Fictitious inhibitory differences:

How skewness and slowing distort the estimation of stopping latencies.

Frederick Verbruggen (University of Exeter, UK)
Christopher D. Chambers (Cardiff University, UK)
Gordon D. Logan (Vanderbilt University, USA)

Psychological Science

Contact information

Frederick Verbruggen (corresponding author)
Psychology, College of Life and Environmental Sciences
Washington Singer Laboratories, Streatham Campus
Exeter, EX4 4QG (United Kingdom)
email: F.L.J.Verbruggen@exeter.ac.uk
phone: +44 1392 725900

Acknowledgments

Verbruggen and Chambers are supported by a grant from the ESRC (ES/J00815X/1). Logan is supported by Grant BCS-0957074 from the National Science Foundation. The authors thank the "*Cognitive Control and Associative Learning*" group at University of Exeter (ccal-exeter.org), Guido Band, and an anonymous reviewer for useful comments on this project.

Abstract

The stop-signal paradigm is a popular method for examining response inhibition and impulse control in psychology, cognitive neuroscience, and clinical domains because it allows the estimation of the covert latency of the stop process (SSRT). In three sets of simulations, we examined to what extent SSRTs that were estimated with the popular *mean* and *integration* methods were influenced by the skew of the reaction time distribution and the gradual slowing of the response latencies. We found that the mean method consistently overestimated SSRT. The integration method tended to underestimate SSRT when response latencies gradually increased. This underestimation bias was absent when SSRTs were estimated with the integration method for smaller blocks of trials. Thus, skewing and response slowing can lead to spurious inhibitory differences. We recommend that the mean method of estimating SSRT be abandoned in favour of the integration method.

Fictitious inhibitory differences

The ability to inhibit planned or ongoing actions is a cornerstone of flexible human behaviour (Verbruggen & Logan, 2008). The stop-signal paradigm (Figure 1A) is currently one of the most popular tasks for examining response inhibition in the laboratory. The last decade has witnessed an exponential rise in stop-signal studies in various research domains (see Figure S1 in the Supplemental Material available online). The paradigm is popular because it allows researchers to estimate the covert latency of the stop process: the *stop-signal reaction time* (SSRT). SSRT has been used to explore the cognitive and neural mechanisms of response inhibition, the development and decline of inhibitory capacities across the life span, and correlations between individual differences in stopping and behaviours such as substance abuse, pathological gambling, risk-taking, and more generally, control of impulses and urges (Chambers, Garavan, & Bellgrove, 2009; Logan, 1994; Verbruggen & Logan, 2008). In the present study, we used simulations to test the reliability and accuracy of SSRT estimates. Previous simulations of Band, van der Molen, and Logan (2003) showed that commonly used SSRT-estimation methods were not influenced much by variability in go reaction time (RT) or SSRT, or by dependency between the go and stop processes. However, we will show that estimates are strongly biased by positive skew and by gradual slowing of reaction times. Because skew and slowing are important characteristics of RT distributions in most stop-signal experiments (see below), our simulations suggest that some of the previously reported differences in stopping may be spurious.

SSRT is estimated based on the assumptions of the independent race model (Logan, 1994; Logan & Cowan, 1984; Verbruggen & Logan, 2009a): performance in the stop task can be modelled as a race between a go process, which is triggered by the presentation of the go stimulus, and a stop process, which is triggered by the presentation of a stop signal (Figure 1B). The stop signal occurs after a variable delay (stop-signal delay; SSD). If the go process finishes before the stop process (i.e. when $RT < SSRT + SSD$; Figure 1B) then response inhibition is unsuccessful and a response is executed; if the stop process finishes before the go process (i.e. when $RT > SSRT + SSD$) then the response is correctly withheld. The race model provides two common methods for estimating SSRT: the *integration* method and the *mean* method (Logan &

Cowan, 1984). In the *integration method*, the point at which the stop process finishes is estimated by integrating the RT distribution and finding the point at which the integral equals the probability of responding [p(respond|signal)] for a specific delay (Figure 1B). SSRT is then calculated by simply subtracting SSD from the finishing time. In the *mean method*, the mean of the inhibition function (a plot of the probability of responding given a stop signal against SSD; see Logan & Cowan, 1984; Verbruggen & Logan, 2009a) is subtracted from the mean of the RT distribution.

In recent years, the majority of stop-signal studies have used a dynamic tracking procedure to determine an SSD at which subjects inhibit 50% of the time. At the beginning of the experiment, SSD is set to a specific value (e.g. 250 ms) and is then constantly adjusted after stop signal trials, depending on the outcome of the race: when inhibition is successful, SSD increases (e.g. by 50 ms); when inhibition is unsuccessful, SSD decreases (e.g., by 50 ms). This one-up/one-down tracking procedure typically results in a p(respond|signal) ≈ .50, which means that the race between the stop process and the go process is tied. Then SSRT is usually estimated with the mean method or the integration method (see Figure S1 in Supplemental Material)[i]. The mean method uses the mean of the inhibition function (see above), which corresponds to the average SSD obtained with the tracking procedure when p(respond|signal) = .50. In other words, the mean method assumes that the mean RT equals SSRT + mean SSD, so SSRT can be estimated easily by subtracting mean SSD from mean RT (e.g. Logan & Cowan, 1984; Logan, Schachar, & Tannock, 1997). The integration method assumes that the finishing time of the stop process corresponds to the nth RT, with n = the number of RTs in the RT distribution multiplied by the overall p(respond|signal) (Logan, 1981); SSRT can then be estimated by subtracting mean SSD from the nth RT (e.g. Ridderinkhof, Band, & Logan, 1999; Verbruggen, Liefooghe, & Vandierendonck, 2004).

---

[i] Some studies estimated SSRT by subtracting mean SSD from the median RT. This 'median' method is a variant of the integration method that assumes that p(respond|signal) is always exactly .50. This is rarely the case. Thus, the median method is most of the time a less accurate version of the integration method; therefore, it is not further considered in the main analyses. We report the results for the median method in Table S8 in the Supplemental Material. The table shows that when 'subjects' did not slow and p(respond|signal) was close to .50, the results were very similar to the results for the integration method (as expected). However, when slowing was implemented, SSRT was overestimated [because p(respond|signal) was often lower than .50; see Tables S3–S5]. The overestimation bias was less pronounced than the bias observed for the mean method because the median is less influenced by the tail of the distribution than the mean.

Simulations and reliability[ii] tests suggest that when the tracking procedure is used, the mean and integration estimates are both reliable (Band et al., 2003; Congdon et al., 2012; Logan et al., 1997; Williams, Ponesse, Schachar, Logan, & Tannock, 1999). However, a recent empirical study reported numerical differences between the two (Boehler et al., 2012). We propose that such discrepancies are due to two factors, namely (1) skewness of the RT distribution and (2) the degree of proactive response slowing in anticipation of stop signals. Indeed, the simulations of Band et al. (2003), and the comparison of mean and integration method by Boehler et al. (2012), suggest that skew and slowing might have an effect on estimations. However, these factors not been systematically explored in the simulations or reliability tests so far. There are often large individual or group differences in the shape of the RT distribution and the degree of response slowing, so it is important to know to what extent these differences influence SSRT estimates.

In our first set of simulations, we examined the effect of positively skewed RT distributions on SSRT estimates. It is well known that the mean is strongly influenced by extreme scores in the tails of the distribution; the median is less affected by the tails. In the stop-signal task, the median corresponds to the nth RT when p(respond|signal) is exactly .50. Because RT distributions are usually positively skewed (Ratcliff, 1993), the right tail of the distribution might explain discrepancies between the 'mean' and 'integration' estimates. As shown in Figure 1C, the mean method would overestimate the finishing time of the stop process (and therefore, the SSRT) when the RT distribution is skewed, whereas the integration method might provide a more accurate estimate. We tested this in the first set of simulations.

In a second and third set of simulations, we explored the effect of response slowing on SSRT estimates. Recent studies have shown that subjects slow responses either *proactively* when they expect that stop signals might occur, or *reactively* when they fail to inhibit their responses (e.g. Aron, 2011; Bissett & Logan, 2011; Leotti & Wager, 2010; Verbruggen & Logan, 2009b; Verbruggen, Logan, Liefooghe, & Vandierendonck, 2008; Zandbelt, Bloemendaal, Neggers, Kahn, & Vink, 2012). Indeed, subjects sometimes slow their RT distributions over the course of the experiment to try to beat the tracking algorithm (see e.g. Leotti & Wager, 2010 for some extreme

[ii] Reliability tests typically used the split-half method: The data set is split in two and SSRTs for both subsets are compared.

examples). These shifts in the RT distribution could result in overestimates of SSRT in the mean method because slowing would primarily influence the right tail of the distribution and underestimates in SSRT in the integration method because the tracking is a step behind when subjects continuously slow down. We tested the effect of slowing in the second and third set simulations.

