.
- Lagasse, P. P., & Hayes, K. C. (1973). Premotor and motor reaction time as a function of movement extent. *Journal of Motor Behavior*, 5(1), 25–32.
- Lakens, D. (2013). Calculating and reporting the effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, 4, 863.
- Lin, D. C., McGowan, C. P., Blum, K. P., & Ting, L. H. (2019). Yank: the time derivative of force is an important biomechanical variable in sensorimotor systems. *Journal of Experimental Biology*, 222(18), jeb180414.
- Milner, A. D. (1986). Chronometric analysis in neuropsychology. *Neuropsychologia*, 24(1), 115–128.
- Mushquash, C., & O'Connor, B. P. (2006). SPSS and SAS programs for generalizability theory and analysis. *Behavior Research Methods*, 38(3), 542–547.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113.
- Oostwoud Wijdenes, L., Brenner, E., & Smeets, J. B. J. (2014). Analysis of methods to determine the latency of online movement adjustments. *Behavior Research Methods*, 46(1), 131–139.
- Posner, M. I. (2005). Timing the brain: Mental chronometry as a tool in neuroscience. *PLoS Biology*, 3(2), e51.
- Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, 109(4), 444.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171(3972), 701–703.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366.
- Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, 81(1), 174.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153(3736), 652–654.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 30, 276–315.
- Solnik, S., Rider, P., Steinweg, K., DeVita, P., & Hortobágyi, T. (2010). Teager-Kaiser energy operator signal conditioning improves EMG onset detection. *European Journal of Applied Physiology*, 110, 489–498.
- Staupe, G., Flachenecker, C., Daumer, M., & Wolf, W. (2001). Onset detection in surface electromyographic signals: A systematic comparison of methods. *EURASIP Journal on Advances in Signal Processing*, 2001(2), 1–15.
- Stone, K. D., Bryant, D. C., & Gonzalez, C. L. R. (2013). Hand use for grasping in a bimanual task: Evidence for different roles? *Experimental Brain Research*, 224(3), 455–467.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662.
- Teasdale, N., Bard, C., Fleury, M., Young, D. E., & Proteau, L. (1993). Determining movement onsets from temporal series. *Journal of Motor Behavior*, 25(2), 97–106.
- Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology*, 14(1), 1–36.
- Tomberg, C., Levarlet-Joye, H., & Desmedt, J. E. (1991). Reaction times recording methods: Reliability and EMG analysis of patterns of motor commands. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 81(4), 269–278.
- Vannozzi, G., Conforto, S., & D'Alessio, T. (2010). Automatic detection of surface EMG activation timing using a wavelet transform based method. *Journal of Electromyography and Kinesiology*, 20(4), 767–772.
- Vispoel, W. P., Morris, C. A., & Kilinic, M. (2018). Applications of generalizability theory and their relations to classical test theory and structural equation modeling. *Psychological Methods*, 23(1), 1–26.
- Westwood, D., & Goodale, M. (2003). Perceptual illusion and the real-time control of action. *Spatial Vision*, 16(3), 243–254.
- Wundt, W. (1880). *Grundzüge der physiologischen psychologie* (2nd ed., 2 Vols.). Engelmann.
- Zhang, X., & Zhou, P. (2012). Sample entropy analysis of surface EMG for improved muscle activity onset detection against spurious background spikes. *Journal of Electromyography and Kinesiology*, 22(6), 901–907.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.