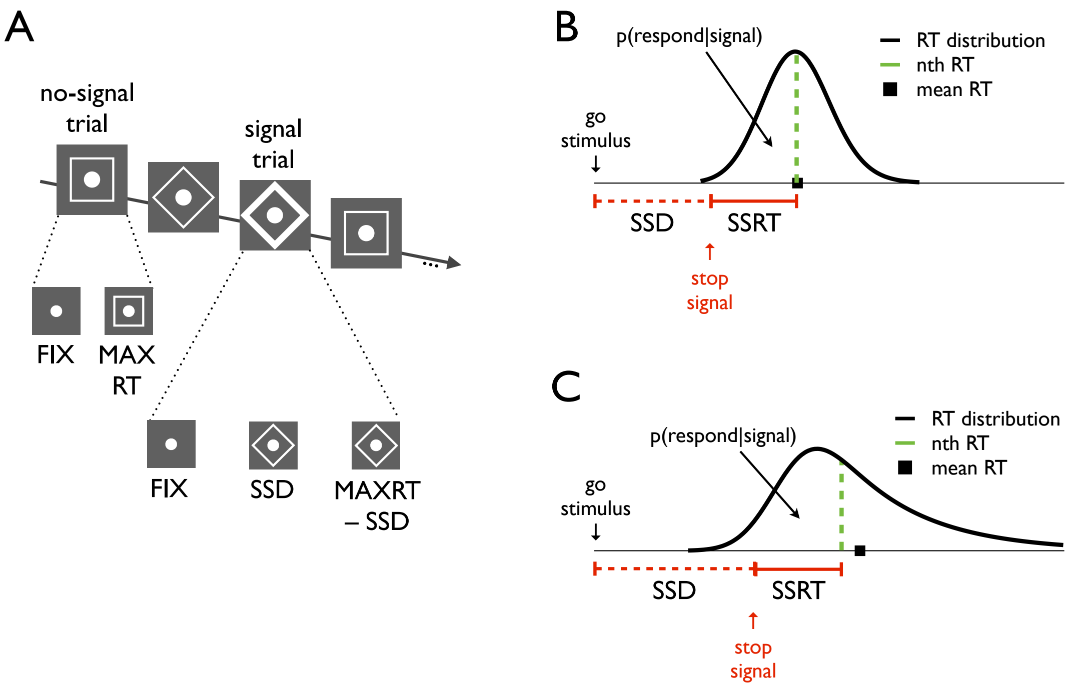


Figure 1. (A) In the stop-signal task, participants perform a go task (e.g. respond to the shape of a stimulus). On a minority of the trials, the go stimulus is followed by a 'stop signal' (e.g. the outline of the shape turning bold) after a variable delay (stop-signal delay; SSD); this stop signal instructs the subject to withhold the planned response. FIX = presentation duration of the fixation sign; MAXRT = response deadline. (B) Graphic representation of the assumptions of the independent horse-race model of Logan and Cowan (1984), indicating how the probability of responding [p(respond|signal)] depends on the distribution of go reaction times, stop-signal delay (SSD) and stop-signal reaction time (SSRT). In this example, p(respond|signal) = .50. The dotted line corresponds to the nth percentile, with n = p(respond|signal); the square shows the mean of the RT distribution. (C) This panel demonstrates that when the distribution is skewed to the right, there is a substantial difference between the mean and the nth RT; this may influence the SSRT estimations (see main text).

Methods

Race model simulations

Performance in the stop-signal task was simulated based on assumptions of the independent race model (Logan & Cowan, 1984): on stop-signal trials, a response was deemed to be withheld (*signal-inhibit*) when the RT was larger than SSRT + SSD; a response was deemed to be erroneously executed (*signal-respond*) when RT was smaller than SSRT + SSD.

All simulations were done using R (http://www.r-project.org). RTs were sampled from an ex-Gaussian distribution, using the rexGaus function (http://gamlss.org). The ex-Gaussian distribution is often used by psychologists to describe RT data (Ratcliff & Murdock, 1976); it has a positively skewed unimodal shape and results from a convolution of a normal (Gaussian) distribution and an exponential distribution. It is characterised by three parameters: mu (mean of the Gaussian component), sigma (SD of Gaussian component), and tau (both the mean and SD of the exponential component) (Figure S2 in the Supplemental Material shows how changes in these three parameters influence the distribution). Sigma approximately represents the rise in the left tail and tau the fall in the right tail of the ex-Gaussian distribution, whose mean = mu + tau and variance = $tau^2 + sigma^2$ (Ratcliff, 1979). Band et al. (2003) also used an ex-gaussian distribution to model reaction times in their simulations.

In the first set of simulations, sigma for the go task (RT-sigma) was 50, 100, or 150; and tau for the go task (RT-tau) was 50, 150, 250 (see e.g. Schmiedek, Oberauer, Wilhelm, Süss, & Wittmann, 2007, for a series of choice-reaction time tasks with tau's in this range). Empirically, sigma is usually not more than one fourth of tau (Ratcliff, 1993); however, we included a wider range of sigma because variability is often increased in clinical populations (e.g.Klein, Wendling, Huettner, Ruder, & Peper, 2006; Leth-Steensen, King Elbaz, & Douglas, 2000). For each combination of RT-sigma and RT-tau, we simulated the data of 100 'subjects'. Mu was different for each subject [mu(subject)]; it was sampled from a normal distribution with mean = 400 (i.e. the population mean) and SD = 25, with the restriction that it was larger than 300.

SSRTs were also sampled from an ex-Gaussian distribution. For all subjects, both SSRT-sigma and SSRT-tau were 10. Mu(subject) was derived from a normal distribution with mean = 200 (population mean), and SD = 10, with the restriction that mu(subject) was larger than 150. Note that we also ran simulations in which SSRT-sigma and SSRT-tau were varied; the results are reported in the Supplemental Material (Table S7). SSRT-sigma and SSRT-tau did not influence the estimates much, and did not interact with the effects of RT-tau and response slowing. Therefore, we used only one value for SSRT-sigma and SSRT-tau in the main simulations reported below.

For each simulated subject, there were 4 blocks of 60 trials; signals randomly 'occurred' on 25% of the trials, resulting in 15 stop-signal trials per block. The 'delay' between the start of the go process and the start of the stop process (SSD) was initially set at 150 + RT-tau (e.g. when RT-tau was 250, the initial SSD was 400), and subsequently adjusted: after a signal-inhibit trial, SSD increased by 50; after a signal-respond trial, SSD decreased by 50. The start value was chosen in such a way that the race between go and stop would be close, but with a small initial 'head start' for the stop process (the finishing time of the go process ≈ mean RT = 400 + RT-tau; the finishing time of the stop process ≈ SSD + mean SSRT = 150 + RT-tau + 200 + SSRT-tau). Because mu was not manipulated across conditions, we only used tau to determine start SSD.

In the second set of simulations, we examined the effect of gradual slowing of reaction times. RTs were again derived from an ex-Gaussian distribution but RT-mu increased linearly over trials. The start value of RT-mu was again derived from a normal distribution with mu = 400 and SD = 25. The slope of the increase depended on a slowing factor, which could be 1, 1.5, or 2.5; these values were roughly based on the degree of slowing for individual subjects in one of our previous studies (Verbruggen & Logan, 2009b). The slope of the increase was (y2 – y1)/(x2 – x1), with y2 = RT-mu(start) * slowing factor, and y1 = RT-mu(start), x2 = 240 (the trial number of the last trial),and x1 = 1 (trial number of the first trial). When the slowing factor was 1, the slope was 0 (i.e. y2 = y1, so no slowing). When the slowing factor was 1.5 or 2.5, the slope was positive and RT-mu increased. For example, with only six trials and slowing factor = 1.5, RT-mu's would be: T1 = mu(start), T2 = mu(start) * 1.1, T3 = mu(start) * 1.2, T4 = mu(start) * 1.3, T5 = mu(start) * 1.4, T6 = mu(start) * 1.5. In this second set of simulations, RT-sigma was 50 or 150, and RT-tau was 50 or 250.

Finally, in the third set of simulations, the slowing factor was different for each subject to allow for individual differences in slowing. For each simulated subject, slowing factor was derived from a uniform distribution with min = 1, and max = 3.

Estimation and analyses

For the first set of simulations, we estimated SSRT over all blocks using the mean (SSRT = mean RT - mean SSD) and integration (SSRT = nth RT - mean SSD) methods. For the second and third set, we also estimated SSRT for each block separately using the integration method and then took the average of these four block estimates[iii]. Trials with a RT > 2000 were considered to be missed responses (in real experiments, there is always a response deadline around this value). These 'missed' trials were excluded when we estimated SSRT using the 'mean' method; for the integration methods, RT for missed responses was set to 2000[iv].

For each estimation method, we calculated the difference between the estimated SSRT and actual SSRT; positive values indicate that SSRT is overestimated, whereas negative values indicate that SSRT is underestimated. Table 1 reports the mean difference scores, confidence intervals, and t-tests that explored whether the SSRT difference was reliably different from zero. Using mixed ANOVAs (see Tables S2, S4, and S6 in the Supplemental Material for overviews), we then tested whether the difference scores were influenced by estimation method, RT-sigma, RT-tau, and slowing (second set of simulations).

Results and discussion

In the first set of simulations, the tracking procedure worked well and p(respond|signal) was close to .50 for all RT-sigma/RT-tau combinations (see Table S1 in Supplemental Material). When we collapsed across all values of RT-sigma and RT-tau, we found that the mean method overestimated SSRT; by contrast, the integration method tended to slightly underestimate SSRT (Table 1).

[iii] Because there was an equal number of signal trials in each block, a block-based variant of the mean method results in the exact same estimate as the estimate obtained using experiment-based mean method.

[iv] As discussed below, the mean method tended to overestimate SSRT. This overestimation would be even more pronounced if 'missed' responses were not excluded or when 'missed' RT was set to 2000.

	Mean difference	95% CI		one-sample t	df	p
		lower	upper			
Experiment 1						
Mean method	23.46	21.15	25.78	19.92	899	< .001
Integration method	-6.16	-8.35	-3.97	5.53	899	< .001
Experiment 2: slow factor = 1						
Mean method	25.89	22.09	29.69	13.39	399	< .001
Integration method	-4.73	-8.19	-1.28	2.70	399	0.01
Integration (blocked)	-3.48	-6.92	-0.04	1.99	399	0.05
Experiment 2: slow factor = 1.5						
Mean method	39.67	35.47	43.88	18.55	399	< .001
Integration method	-1.13	-4.62	2.36	0.64	399	0.52
Integration (blocked)	-0.62	-4.15	2.91	0.35	399	0.73
Experiment 2: slow factor = 2.5						
Mean method	62.67	58.27	67.07	27.99	399	< .001
Integration method	-14.10	-18.38	-9.82	6.48	399	< .001
Integration (blocked)	-1.78	-5.38	1.81	0.97	399	0.33
Experiment 3						
Mean method	51.07	46.81	55.33	23.55	399	< .001
Integration method	-6.59	-10.37	-2.80	3.42	399	< .001
Integration (blocked)	0.84	-2.52	4.19	0.49	399	0.62

Table 1: Overview of analyses of the difference scores (i.e. difference between the estimated SSRT and actual SSRT; positive values indicate that SSRT is overestimated, whereas negative values indicate that SSRT is underestimated). One-sample t tests were performed to examine whether the scores were significantly different from zero.

We used box-plots of differences scores to examine the accuracy of SSRT estimates and to explore the estimation bias: a leftward shift of a box indicates underestimation; rightward shift indicates overestimation. The plots (Figure 2) demonstrate that when RT-sigma and RT-tau were small, the difference between the estimated and actual SSRTs was small for most of subjects. An increase in RT-sigma led to more noisy estimates, but did not induce a systematic bias (i.e. the box widened but was still centred around zero). Changes in RT-tau, which influenced the right tail (positive skew) of the RT-distribution, had a more pronounced effect on SSRT estimations. A comparison of the bottom- and top-row box-plots shows that when RT-tau increased, estimates became noisier and, more importantly, became biased. For the mean method, the rightward shift of the top-row boxes shows that SSRT was overestimated when RT-tau increased. The integration method had a small tendency to underestimate SSRT when RT-tau increased, but this effect was less pronounced. Thus, the integration method seemed more robust and less biased than the mean method. These conclusions are supported by significant main effects of estimation method, RT-tau, and an interaction between method and RT-tau (see Table S2 in the Supplemental Material).

The overestimation bias for large RT-tau's is problematic when SSRTs of different groups or conditions are compared. Often, RT distributions differ between groups or conditions. For example, a recent study showed that RT-tau was approximately 251 ms for children with ADHD and 162 ms for control children (Epstein et al., 2011). Such RT-tau group differences could influence the SSRT estimates. We further tested this by randomly selecting 20 subjects in the RT-sigma=100/RT-tau=150 condition and 20 subjects in the RT-sigma=100/RT-tau=250 condition. As expected, there was no difference between the true stop latencies in both conditions [208 vs. 206; $F(1,38) = 0.11$, $p = 0.750$]. However, there was a significant 31 ms difference between the estimated SSRTs [RT-tau=150: 229 ms, RT-tau=250: 260 ms; $F(1,38) = 6.60$, $p = .014$]. Thus, when there are differences in RT-tau, the mean method may lead to incorrect conclusions about group differences in SSRTs. Note that there was no difference between the SSRTs estimated using the integration method [200 vs. 204; $F(1,38) = 0.07$, $p = 0.798$].

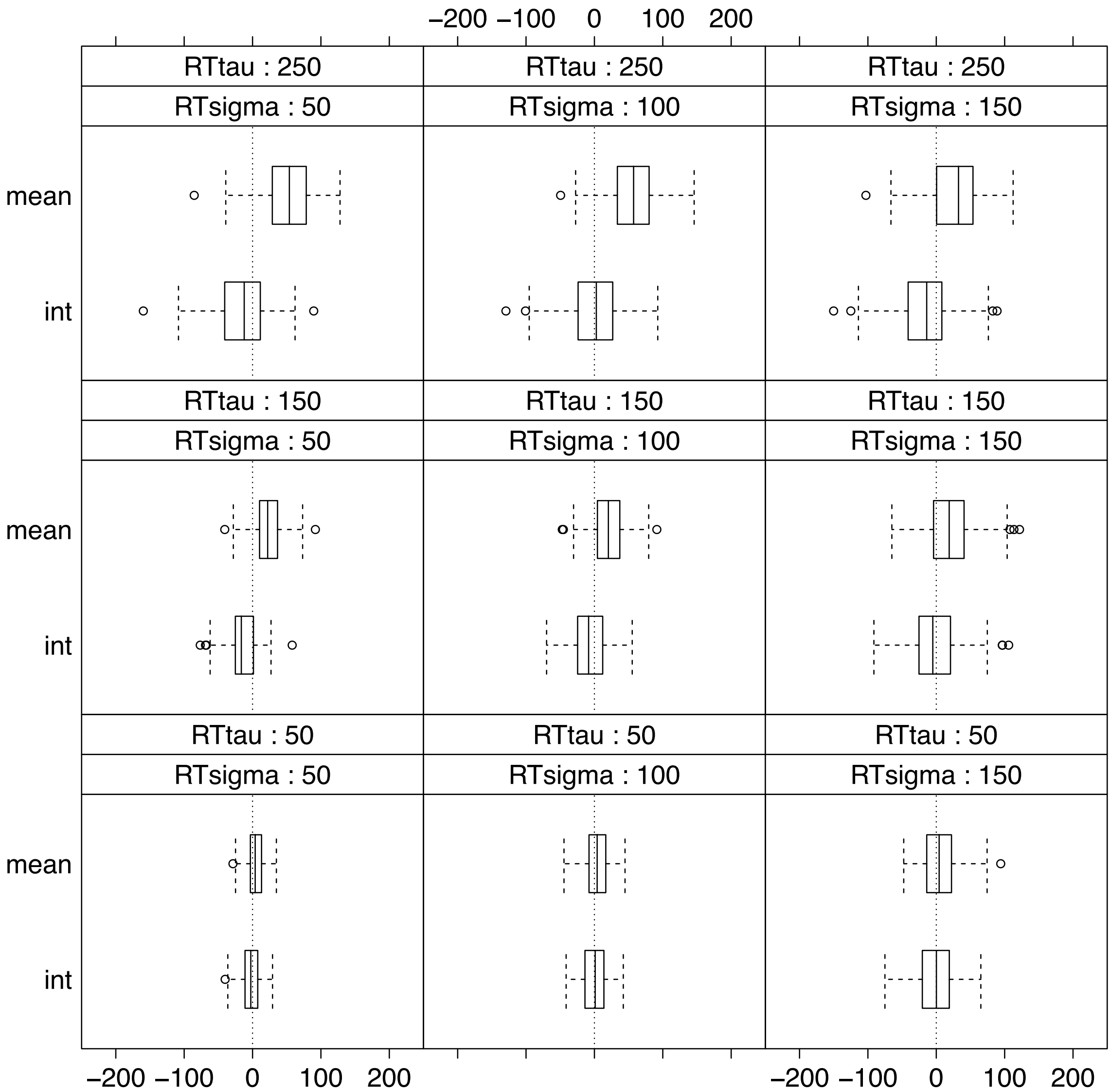


Figure 2. This series of box-plots shows the difference between the estimated stop latency (using the mean and integration methods) and the true stop latency for each combination of RT-sigma and RT-tau. Negative values indicate the estimated value is an underestimation of the true SSRT; positive values indicate that the estimated SSRT is longer than the actual stop latency. The plots show that estimates become noisier when RT-sigma and RT-tau increase. Importantly, the mean method substantially *overestimates* the stop latency when RT-tau increases.

In the second set of simulations, we tested how gradual slowing of RTs over trials influences the SSRT estimates. Here, we used two variants of the integration method: (1) the variant that we used in the first set of simulations and that uses all trials to obtain a single SSRT estimate (henceforth, the experiment-wide integration method), and (2) a block-based integration method that estimated SSRT for each block separately (there were 60 trials per block, 15 of which were signal trials) and then took the average of these four estimates.

The box-plots in Figure 3 show that the mean method overestimates SSRT when RT-tau increases or when mean RT gradually increases over trials (see also Table 1). By contrast, the experiment-wide integration method tended to underestimate SSRT, especially when the slowing factor increased (see Figure 3 and Table 1). The block-based integration method did not show such a consistent bias. These conclusions were supported by the ANOVAs reported in Table S4 of the Supplemental Material.

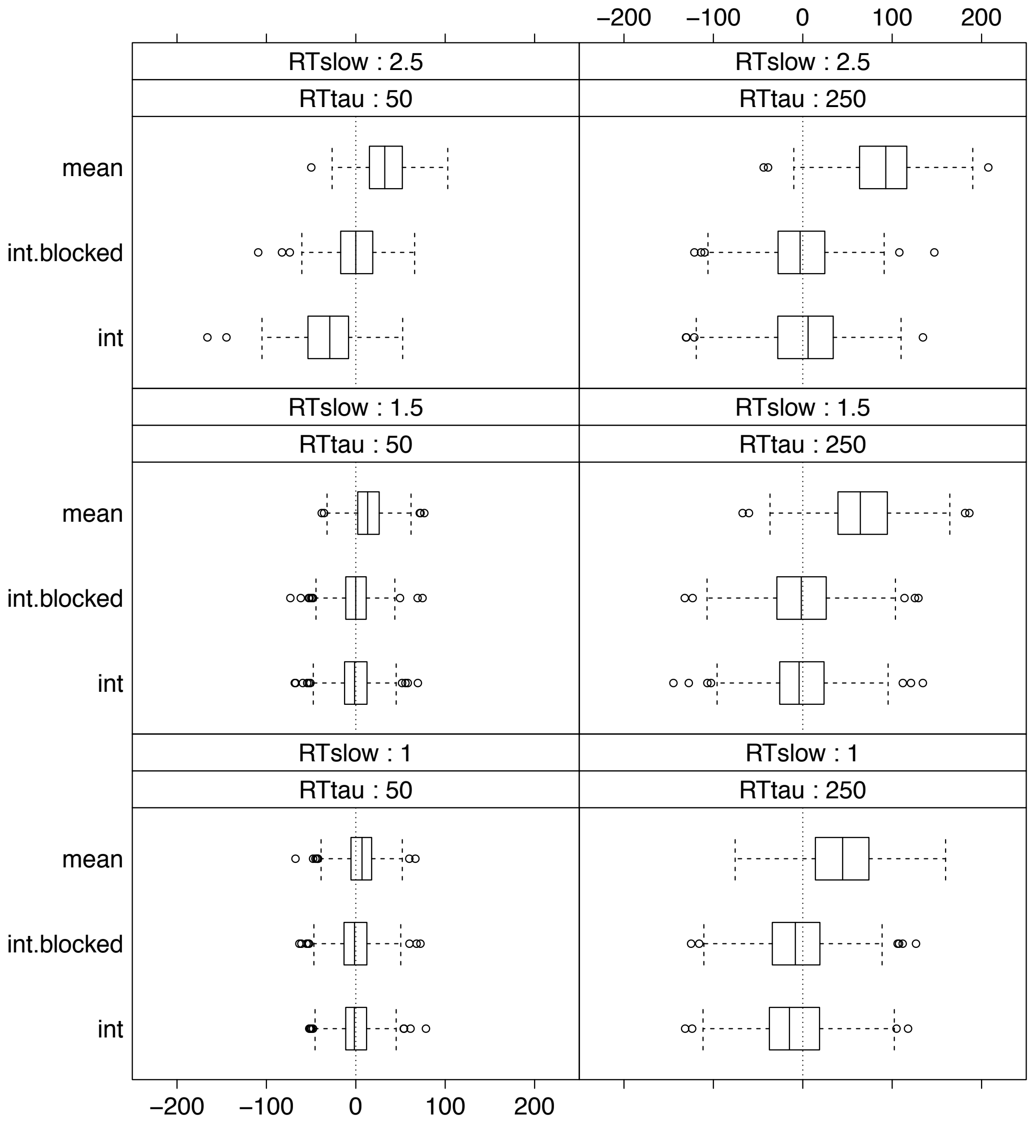


Figure 3. This series of box-plots shows how the SSRT estimates are influenced by RT-tau and response slowing. The mean method is influenced by both, which leads to *over*estimations for most RT-tau and slowing-factor combinations. By contrast, the integration method, which estimates SSRT based on all trials, tends to *under*estimate SSRT. Finally, the block-based integration method, which estimates SSRT for each block separately first, appears relatively immune to changes in RT-tau and slowing.

We found that the mean method was strongly influenced by response slowing. One possible explanation for this finding is that the mean method assumes that the probability of responding approximates .50. However, Table S3 (Supplemental Material) shows that when the slowing factor increased, p(respond|signal) tended to decrease: when the slowing factor is large, the tracking procedure cannot keep up with the changes in RT, so p(respond|signal) will be lower than .50. Therefore, we re-estimated SSRT using only those simulated subjects for which .40 < p(respond| signal) < .60; these values are based on the criterion discussed in Verbruggen, Logan, & Stevens (2008). The results are shown in Figure S3 (Supplemental Material); as can be seen, the RT-tau and slowing biases were still present, even when only the central estimates were included.

The second set of simulations demonstrated that the mean method and experiment-wide integration method were influenced by response slowing. In a third set of simulations, we used a random slowing factor for each simulated subject to explore the correlation between slowing and the degree of over/underestimation. Figure 4 shows that when RT-tau was low and the experiment-wide integration method was used, the estimated SSRT correlated negatively with the degree of slowing[v,vi]. Researchers have previously argued that such negative correlations could be due to proactive suppression of motor output or changes in task priorities (e.g. Jahfari, Stinear, Claffey, Verbruggen, & Aron, 2010; Leotti & Wager, 2010). Our simulations suggest that this negative correlation could be due to a bias in SSRT estimation. Importantly, this bias was not observed when SSRT was estimated for each block separately (middle row Figure 4). As expected based on the previous sets of simulations, we found a positive correlation between response slowing and degree of overestimation for the mean method.

---

[v] We obtained very similar results when we implemented slowing differently: On each trial, we obtained an RT from a single ex-gaussian distribution with constant parameters [mu(subject), sigma, tau]. This RT was then multiplied by the slowing factor. For example, if the slowing factor would have been 1.5 and the number of trials = 3, the RTs would be: T1 = RT(sampled), T2 = RT(sampled) * 1.25, T3 = RT(sampled) * 1.5. The negative correlations and underestimation bias were still present for the standard integration method (if anything, the effects were more pronounced), but not for the block-based method.

[vi] For the mean method, overall correlation was 0.26 [t(398) = 5.48, p < .001], which suggests that the mean method will overestimate SSRT when subjects slow down. For the experiment-wide integration method, the overall correlation was -0.19 [t(398) = -3.83, p < .001]; this suggests that the experiment-based integration method will underestimate SSRT when subjects slow down. Finally, for the block-based integration, overall correlation was 0.02 [t(398) = 0.34, p = 0.74], which suggests that the estimates are not influenced by slowing. We obtained very similar correlations between the degree of slowing and the difference between SSRT(estimate) and SSRT(true): for the mean method, the overall correlation was 0.29 [t(398) = 5.98, p < .001]; for the integration method, the overall correlation was -0.21 [t(398) = -4.36, p < .001]; finally, for the block-based integration, overall correlation was 0.02 [t(398) = 0.45, p = 0.66]. This confirms that the observed correlations between slowing and SSRT were due to an estimation bias.

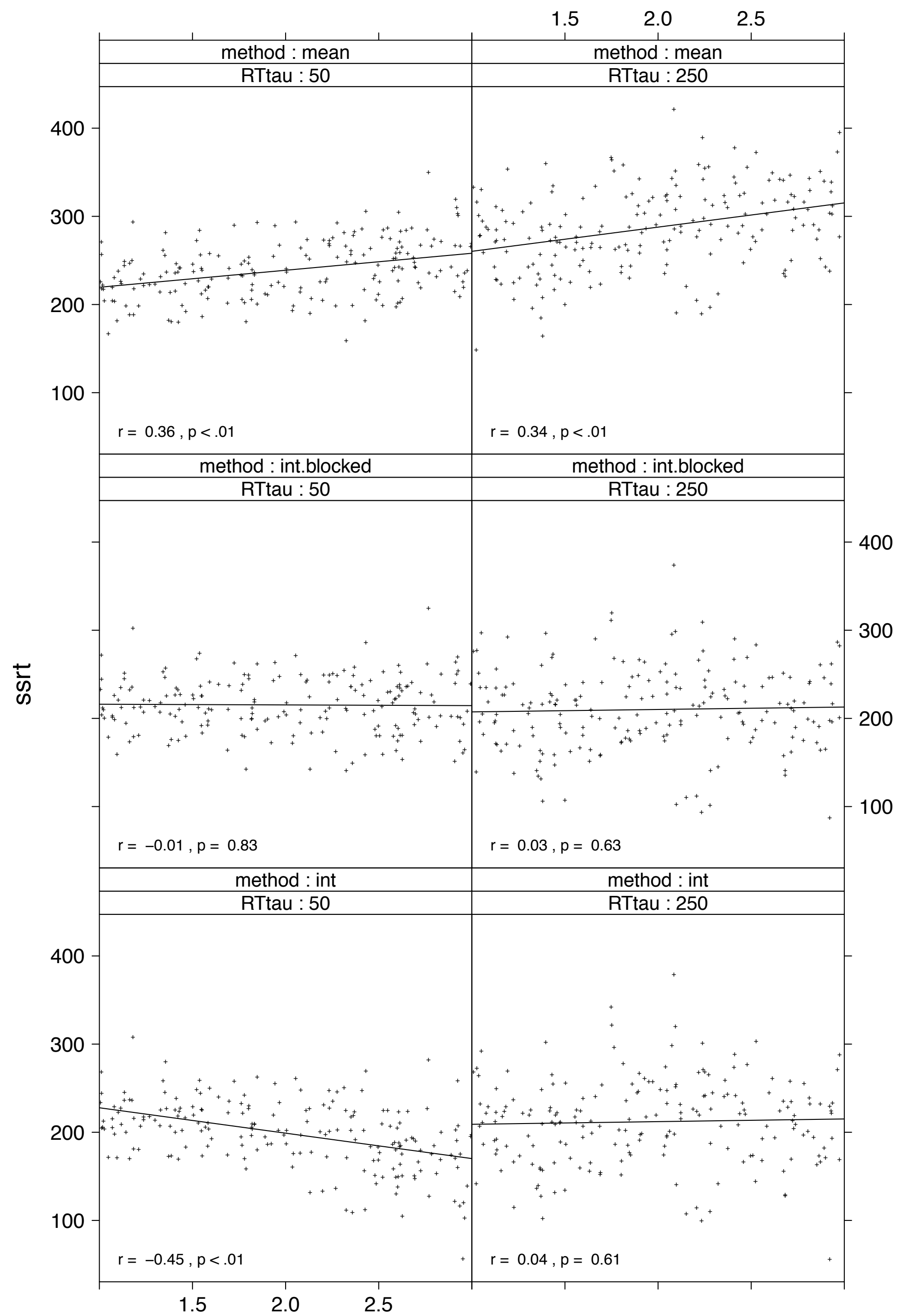


Figure 4: Correlation between the estimated SSRT and the slowing factor.

Conclusions and practical guidelines

In the present study, we explored to what extent the skew of the RT distribution and gradual slowing of response latencies influences the 'mean' and 'integration' SSRT estimates. The mean method is often used (see Figure S1 in the Supplemental Material) because it is very easy: SSRT can be estimated simply by subtracting the mean SSD from the mean RT. However our simulations show that this approach overestimates SSRT when the RT distribution is skewed to the right (i.e. when RT-tau is large) or when RTs increase gradually over the course of the experiment. We demonstrated that individual or group differences in RT skew or response slowing could result in spurious inhibitory differences. Unfortunately, such RT differences may occur frequently. For example, studies showed that SSRT is longer for children with ADHD than for control children (Lijffijt, Kenemans, Verbaten, & van Engeland, 2005; Oosterlaan & Sergeant, 1998; Schachar & Logan, 1990). However, a recent study estimated that tau was much higher in children with ADHD than in control children (Epstein et al., 2011). Thus, the mean method will overestimate SSRT differences between ADHD children and control children, and possibly produce spurious differences. Thus, **our first take-home message is that the mean method should be abandoned** because it is overly susceptible to the shape of the RT distribution.

The integration method fared better in the first set simulations: there was a trend to underestimate SSRT slightly (approximately 4 ms), but there were no obvious group differences caused by changes in the shape of the RT distribution. This is consistent with a recent reliability analysis that used split-half reliability measures (Congdon et al., 2012). However, the second and third set of simulations showed that the small underestimation bias for the integration method became more pronounced when there is gradual slowing of RTs across blocks. This underestimation bias may explain the previously observed negative correlations between SSRT and response slowing(e.g. Jahfari et al., 2010; Leotti & Wager, 2010). Thus, **our second take-home message is that the experiment-wide integration method results in reliable and unbiased estimates unless subjects slow their RT gradually.**

The gradual slowing of reaction times may be reduced by clear advance instructions (for example, by stressing speed in the go task and explaining the staircase tracking procedure) and by providing feedback after every trial (e.g. Ridderinkhof et al., 1999; Verbruggen et al., 2004) or after every block (e.g. Verbruggen et al., 2008). Thus, **our third take-home message is that in standard stop tasks researchers should provide clear instructions and implement feedback procedures to discourage excessive strategic slowing.**

Even when feedback is provided, slowing may still be observed in certain subjects (e.g. Verbruggen et al., 2004; Verbruggen et al., 2008). Researchers can exclude those subjects who slow their responses substantially; our simulations suggest that the underestimation bias appeared when the mean of the normal part of the distribution doubled[vii]. However, this may result in the exclusion of a large number of subjects in some experiments, which could induce an exclusion bias. Also, researchers may be specifically interested in the correlation between slowing and SSRT. Recently, several authors have argued that strategy adjustments may be an important aspect of successful stop performance and more generally, impulse control in everyday life (e.g. Aron, 2011; Bissett & Logan, 2011; Leotti & Wager, 2010; Verbruggen & Logan, 2009b). Feedback about slowing may not be provided when such strategic adjustments are examined. Furthermore, excluding subjects who slow substantially is not be appropriate in such studies. The second and third set of simulations show that a block-based version of the integration method is less susceptible to bias from response slowing. When SSRT was estimated for each block separately (number of no-signal trials per block = 45; number of signal trials per block = 15) and then averaged, we obtained a reliable and unbiased SSRT even when there was substantial response slowing. Additional analyses (Supplementary Material; Figures S5–S6) suggest that approximately 40-80 trials are required per block (25% of which are signal trials). If there are fewer trials, the estimates become too noisy; if there are more trials, the underestimation bias starts to emerge. We recommend that there are at least 50 signals in total. Thus, **our fourth take-home message is that researchers should estimate SSRT for each block separately when strategic slowing is observed and subjects cannot be excluded**.

It should be noted that slowing could be interpreted as a violation of the *context independence* and the *stochastic independence* assumptions of the race model (Logan & Cowan,

[vii] When RT-tau is small, the mean of the normal part of the distribution will not differ much from the global average.

1984). Context independence (also referred to as signal independence) refers to the assumption that the RT distribution is the same for no-signal trials and stop-signal trials. Stochastic independence refers to the assumption that trial-by-trial variability in RT is unrelated to trial-by-trial variability in SSRT. Gradual slowing of RT does not necessarily violate these assumptions: Because subjects cannot predict whether a stop signal will occur in the standard version of a stop task, they are expected to slow down on all trials (including no-signal trials). In other words, the assumptions of the race model hold as long as slowing occurs to a similar degree on both signal and no-signal trials. Note also that the race model does not make assumptions about the shape of the finishing-time distributions. Thus, skew should not influence the SSRT estimations. The results of the first set of simulations demonstrated that this was the case for the integration method.

To conclude, our results demonstrate that the central SSRT estimates, which were previously thought to be most reliable, are strongly influenced by the right tail of the RT distribution and gradual slowing of RTs. Therefore, we recommend that researchers abandon the mean method to estimate SSRT and instead use the experiment-wide or block-based integration method to reliably estimate the latency of response inhibition.

References

Aron, A. R. (2011). From reactive to proactive and selective control: Developing a richer model for stopping inappropriate responses. *Biological Psychiatry*, *69*(12), e55-68. doi:10.1016/j.biopsych.2010.07.024

Band, G. P., van der Molen, M. W., & Logan, G. D. (2003). Horse-race model simulations of the stop-signal procedure. *Acta Psychologica*, *112*(2), 105-42.

Bissett, P. G., & Logan, G. D. (2011). Balancing cognitive demands: Control adjustments in the stop-signal paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 392-404. doi:10.1037/a0021800

Boehler, C. N., Appelbaum, L. G., Krebs, R. M., Hopf, J. M., & Woldorff, M. G. (2012). The influence of different stop-signal response time estimation procedures on behavior-behavior and brain-behavior correlations. *Behavioural Brain Research.* doi:10.1016/j.bbr.2012.01.003

Chambers, C. D., Garavan, H., & Bellgrove, M. A. (2009). Insights into the neural basis of response inhibition from cognitive and clinical neuroscience. *Neuroscience and Biobehavioral Reviews*, *33*(5), 631-646. doi:10.1016/j.neubiorev.2008.08.016

Congdon, E., Mumford, J. A., Cohen, J. R., Galvan, A., Canli, T., & Poldrack, R. A. (2012). Measurement and reliability of response inhibition. *Frontiers in Psychology*, *3*, 37. doi:10.3389/fpsyg.2012.00037

Epstein, J. N., Langberg, J. M., Rosen, P. J., Graham, A., Narad, M. E., Antonini, T. N., . . . Altaye, M. (2011). Evidence for higher reaction time variability for children with adhd on a range of cognitive tasks including reward and event rate manipulations. *Neuropsychology.* doi:10.1037/a0022155

Jahfari, S., Stinear, C. M., Claffey, M., Verbruggen, F., & Aron, A. R. (2010). Responding with restraint: What are the neurocognitive mechanisms? *Journal of Cognitive Neuroscience*, *22*(7), 1479-92. doi:10.1162/jocn.2009.21307

Klein, C., Wendling, K., Huettner, P., Ruder, H., & Peper, M. (2006). Intra-subject variability in attention-deficit hyperactivity disorder. *Biological Psychiatry*, *60*(10), 1088-97. doi:10.1016/j.biopsych.2006.04.003

Leotti, L. A., & Wager, T. D. (2010). Motivational influences on response inhibition measures. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 430-47. doi:10.1037/a0016802

Leth-Steensen, C., King Elbaz, Z., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica*, *104*(2), 167-190.

Lijffijt, M., Kenemans, J. L., Verbaten, M. N., & van Engeland, H. (2005). A meta-analytic review of stopping performance in attention-deficit/hyperactivity disorder: Deficient inhibitory motor control? *Journal of Abnormal Psychology*, *114*(2), 216-22. doi:10.1037/0021-843X.114.2.216

Logan, G. D. (1981). Attention, automaticity, and the ability to stop a speeded choice response. *Attention and Performance IX*, 205-222.

Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189-239). Academic Press.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*(3), 295-327.

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and inhibitory control. *Psychological Science*, *8*(1), 60-64.

Oosterlaan, J., & Sergeant, J. A. (1998). Response inhibition and response re-engagement in attention-deficit/hyperactivity disorder, disruptive, anxious and normal children. *Behavioural Brain Research*, *94*(1), 33-43.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*(3), 446-61.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510-32.

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review; Psychological Review*, *83*(3), 190-214.

Ridderinkhof, K. R., Band, G. P. H., & Logan, G. D. (1999). A study of adaptive behavior: Effects of age and irrelevant information on the ability to inhibit one's actions. *Acta Psychologica*, *101*(2-3), 315-337.

Schachar, R., & Logan, G. D. (1990). Impulsivity and inhibitory control in normal development and childhood psychopathology. *Developmental Psychology*, *26*(5), 710-720.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süss, H. M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*(3), 414-29. doi: 10.1037/0096-3445.136.3.414

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, *12*(11), 418-24. doi:10.1016/j.tics.2008.07.005

Verbruggen, F., & Logan, G. D. (2009a). Models of response inhibition in the stop-signal and stop-change paradigms. *Neuroscience and Biobehavioral Reviews*, *33*(5), 647-61. doi:10.1016/j.neubiorev.2008.08.014

Verbruggen, F., & Logan, G. D. (2009b). Proactive adjustments of response strategies in the stop-signal paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 835-54. doi:10.1037/a0012726

Verbruggen, F., Liefooghe, B., & Vandierendonck, A. (2004). The interaction between stop signal inhibition and distractor interference in the flanker and stroop task. *Acta Psychologica*, *116*(1), 21-37. doi:10.1016/j.actpsy.2003.12.011

Verbruggen, F., Logan, G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods*, *40*(2), 479-83.

Verbruggen, F., Logan, G. D., Liefooghe, B., & Vandierendonck, A. (2008). Short-term aftereffects of response inhibition: Repetition priming or between-trial control adjustments? *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 413-26. doi: 10.1037/0096-1523.34.2.413

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental Psychology*, *35*(1), 205-13.

Zandbelt, B. B., Bloemendaal, M., Neggers, S. F., Kahn, R. S., & Vink, M. (2012). Expectations and violations: Delineating the neural network of proactive inhibitory control. *Human Brain Mapping.* doi:10.1002/hbm.22047

Supplementary Material

Table S1: p(respond|signal) for first set of simulations

		M	SD	min	max
Sigma	Tau				
50	50	0.495	0.011	0.467	0.517
100	50	0.494	0.013	0.467	0.517
150	50	0.493	0.011	0.467	0.517
50	150	0.498	0.011	0.467	0.533
100	150	0.496	0.013	0.467	0.517
150	150	0.496	0.013	0.467	0.517
50	250	0.504	0.015	0.467	0.533
100	250	0.500	0.015	0.467	0.533
150	250	0.499	0.016	0.467	0.533

Table S2: Overview of analyses of difference scores for the first set of simulations. We conducted a 2 (RT-tau) x 2 (RT-sigma) x 2 (method) ANOVA to examine effects of sigma, and tau, and to compare methods; we followed this up with separate 2 (RT-tau) x 2 (RT-sigma) ANOVAs for each method.

	df	*MSE*	*F*	*p*
Combined				
Method (M)	1, 891	62725	5609.24	< .001
Sigma (S)	2, 891	1944	3.84	0.022
Tau (T)	2, 891	1944	16.71	< .001
M x S	2, 891	62725	112.061	< .001
M x T	2, 891	62725	1326.78	< .001
S x T	4, 891	1944	5.74	< .001
M x S x T	4, 891	62725	21.24	< .001
Mean only				
Sigma (S)	2, 891	939	9.66	< .001
Tau (T)	2, 891	939	125.23	< .001
S x T	4, 891	939	8.64	< .001
Integration only				
Sigma (S)	2, 891	1075	5.83	< .001
Tau (T)	2, 891	1075	7.68	< .001
S x T	4, 891	1075	4.22	< .001

Table S3 : p(respond|signal) for second set of simulations

Sigma	Tau	Slowing	M	SD	min	max
50	50	1	0.494	0.010	0.467	0.517
150	50	1	0.494	0.013	0.467	0.533
50	250	1	0.505	0.015	0.467	0.533
150	250	1	0.499	0.015	0.467	0.533
50	50	1.5	0.460	0.012	0.433	0.483
150	50	1.5	0.463	0.014	0.433	0.500
50	250	1.5	0.473	0.014	0.433	0.500
150	250	1.5	0.472	0.017	0.433	0.517
50	50	2.5	0.398	0.013	0.367	0.417
150	50	2.5	0.401	0.014	0.367	0.433
50	250	2.5	0.410	0.017	0.367	0.450
150	250	2.5	0.410	0.019	0.367	0.467

Table S4: Overview of analyses of difference scores for the second set of simulations. We conducted a 2 (RT-tau) x 2 (RT-sigma) x 3 (method) x 3 (slowing) ANOVA to examine effects of sigma, tau, slowing and estimation method; we followed this up with separate 2 (RT-tau) x 2 (RT-sigma) x 3 (slowing) ANOVAs for each method.

	df	*MSE*	*F*	*p*
Combined				
Method (M)	2,2376	104	8594.58	< .001
Sigma (S)	1,1188	3533	0.01	0.922
Tau (T)	1,1188	3533	77.54	< .001
Slowing (SL)	2,1188	3533	8.405	0.000
M x S	2,2376	104	16.54	< .001
M x T	2,2376	104	2275.63	< .001
M x SL	4,2376	104	612.43	< .001
S x T	1,1188	3533	7.026	0.008
S x SL	2,1188	3533	0.247	0.781
T x SL	2,1188	3533	8.163	< .001
M x S x T	2,2376	104	67.92	< .001
M x S x SL	4,2376	104	20.30	< .001
M x T x SL	4,2376	104	119.00	< .001
S x T x SL	2,1188	3533	8.338	< .001
M x S x T x SL	4,2376	104	15.64	< .001
Mean only				
Sigma (S)	1,1188	1122	1.70	0.192
Tau (T)	1,1188	1122	649.62	< .001
Slowing (SL)	2,1188	1122	123.12	< .001
S x T	1,1188	1122	28.11	< .001
S x SL	2,1188	1122	2.37	0.094
T x SL	2,1188	1122	5.61	0.004
S x T x SL	2,1188	1122	6.32	0.002
Integration only				
Sigma (S)	1,1188	1342	0.88	0.347
Tau (T)	1,1188	1342	10.17	0.001
Slowing (SL)	2,1188	1342	13.35	< .001

S x T	1,1188	1342	5.44	0.020
S x SL	2,1188	1342	0.84	0.434
T x SL	2,1188	1342	34.17	< .001
S x T x SL	2,1188	1342	14.97	< .001
Integration (blocked) only				
Sigma (S)	1,1188	1276	0.29	0.589
Tau (T)	1,1188	1276	3.20	0.074
Slowing (SL)	2,1188	1276	0.65	0.523
S x T	1,1188	1276	0.08	0.782
S x SL	2,1188	1276	1.02	0.360
T x SL	2,1188	1276	1.08	0.340
S x T x SL	2,1188	1276	4.33	0.013

Table S5 : p(respond|signal) for third set of simulations

		M	SD	min	max
Sigma	Tau				
50	50	0.432	0.037	0.367	0.500
150	50	0.431	0.042	0.333	0.517
50	250	0.446	0.039	0.367	0.517
150	250	0.442	0.039	0.350	0.517

Table S6: Overview of analyses of difference scores for the third set of simulations. We conducted a 2 (RT-tau) x 2 (RT-sigma) x 3 (method) ANOVA to examine effects of sigma, tau, and estimation method; we followed this up with separate 2 (RT-tau) x 2 (RT-sigma) ANOVAs for each method.

	df	*MSE*	*F*	*p*
Combined				
Method (M)	2,792	259	1520.26	< .001
Sigma (S)	1,396	3251	0.68	0.410
Tau (T)	1,396	3251	49.71	< .001
M x S	2,792	259	1.0223	0.360
M x T	2,792	259	274.94	< .001
S x T	1,396	3251	0.23	0.631
M x S x T	2,792	259	15.27	< .001
Mean only				
Sigma (S)	1,396	1194	0.76	0.385
Tau (T)	1,396	1194	226.91	< .001
S x T	1,396	1194	4.76	0.030
Integration only				
Sigma (S)	1,396	1410	0.07	0.798
Tau (T)	1,396	1410	23.34	< .001
S x T	1,396	1410	0.34	0.559
Integration (blocked) only				
Sigma (S)	1,396	1164	1.50	0.222
Tau (T)	1,396	1164	0.03	0.867
S x T	1,396	1164	2.14	0.144

Table S7: We reran the second set of simulations to test whether the effect of RT-tau and response slowing were influenced by SSRT-sigma and SSRT-tau (i.e. rise of the left tail and fall of the right tail of the SSRT distribution). This table shows the average (avg; in grey) and standard deviation (sd) of the difference scores for each combination of SSRT-sigma * SSRT-tau * estimation method * RT-tau * slowing factor. We replicated the effects of RT-tau and response slowing: the mean method overestimated SSRT when RT-tau and/or the slowing factor increased; the experiment-wide integration method [int(exp)] underestimated SSRT when RTs gradually increased; finally, the block-based integration method seemed relatively immune against effects of RT-tau and response slowing. Importantly, SSRT-sigma and SSRT-tau had very little effect on the differences scores and did not interact with the effects of RT-tau or slowing.

	slowing factor = 1				slowing factor = 1.5				slowing factor = 2.5			
	RTtau=50		RTtau=250		RTtau=50		RTtau=250		RTtau=50		RTtau=250	
	avg	sd	avg	sd	avg	sd	avg	sd	avg	sd	avg	sd
SSRT-sigma = 10 SSRT-tau = 10												
int(exp)	1.92	24.8	-5.6	42.8	0.69	22.4	-1.1	44.1	-31	33.7	9.6	45.9
int(block)	1.74	25.3	-2.5	43.3	0.25	22.7	-2.9	42.8	0.83	29.2	3.63	43.9
mean	7.04	21.6	48.2	43.6	15.2	21.3	65.3	42.4	34.2	28.8	96.1	41.8
SSRT-sigma = 25 SSRT-tau = 10												
int(exp)	-2.6	26	-11	45.7	1.6	21.8	-6.5	42.3	-28	31.3	4.35	43.4
int(block)	-2.2	25.5	-6.3	46.8	1.92	22.2	-7	42.5	0.81	26.2	0.65	40.1
mean	3.48	24.1	44.3	43.5	15.9	21	60.9	40.1	36.1	25.6	93.2	37.6
SSRT-sigma = 10 SSRT-tau = 50												
int(exp)	-6.3	24.8	-10	45.2	-1.3	23.3	-2.5	38.4	-27	33.8	2.92	41.7
int(block)	-5.5	25.2	-6.9	46.1	-1.4	23.9	-2	37.5	1.12	28.7	-3.9	37.8
mean	-2	23.3	39.4	41.6	11.6	22.1	58.4	38	33.2	29.6	83.5	36.7
SSRT-sigma = 25 SSRT-tau = 50												
int(exp)	-5.7	22.4	-9.5	41.2	1.32	22.5	-3.4	45.4	-27	32.3	7.4	45.6
int(block)	-5.9	22.1	-7.3	41.2	1.15	22.4	-4	45	1.44	30.1	0.87	43.8
mean	-1.5	21.1	41.6	37.5	13.5	21.2	61.1	41.4	34.5	28.5	89.5	39.9

Table S8: Difference scores when SSRT was analysed using the median RT instead of the nth RT. This table shows the average (avg; in grey) and standard deviation (sd) of the difference scores (estimated SSRT - true SSRT) for each combination of RT-tau * RTsigma. The final column shows the global average of the difference scores and 95% confidence intervals (when the interval does not include 0, the global average is significantly different from zero). For the second set of simulations, we calculated difference scores for each slowing factor separately. Note that in the third set of simulations, we also found that the estimated SSRT correlated with the amount of slowing (r = 0.46, p < .001), which suggests that the estimated SSRT will be an overestimation when subjects slow down.

	RTtau = 50				RTtau = 250				Average difference (lower & upper CI)
	RTsigma = 50		RTsigma = 150		RTsigma = 50		RTsigma = 150		
	avg	sd	avg	sd	avg	sd	avg	sd	
Simulation 1	-2.22	13.95	3.36	29.42	-17.9	40.28	-16.6	42.28	-8.3 (-11.7, -4.9)
Simulation 2									
slow factor = 1	-3.33	13.5	-1.34	31.52	-13.9	36.82	-9.79	45.77	-7.1 (-8.8, -5.4)
slow factor = 1.5	8.45	15.06	15.47	27.3	6.18	36.5	16.14	45.66	11.6 (9.9, 13.2)
slow factor = 2.5	25.7	22.09	41.44	34.95	69.96	38.81	71.89	42.69	52.2 (50.3, 54.2)
Simulation 3	18.43	22.13	29.43	32.63	38.65	49.76	48.35	51.08	33.7 (29.6, 37,9)

Figure S1: To demonstrate the increasing popularity of the stop-signal paradigm, we performed a search in Web of Science (topic = 'stop-signal task'). **(A)** The search results confirmed that the task is currently very popular in different research areas (the research areas correspond to the *Web of Science Categories*; note that papers can belong to multiple categories). **(B)** Since the year 2000, there has been an exponentially increasing number of stop-signal studies. **(C)** We checked the method section of 170 stop-signal articles published since 2010. 128 studies reported how SSD was determined and how SSRT was estimated. The majority of these studies used the tracking method. Approximately half of the tracking studies used the mean RT to estimate the SSRT; the other studies used the integration method or used the median of the RT distribution as an approximation for the nth RT.

Figure S2: Effect of changes in mu, sigma, and tau on the shape of the RT distribution. **A.** When mu increases, the RT distribution shifts to the right (left upper panel vs. left lower panel). **B.** When tau increases, the distribution becomes positively skewed (left upper panel vs. right upper panel); note that when RT-tau is small, the distribution is almost symmetrical. **C.** When sigma increases, the spread of the distribution increases (left upper panel vs. right lower panel).

Figure S3: Results for the mean method (second set of simulations) when only the central estimates (.40 < presp < .60) are included. There was still an effect of RT-tau and slowing (both p-values < .001).

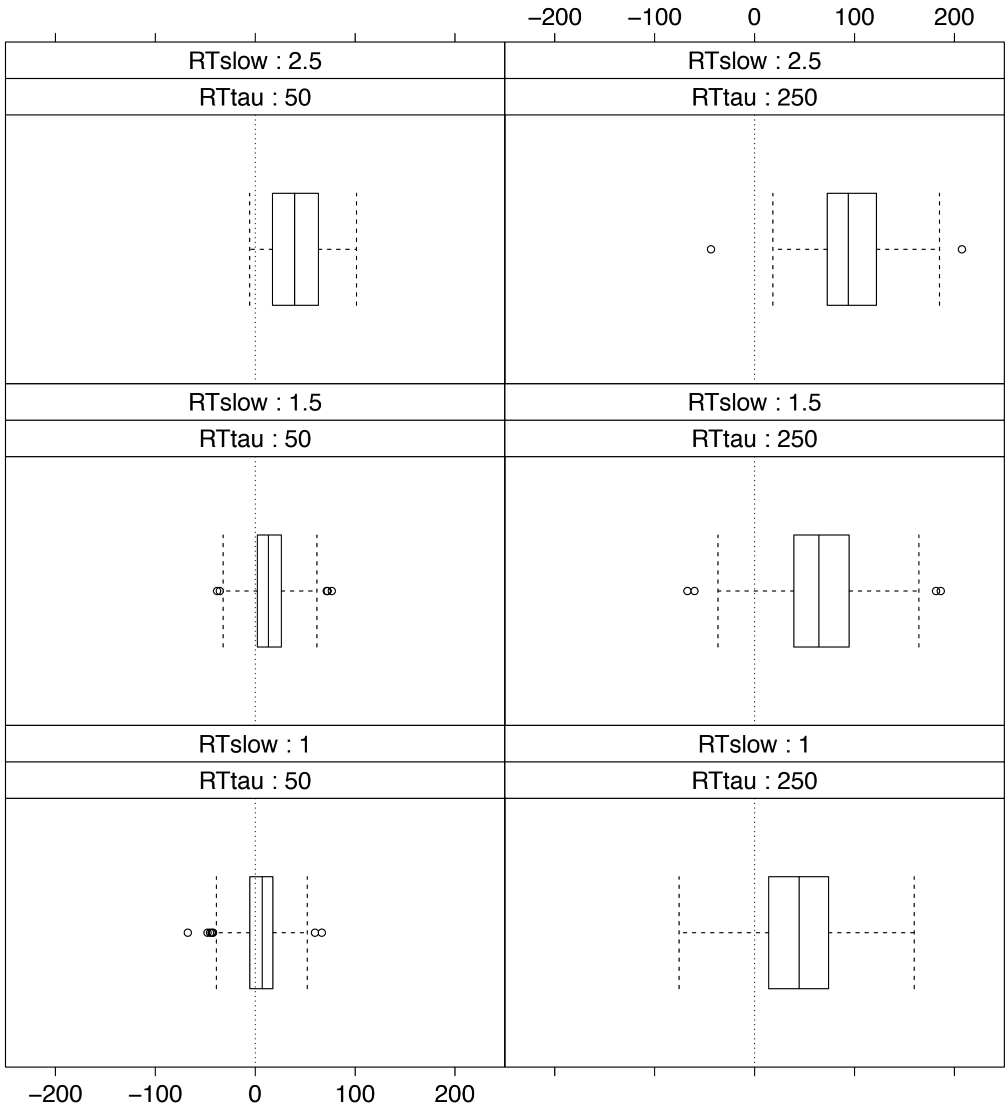
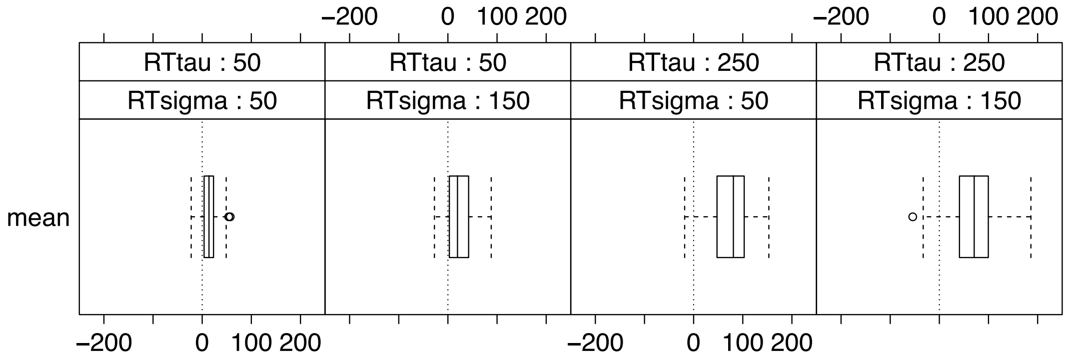


Figure S4: Results for the mean method (third set of simulations) when only the central estimates (.40 < presp < .60) are included. There was still an effect of RT-tau (p < .001), and the degree of slowing still correlated with overestimation.



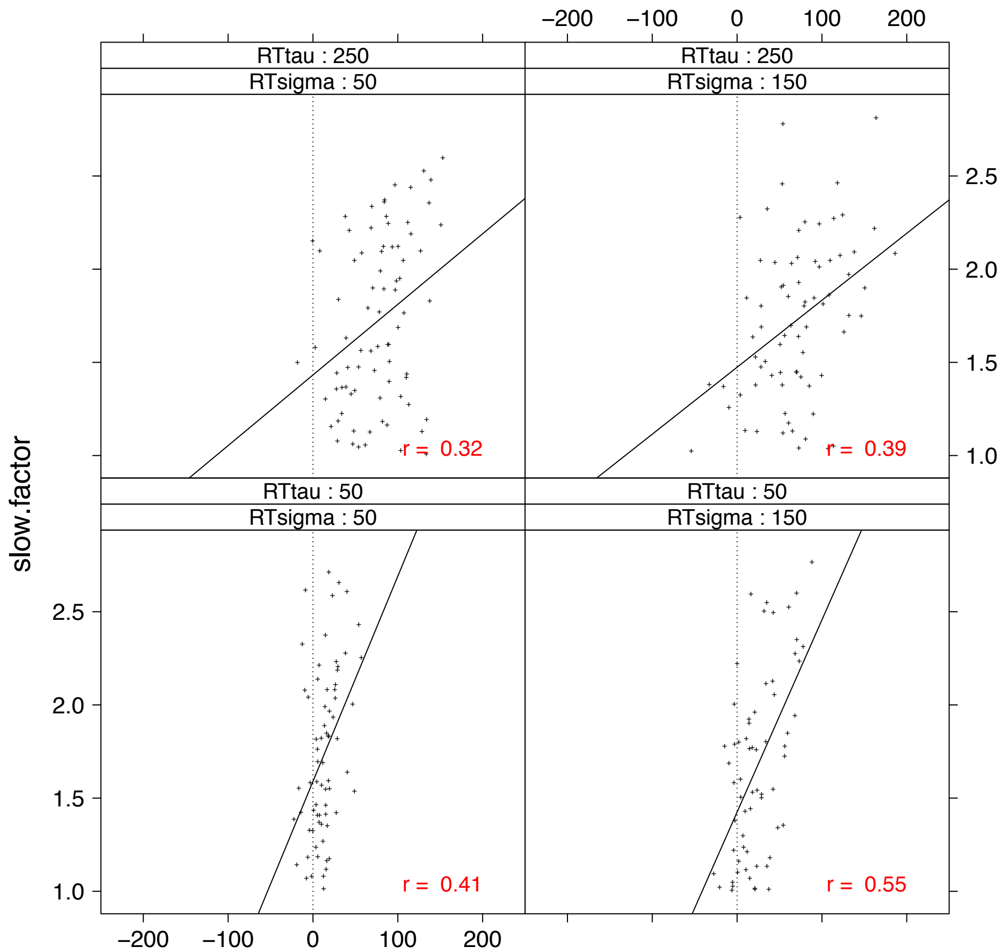


Figure S5: The effect of the number of trials per block for the ‘block-based’ integration method. There were 240 trials per subject. For n=10, we split the data file in 24 blocks of 10 trials; for n=20, we split the data file in 12 blocks of 20 trials; and so on. N=240 corresponds to the estimates using the default integration method. N=60 corresponds to the values for the ‘blocked-based’ integration method reported in the mean manuscript.

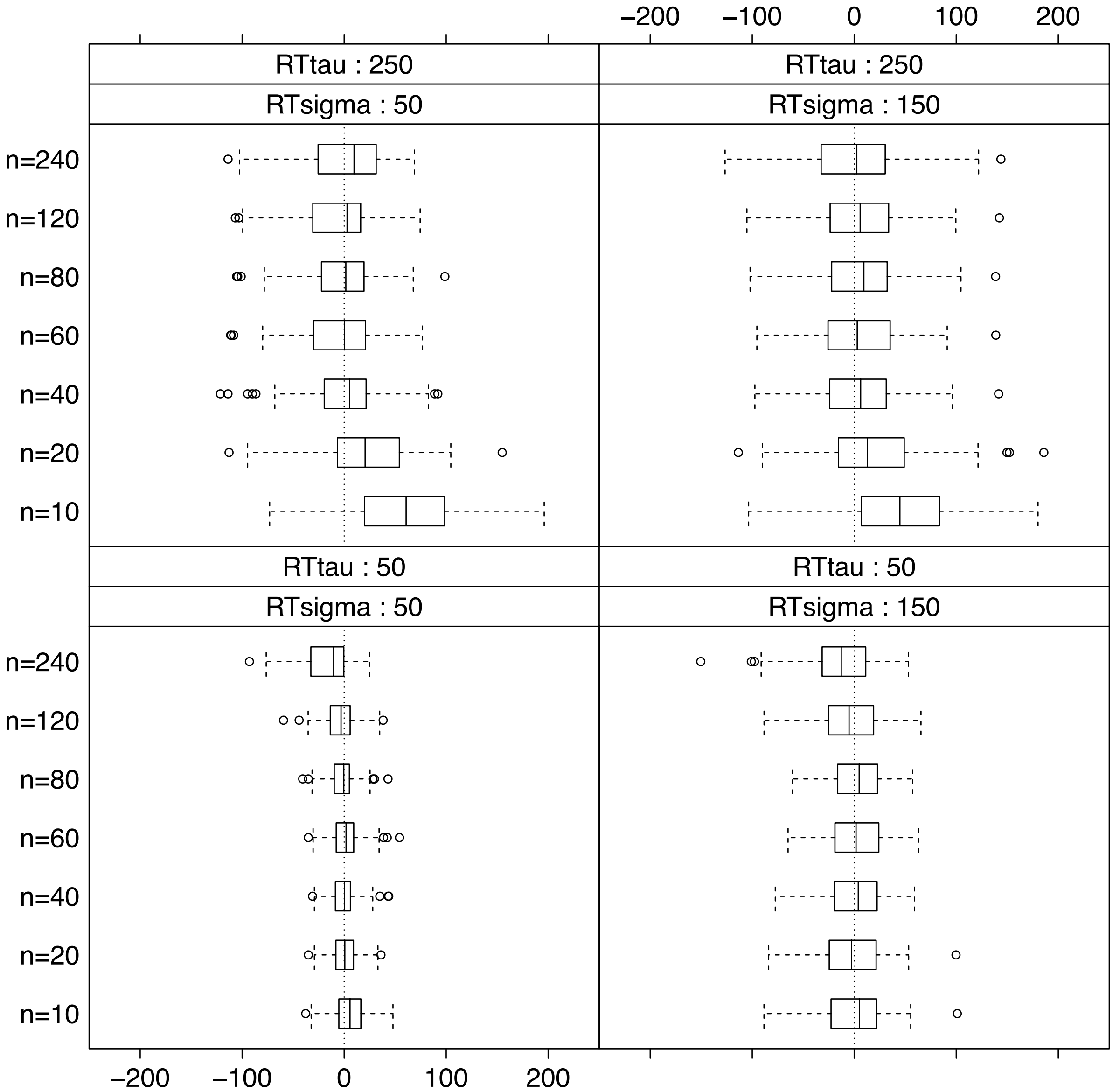


Figure S6: Correlations between difference SSRT(true) and SSRT(estimated) for the different block-based integration methods (see caption Figure S5 for discussion of methods). This figure suggests that when the number of trials is larger than 80 trials, the previously observed negative correlation (see main manuscript) appears again.